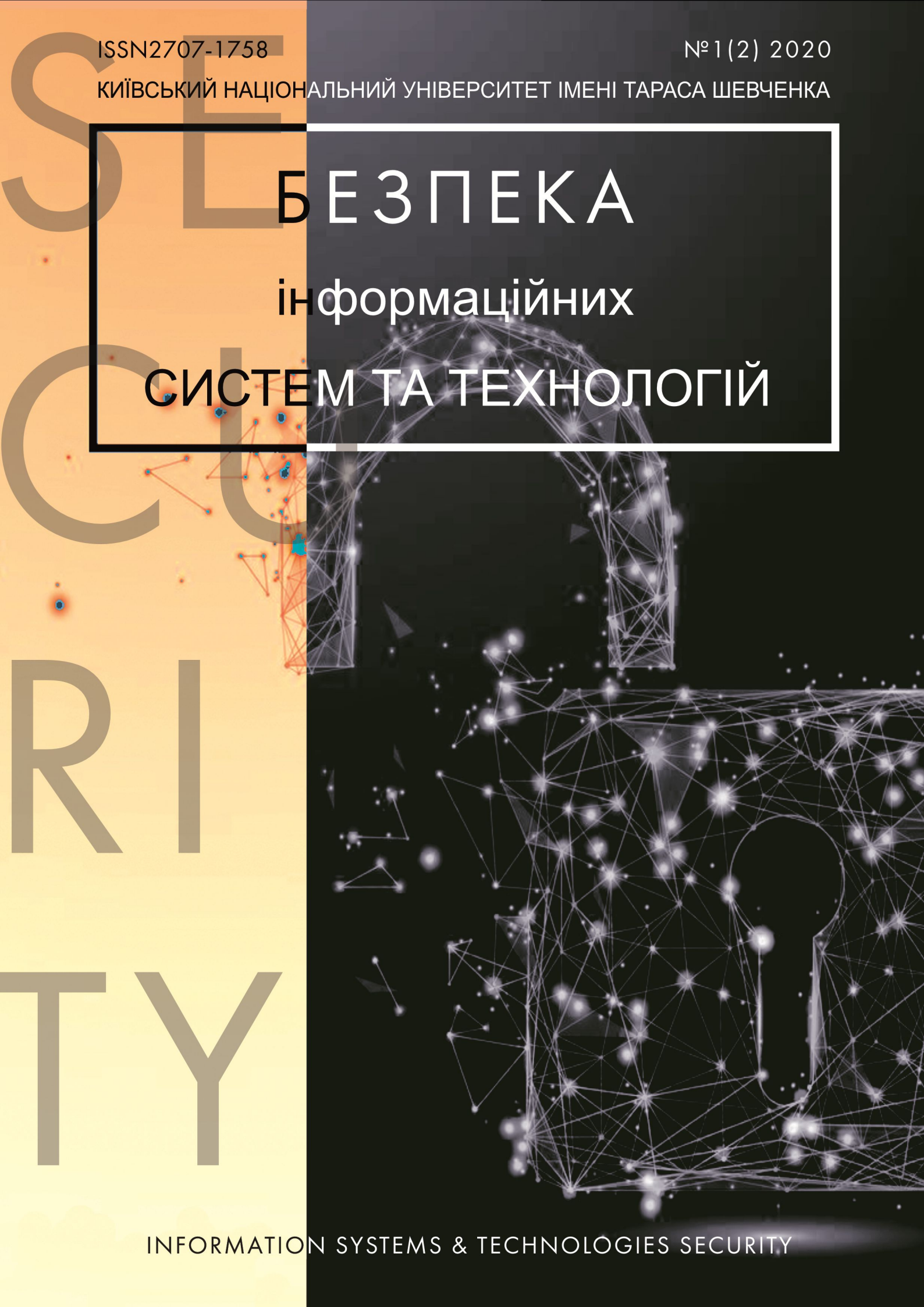
ISSN2707-1758 №1(2) 2020

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

БЕЗПЕКА інформаційних СИСТЕМ ТА ТЕХНОЛОГІЙ



INFORMATION SYSTEMS & TECHNOLOGIES SECURITY

Безпека інформаційних систем і технологій

№ 1 (2) 2020 **НАУКОВИЙ ЖУРНАЛ** **ISSN 2707-1758**

Information systems and technologies security

Інформаційна та кібернетична безпека

Комп'ютерні науки та інформаційні технології

БЕЗПЕКА ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ
№ 1 (2) 2020

Свідоцтво про державну реєстрацію друкованого засобу масової інформації КВ № 23831-13671Р від 15.03.2019

Рекомендовано до друку Вченою радою факультету інформаційних технологій Київського національного університету імені Тараса Шевченка (протокол № __ від 01.02.2020)

Адреса редакційної колегії:
04116, Київ,
вулиця Богдана Гаврилишина, 24

Інформаційно-комунікаційні системи

Журнал «Безпека інформаційних систем і технологій» було створено з ініціативи кафедри кібербезпеки та захисту інформації факультету інформаційних технологій Київського національного університету імені Тараса Шевченка

Підп. до друку 01.02.2020. Формат 60×84/8. Ум. друк. арк. 10,46. Обл.-вид. арк.11,51.
Тираж 100 пр. Замовлення № 456 -2.

Видавець і виготівник
Київський національний університет імені Тараса Шевченка
01033, Київ, вулиця Володимирська, 60

DOI:
UDC 004.56.5

ANALYSIS OF THE PRIMARY TRENDS IN CYBERSECURITY

Oksiiuk Oleksandr [1]
orcid.org/0000-0001-9797-6015

Zerko Andriy [2]
orcid.org/0000-0003-2000-7835

Fesenko Andriy [3]
orcid.org/0000-0001-5154-5324

[1] Kyiv, Taras Shevchenko National University of Kyiv, oksiuk@ukr.net, bos_2006@ukr.net
[2] Kyiv, Taras Shevchenko National University of Kyiv, a.l.zerko@ukr.net
[3]Kiev, Kiev National Taras Shevchenko University, aafesenko88@gmail.com

Article history:
Received by the editor 04.02.2020
Accepted 12.02.2020

Key words:
cybersecurity;
cyberspace;
information protection;
cyber-attacks;
threats

Abstract: Open and free cyberspace increases the freedom of people and social communications, in such conditions it becomes especially important to search for new possibilities of ensuring the state security in view of the formation of a new confrontation field - cyberspace. It is important to analyze the actual problems of information security, actions of the world governments and world organizations for identifying the current state of modern trends in the cybersecurity field. Cybersecurity incidents affect the lives of consumers of informational and many other services, and cyberattacks targeting various objects of electronic communications infrastructure or process management. This article covers in detail the factors that influence the state of cybersecurity in the country, its cyberspace and the protection of information objects. The rapid development of malicious software in the world and the lines of action by famous hacker groups are analyzed. The tendencies of active legislative updates in the cybersecurity field of the world's leading countries, such as creating new structural groups, increasing the number of existing ones and increasing their funding, are identified. The reasons for attackers concentrating their efforts on the search for assets vulnerabilities and the development of a unique multifunctional malware and technologies for unauthorized assets are considered. Structured information about the status of modern trends in the field of cybersecurity and information protection is presented in this article. The situation that has evolved to date with cybercrime requires continuous improvement of cybercrime fighting methods, development of information systems and methods aimed at ensuring the country's cybersecurity. Therefore, the issue of cyberspace security, cybercrime fighting is relevant internationally as well as at the national level and therefore needs further consideration.

Анотація: Відкритий та вільний кіберпростір збільшує свободу людей та соціальну комунікацію, в таких умовах особливого значення набуває пошук нових можливостей забезпечення безпеки держави з огляду на формування нового поля протиборства – кіберпростору. Важливо проаналізувати актуальні проблеми інформаційної безпеки, дії урядів світу та світових організацій для виявлення сучасного стану сучасних тенденцій у сфері кібербезпеки. Інциденти в сфері кібербезпеки позначаються на життєдіяльності споживачів інформаційних і багатьох інших послуг та кібератаки, націлені на різноманітні об'єкти інфраструктури систем електронних комунікацій чи управління технологічними процесами. У цій статті детально розглядаються фактори, що впливають на положення кібербезпеки країни, її кіберпростір та захист інформаційних об'єктів. Проаналізовано стрімкий розвиток шкідливого програмного забезпечення у світі та напрямок дій відомих хакерських угрупувань. Виявлено тенденції активних оновлень законодавства у сфері кібербезпеки провідних країн світу шляхом створення нових структурних груп, збільшення кількості існуючих та збільшення їх фінансування. Розглянуто причини концентрації зусиль зловмисників на пошуку вразливості активів та розробці унікальної багатофункціональної зловмисної програми та технологій для несанкціонованих активів. Структурована інформація про стан сучасних тенденцій у сфері кібербезпеки та захисту інформації представлена в цій

статті. Ситуація, яка склалася на сьогоднішній день з кіберзлочинністю, вимагає постійного удосконалення методів боротьби з кіберзлочинами, розробки інформаційних систем та методів, спрямованих на забезпечення кібербезпеки країни. Отже, питання безпеки кіберпростору, боротьби з кіберзлочинністю є актуальним як на міжнародному рівні, так і на рівні окремої країни, а тому потребує подальшого розгляду.

1. INTRODUCTION

The rapid development of information technology (IT) is transforming the world. An open and free cyberspace increases people`s freedom and social communications, makes relationships between them easier and creates a new global interactive place for many ideas, researches and innovations.

At the same time, access to cyberspace created because of compatible communication systems and electronic communications that are using the Internet or other global data networks, is allowing to gain advantage in the political, economic, military, social, scientific, technological and other spheres.

New technologies and new users will reshape cyber-risks in 2020. The emergence of 5G networks in 2020 will result in substantially broader access for both devices and people. Greater and more convenient broadband at higher speeds will encourage the development and deployment of everything from connected devices and ubiquitous computing to virtual as well as augmented reality and artificial intelligence.

The number of national and municipal laws and regulations addressing cybersecurity – through either new proposals or updates to existing measures – will also increase. While calls for greater alignment across regulatory regimes will continue to grow, there will be little change on this front in the near term owing to the pace and complexities of law-making process and continued debate about the underlying cybersecurity requirements. Continued volatility across the geopolitical landscape will also add to this delay and cybersecurity threats will evolve to exploit the ever-changing environment.

However, the advantages of modern digital world and the growth of information technology led to the emergence of new national and international security threats. Besides natural incidents, there is also an increase in the number and power of threats that benefit individual states, groups or individuals. Depending on their intentions and motivation, these threats can include:

- Collection and theft of information resources for further use or sale;
- Industrial espionage and diversion;
- Hacking networks in order to gain access to their information and use them to implement cyber-attacks;
- Violation of network processes with cyber-attacks or malicious software;
- Attacking or exploiting important infrastructure objects;
- Personal data theft.

2. ACTIONS IN THE CYBERSPACE OF THE WORLD`S COUNTRIES

The leading countries of the world are focusing on the creation and improvement of legislations in the cybersecurity sphere in order to increase the level of digital information protection.

At the same time, the continued aggression in cyberspace from Russian Federation was the reason for several incidents and other changes in the protected external and internal environment of the world`s leading countries.

Let us analyze some actual examples of actions in cyberspace of the countries in the world that deserve attention:

1. USA.

In December 18, 2017, the updated US National Security Strategy has been published; it included directives that might increase the level of country’s protection against cyber threats.

2. EU.

The Research Center of the European Parliament reported that one of the main directions of EU policy in 2018 is to increase the cybersecurity because of the threat from Russia.

3. Great Britain.

The British Cabinet has produced a transitional national science and technology strategy in the cyber security sphere aimed at ensuring the UK`s technological capacities’ resistance to cyber-threats.

4. France.

In February, 2018, the Government of France has approved the Strategic Review of Cybersecurity (Review) containing the country`s strategic tasks in the sphere of digital and information technology taking into consideration the current geopolitical state, as well as improving the new instruments of warfare.

In order to adhere to the standard of the country`s cybersecurity, a consecutive increase of personnel in the Cyber Command (CYBERCOM) of the General Staff of the Armed Forces of France is in order: to 3200 people by 2019 and to 4K people by 2025.

5. Germany.

The government of Germany is considering the possibility of introducing corrections to the country`s constitution regarding the problem of hacker`s attacks aimed at private computer networks. The Germany's Ministry of International Affairs delegate stated that the relevant reforms were to be completed in 2018. The experts think that the possible methods of protection from hackers include the probability of disabling servers used by attackers in offensive.

6. Japan.

During the second quarter of 2018, the Japanese government intended to develop criteria for the danger of cyberattacks on critical infrastructure of the country (this includes railways, electricity, financial institutions and so on). This will help the government use appropriate measures to fight crisis.

The concrete ways of government`s response to a cyberattack will change depending on the six-level scale of threats.

7. Georgia.

At the end of November 2017, the Government of Georgia has announced about additional $35 million to fund the creation of the Innovations Center and Cybersecurity Training.

The results of the analysis show the tendency of active legislative updates in the cybersecurity sphere of the leading countries of the world, the increase of cybersecurity level achieved by the creation of new structural divisions and increased amount of funding, mainly because of the threat from Russia.

3. ACTIONS OF INTERNATIONAL ORGANIZATIONS AND LEADING COMPANIES IN THE CYBERSECURITY SPHERE

International organizations and leading companies in the cybersecurity sphere are actively exploring the cyberspace usage to increase the protection level of the vital interests of citizens, society and state, as well as to protect their assets by creating cyber security systems, cyber security centers, cyber defense hardware and software complexes.

The following actual examples of international organizations' and leading cyber-security companies' actions are given for the analysis:

UN Secretary General Antonio Guterres at The Munich Security Conference during his speech said that our world needs determination in responding to cyber-attacks on the international field.

He proposed to start a discussion on that matter at the UN General Assembly residency.

As part of reforming the NATO command structure, they are reevaluating cyber capabilities of the Alliance for strategic cybersecurity.

On February 14 – 15, 2018 at the meeting at the NATO Headquarters in Brussels, the North Atlantic Council at the level of defense ministers of NATO`s member states decided to create a NATO Center for Cyber Operations (Cyber Operations Center). This Center will be integrated with Headquarters of the NATO Operations Command (SHAPE) in Mons, Belgium.

The Press Service of the Cybersecurity Center of the NATO Center (CCD COE), based in Estonia reported on January 30, 2018 that the Center is starting to identify the need for training new specialists in the field of cybersecurity and coordinating their education.

Experts from American Network Protection "FireEye" have discovered a new Triton (Trisis) malware family with status of rare malicious electronic medium, which runs often in the Middle East.

The main object of this wrecker is to prevent the work of industrial defense systems that are protecting workers' lives.

Hackers can use Triton in the system to create a dangerous situation, even dealing physical damage.

The cybersecurity researchers from the American company Cisco Talos were able to detect the SPO that was used to attack the sites of the 2018 Olympics.

"During the attack, hackers interrupted the work of digital interactive television at the main Press Center of the Winter Olympics in Pyeongchang, South Korea, and caused a crash in the Wi-Fi network at the stadium".

Software called "Olympic Destroyer" is a malicious program for operation systems based on Windows that can complete a lot of hacker`s tasks, especially infecting a device with multiple files that steal stored passwords from Internet Explorer, Firefox and Chrome browsers, as well as computer system passwords.

The National Institute of Standards and Technology of the USA (NIST) has published a document project "The Status of International Cybersecurity Standardization for IoT", that can help in developing security standards for IoT.

NIST proposes to divide IoT into five functional areas:

- Device connection;
- IoT of the consumer class;
- Medical equipment and devices used in health care area;
- "Smart" homes;
- "Smart" production (including ACS);

Standards should be developed for each area, taking into account its specifics.

Taking into consideration all of the above and according to the results of the analysis, it can be

noted that international organizations and leading companies in the field of cybersecurity are taking active measures protecting the vital interests of citizens, society and state in cyberspace.

4. ACTUAL THREATS IN CYBERSPACE AND ASSETS VULNERABILITY

Attackers are developing new computer viruses and malware to gain access to assets owned by a person, organization or state.

The following cyber threats and vulnerable assets that deserve attention are given for analysis.

GandCrab malware

Experts of the company "Malwarebytes" reported about new ransomware that is spreading in a very unusual way – with the help of two sets of exploits.

The activity of ransomware "GandCrab" was first recorded on January 26, 2018. This malware was distributed by two separate exploits – RIG and GrandSoft.

The RIG contains exploits for vulnerabilities in Internet Explorer and Flash Player to execute JavaScript, Flash or VBscript attacks. The RIG is probably distributing GandCrab through malicious advertisements on compromised sites.

The second set of GandCrab`s exploits, GandSoft, exploits vulnerabilities in the Java Runtime Environment, which allows to remotely execute the code.

After installing it into the system, GandCrab works like most of ransomware – encrypting files that are stored on the computer using RSA algorithm, adding GDCB extension to them and then demanding a payment for the recovery tool. However, unlike most cryptographers, GandCrab requires payment not in Bitcoin, but in the cryptocurrency Dash. This fact is another proof that cybercriminals are ceasing the use of Bitcoin little by little in favor of other cryptocurrencies.

Coinhive malware

On February 11, 2018, thousands of government websites in the UK and Australia were attacked by the Coinhive malware, which exploited the potential of infected computers to demand cryptocurrency. With this, the National Cybersecurity Center of the United Kingdom (NCSC) published a tutorial that describes attacks using third-party JavaScript archives and gives advice for website administrators and community agents to counter the attack.

The experts of the center note that hackers are focusing on discrediting additional computer systems, since this allows them to initiate a much more complicated criminal operation.

Dridex malware

Security researchers from Forcepoint reported about a new fishing campaign, in which attackers use discredited FTP resources to spread the banking trojan, Dridex.

This trojan spreads by phishing emails and by deceptive offers to download and execute malicious macros that are hidden in Microsoft Office documents. Once in the system, Dridex steals credentials for online banking, which attackers can then use for stealing money from the victim's bank account.

This campaign has begun on January 17, 2018. Phishing letters were sent mainly to top level domains, such as .com, .fr, and .co.uk. The highest number of victims was recorded in France, the United Kingdom and Australia.

According to experts, the Dridex campaign uses two types of documents: an XLS file with malicious macros that downloads trojan on the device, and a DOC file that exploits a vulnerability in Dynamic Data Exchange (DDE) to execute attacker`s commands.

Cisco ASA firewall software vulnerability (CVE-2018-0101)

Cisco Company informed their clients without going into details that the attackers are actively exploiting the critical vulnerability CVE-2018-0101, which affects Cisco Adaptive Security Appliance (ASA) an operating system run by the Cisco ASA family of firewalls. This vulnerability allows a remote unauthorized hacker to execute any code or evoke a denial of service.

Hacker groups continue to actively search for vulnerabilities in assets and control systems in order to realize new cyber threats.

The actions of hacker groups that deserve attention can be determined:

A cyber espionage group Fancy Bear (APT28)

The cybersecurity researchers from ESET have reported about appearance of a new Xagent malware feature – one of the main tools in the Fancy Bear hacker group (APT28).

They also noted that the main objects of this group`s attack are still government agencies and embassies around the world, and especially in Eastern Europe.

The "Talos Group" has exposed a new criminal cyberattack from hacker group APT28. The document created by this group is the leaflet of the CyCon U.S. Conference, organized with the Cyber Institute of the US Military Academy and the CCD COE (7 – 8 November 2017, Washington). The leaflet does not contain malicious software, but activates the execution of a malicious code inside a fake document when it`s opened.

Hackers group Dark Caracal

This hacker group, which is allegedly related with the Lebanese government, has stolen hundreds of gigabytes of information from thousands of

victims around the world using only phishing emails and simple malware. This is stated in the common report of the human rights organization Electronic Frontier Foundation and cybersecurity company Lookout.

5. CYBERCRIME

Cybercriminals continue to realize socially dangerous attacks in cyberspace.

The Council of Economic Advisers under the President of the United States created a list of "hacker states", which included:

- Russia;
- China;
- Iran;
- The DPRK.

This was reported in the 62-page report of the Council "Cost of Malicious Cyber Activity to the U.S. Economy".

In the report, economists identified six categories of cybercriminals depending on their targets:

The first category includes Russia, Iran, China and the DPRK, whose state hackers have political, economic, technological, and military targets.

The second category includes corporations who wish to acquire their competitors' industrial secrets and intellectual property. Many of them are funded by the state.

The third category is "Hacktivists", whose activity in the cyberspace is a protest action. Their actions have a propagandistic nature, and they cause losses for organizations for ideological reasons.

The fourth category includes organized cybercrime groups that carry out targeted attacks for the purposes of acquiring profit.

Next are the "Opportunists" – unprofessional hackers who want popularity. In their attacks, they use widely available techniques and codes.

The last category consists of insiders – present or former employees of companies, prompted by revenge or profit.

Considering all of the above, it can be noted that cybercriminals of different categories are actively searching for vulnerabilities in assets and management systems to reach their political, economic, technological, military and other objectives.

6. Conclusion

According to the results of the analysis, the main trends in the cybersecurity sphere are the following:

The leading countries of the world and foreign companies are actively researching the aggressive actions from RF in cyberspace and note that such cyberattacks lead to losses in the political, economic, technological and military spheres (USA, UK, Australia, Canada and New Zealand are blaming Russia for organizing a cyberattack using NotPetya malware in the summer of 2017. Estonia has published an annual report that analyzes Russia`s cyberbullying activities).

The leading countries of the world continue the process of legislative consolidation in the sphere of cybersecurity (France approved the Strategic Review on cybersecurity; in Poland the bill about nation system of cybersecurity was developed).

The leading countries increase their operational capabilities in order to raise the level of cybersecurity, create new structural units and increase the number of existing ones (NATO has provided the yearly cyber training "Crossed Swords", the USA`s Ministry of energy created a new unit for cybersecurity, energy security and emergency responding. France plans to increase the number of those units up to 4,000 by 2025. The Australian Defense Forces have created a cyber-command).

The leading countries of the world are concentrating attention on the interactions with state bodies and with private organizations to increase the protection level of critical infrastructure objects (France is planning a full engagement of private companies to the common process of the securing cyberspace).

Attackers are focused on finding vulnerabilities in assets (control systems) and developing multifunctional malware with unique properties and technologies for unauthorized access to assets (attackers focus on discrediting additional computer systems, allowing them to gain access to the main asset in the future; the ransomware from GandCrab is spread by unusual way – with two sets of exploits).

Russian-oriented hacker groups continue to carry out cyberattacks on the assets of the Ukrainian segment of the Internet and on assets of the Ukraine-oriented countries of the world, foreign companies, institutions, organizations. Notably, for some of the hacker groups the targets of their attacks are the same as actions of the Russian government policy (for example, interference by hacker groups ATP28 and ATR29 in the US election process).

The leading countries of the world are researching to determine the legal status of cryptocurrencies and legal regulation of transactions with them. At the same time, attackers are actively searching for vulnerabilities in assets and control systems to use their resources to get the cryptocurrency or steal it.

Cyber defense is the only thing that can prevent the loss of information and the interference of some countries into the security of others. The analysis identified the main areas of protection against cyber threats, protection of sovereignty of cyberspace and

national security in the leading countries of the world.

7. References

[1] Cybersecurity Trends Report 2019 // elevenpaths. – 2019. – URL: https://www.elevenpaths.com/ cybersecurity- trends-report-2019/index.html. The Hacker News — Online Cyber Security News & Analysis. URL: https://thehackernews.com/.

[2] North Atlantic Treaty Organization official website. URL: https://www.nato.int/.

[3] United Nations official website. URL: http://www.un.org/.

[4] FireEye, Threat Research Blog. URL: https://www.fireeye.com/blog/threat-research.html.

[5] Cyber Security Essentials. James Graham, Ryan Olson, Rick Howard. N.Y.: CRC Press, 2010.

[6] Michelle Nichols. North Korea took $2 billion in cyberattacks to fund weapons program: U.N. report / Michelle Nichols // reuters – URL: https://www.reuters.com/article/us-northkorea-cyber-un/north-korea-took-2-billion-in-cyberattacks-to-fund-weapons-program-u-n-report-idUSKCN1UV1ZX.

[7] State of Cybersecurity Report 2018 // wipro. – 2018. – URL: https://www.wipro.com/ content/dam/nexus /en/service-lines/ applications/latest-thinking/state-of-cybersecurity-report-2018.pdf..

[8] State of Cybersecurity 2019, Part 1: Current Trends in Workforce Development // isaca. – 2019. – URL: https://www.isaca.org/ bookstore/bookstore-wht_ papers-digital/whpsc191..

[9] Bernard Marr. The 5 Biggest Cybersecurity Trends In 2020 Everyone Should Know About / Bernard Marr // www.forbes.com. – 2020. – URL: https://www.forbes.com/sites/bernardmarr/2020/01/10/the-5-biggest-cybersecurity-trends-in-2020-everyone -should-know-about/#4e86a2857ecc

[10] Cybersecurity Trends for 2019 // cisecurity.org. – 2018. – URL: https://www.cisecurity.org/blog/ cybersecurity-trends-for-2019/.

[11] Elias chachak. Top 10 Countries Best Prepared Against Cyber Attacks / ELIAS CHACHAK // cyberdb.co. – 2019. – URL: https://www.cyberdb.co/top-10-countries-best-prepared-cyber-attacks/.

Oksiiuk Oleksandr
Dr.Sc., Professor, Head of Cybersecurity and Information Protection Department, Faculty of Information Technology, Taras Shevchenko National University of Kyiv.
Interests: theory and practice of developing automated systems and conducting expert assessments, models and methods of building integrated information security.

Zerko Andriy
Graduate student of Cybersecurity and Information Protection Department, Faculty of Information Technology, Taras Shevchenko National University of Kyiv.
Interests: information security, information technology

Fesenko Andriy
PhD., Assistant Professor of Cybersecurity and Information Protection Department, Faculty of Information Technology, Kiev National Taras Shevchenko University
Interests: information security, cybersecurity

DOI:
УДК 519.816

ВИЗНАЧЕННЯ ПРІОРИТЕТНОСТІ ЗАХОДІВ КІБЕРБЕЗПЕКИ ПРИ НЕПОВНИХ ЕКСПЕРТНИХ РАНЖУВАННЯХ

Гнатієнко Григорій[1]
orcid.org/0000-0003-3224-40614

Тмєнова Наталія[2]
orcid.org/0000-0002-0247-5400

[1] Україна, м. Київ, Київський національний університет імені Тараса Шевченка, g.gna5@ukr.net
[2] Україна, м. Київ, Київський національний університет імені Тараса Шевченка, tmyenovox@gmail.com

Історія статті:
Надійшла до редакції
13.02.2020
Прийнято
23.02.2020

Ключові слова:
організаційна система; інформаційна безпека; метрика; відстань; медіана; групове упорядкування об'єктів; неповне експертне ранжування.

Анотація. Якісне функціонування системи інформаційної безпеки та вирішення проблем, що виникають при захисті інформації, наразі є актуальним напрямом в різних сферах людської життєдіяльності. Успішний кіберзахист полягає у створенні та реалізації багаторівневої системи заходів, які охоплюють різні аспекти, що мають комплексно взаємодіяти та взаємодоповнювати один одного. Ці заходи мають різний характер, і їх пріоритетність з точки зору різних служб організації може суттєво відрізнятися, тому розробку послідовності виконання заходів кібербезпеки логічно формалізувати у класі задач групового вибору. В роботі пропонується гнучкий математичний апарат для моделювання задач інформаційної безпеки та адекватного застосування аналізу думок колективу експертів на практиці. Описано підхід до знаходження результуючого ранжування пріоритетності заходів як розв'язку задачі багатокритеріальної оптимізації, де послідовність виконання заходів може передбачати взаємодію виконавців і вимагати регламентування послідовності дій усіх елементів та підсистем організаційної системи. Такий підхід дозволяє об'єднати різні за складом та пріоритетністю заходи інформаційної безпеки, запропоновані експертами різних підрозділів; знаходити компромісне рішення для різнорідної групи експертів; не порушувати задані переваги жодного експерта при обчисленні компромісного ранжування заходів кібербезпеки. Запропонований підхід може бути корисним при розробці заходів забезпечення належного рівня кібербезпеки та сприятливим при розробці і впровадженні процедур швидкого реагування на загрози, а також бути незамінним при комплексній побудові системи безпеки організації або її удосконаленні та містити елементи навчання, узгодження, координації, комплексності учасників експертної групи, які є керівниками підрозділів єдиної організаційної системи.

Abstract. High-quality functioning of the information security system and solving problems that arise in the information protection, is currently a topical trend in various areas of human life. Successful cyber protection consist in creating and implementing a multi-level system of measures that cover various aspects with complex interact and complement each other. These measures have a different nature, and their priorities may differ significantly in terms of different services of the organization, so it is logical to formalize the sequence of cybersecurity implementation in a class of group choice tasks. The paper proposes a flexible mathematical apparatus for modeling information security problems and adequate application of the opinion analysis of experts' team in practice. The approach to finding the resultant ranking of measures priority is described as a solution to the problem of multicriteria optimization, where the sequence of measures implementation may involve the interaction of performers and require regulation of the actions sequence of all elements and subsystems of the organizational system. This approach allows to combine different information security measures proposed by the experts of various departments; to find a compromise solution for a diverse group of experts; not to violate any expert's preferences under calculating the compromise ranking of cyber security measures. The proposed approach can be useful in developing appropriate cybersecurity measures and favorable in developing and implementing of rapid response procedures to threats, as well as it can be indispensable in the overall building or improving organization security system and it can contain elements of training, coordination, and complexity of expert team members, who are the heads of units of a single organizational system.

1. ВСТУП

Важливість проблеми підвищення рівня інформаційної безпеки зростає разом із інтенсивним розвитком інформаційних технологій та посиленням конкуренції у сучасному світі. Забезпечення функціонування безпекової складової будь-якої сучасної інформаційної системи є актуальною у міждержавних відносинах, у будь-якій галузі народного господарства, у різноманітних напрямках людської життєдіяльності тощо. Вирішення проблем, що виникають при захисті інформації, є досить складним процесом, який базується на системному підході, що застосовується при створенні комплексної системи захисту інформації. Якість функціонування системи інформаційної безпеки залежить від якості функціонування елементів системи, тому моделювання інтегрального рівня безпеки інформаційної системи є актуальним напрямком досліджень.

Наявність значної кількості способів негативного впливу на системи інформаційної безпеки потребує вдосконалення заходів кібербезпеки. Успішний кіберзахист полягає у створенні та реалізації багаторівневої системи заходів, які охоплюють такі основні аспекти: організаційний, персонал організації, комп'ютери, локальні мережі, програмне забезпечення, бази даних та різноманітні сховища даних, які слід комплексно убезпечити. Для вдалого захисту від загроз усі ці аспекти кіберзахисту мають комплексно взаємодіяти та взаємодоповнювати один одного.

2. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Оцінка захищеності системи потрібна для створення механізму та умов оперативного реагування на загрози інформаційній безпеці та прояви негативних тенденцій функціонування системи. Для цього слід застосовувати комплекс заходів та протидій. Найбільш розповсюдженими методиками в галузі захищеності інформаційних систем є три основних метода захищеності системи: формальний, статистичний та класифікаційний [1, 2]. Перспективним підходом є підвищення об'єктивності і комплексності оцінки засобів захисту інформації на базі формалізації експертних даних.

Задачі групового (колективного, колегіального, демократичного) упорядкування об'єктів дозволяють моделювати практичні ситуації у різних предметних областях. Особливістю загальноприйнятих постановок таких задач є зафіксована для усіх членів експертної групи множина об'єктів, що у багатьох випадках суттєво знижує якість експертизи. Наразі існують дослідження, в яких допускається можливість задання неповних ранжувань у задачах колективного ранжування [3], але алгебраїчні методи обчислення медіани до таких задач на сьогодні ще не застосовувалися.

Заходи щодо інформаційної безпеки мають різнопланновий характер, і їх пріоритетність, з точки зору різних служб організації, може суттєво відрізнятися. Тому розробку послідовності виконання цих заходів логічно формалізувати у класі задач групового вибору, орієнтованого на визначення кількох рівноцінних розв'язків або упорядкування кількох виділених чи усіх заданих альтернатив [4]. Гнучкий підхід до упорядкування об'єктів, коли кожен експерт може пропонувати часткове упорядкування вибраної ним підмножини з усієї множини альтернатив, та пошук повного результуючого ранжування об'єктів алгебраїчними методами дослідниками не пропонувалися. Адже при цьому з'являється можливість враховувати ситуацію, коли деякі експерти можуть не знати про особливості функціонування окремих елементів з множини заходів безпеки, що стосуються загальної проблеми інформаційної безпеки усієї організаційної системи.

3. МЕТА ДОСЛІДЖЕННЯ

У дослідженні пропонується гнучкий математичний апарат для моделювання задач інформаційної безпеки та адекватного застосування аналізу думок колективу експертів на практиці. Мета дослідження вимагає розробки алгоритму визначення пріоритетності заходів інформаційної безпеки та забезпечення незалежності визначення та задання індивідуальних переваг експертів для успішного застосування комплексних заходів безпеки. Зокрема, експерт не повинен залежати від множини заходів, які можуть бути ним пропоновані для розв'язання деякої загальної проблеми. Це дозволяє апріорно підвищити професійність групового розв'язку, оскільки експерти мають можливість робити свої судження у межах своєї високої професійності на

множині критеріїв, де вони є безсумнівними спеціалістами, а не у жорстких рамках визначення пріоритетів на усій множині заходів безпеки, заданої для усіх членів експертної групи.

4. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

На сьогодні традиційно виділяються такі види безпеки: економічна, юридична, фізична та інформаційна. Водночас до інформаційних загроз традиційно відносяться такі:
– розголошення конфіденційної інформації;
– витік конфіденційної інформації через технічні засоби;
– несанкціонований доступ до охоронюваних відомостей;
– знищення або несанкціоновані зміни інформації;
– блокування або руйнування технічних засобів прийому, передачі, обробки та зберігання інформації.

Усі зазначені вище аспекти інформаційної безпеки можуть належати до зони відповідальності різних служб та підрозділів організації. Тому природно, що керівники таких підрозділів не можуть бути однаково компетентними у всіх питаннях безпеки. Комплексність є одним із основних принципів організації та функціонування системи інформаційної безпеки. Складність побудови політики комплексної інформаційної безпеки обумовлена тим, що інформаційна безпека досягається комплексом організаційних, технічних та інших заходів.

Слід зазначити, що часто саме послідовність виконання кроків при прийнятті рішення є найсуттєвішим чинником як захисту, так і забезпечення надійності та усунення загроз.

Нехай розв'язується задача визначення послідовності впровадження заходів інформаційної безпеки деякої складної системи. Для цього різними підрозділами здійснюється підготовка та обґрунтування рішень щодо послідовності або пріоритетності заходів комплексної безпеки великої та багатопрофільної організації. Нехай колективом експертів задано k часткових порядків (неповних ранжувань) $R^i, i = 1,...,k.$

Представники різних підрозділів та служб, які входять до експертної групи, у межах своєї компетенції та уявлення про важливість заходів надають індивідуальні (частково впорядковані) пропозиції щодо послідовності запропонованих ними заходів. Серед членів експертної групи часто не співпадають множина запропонованих заходів, компетентність, інтереси, пріоритети, акценти, аспекти, уявлення про ідеалізацію моделі тощо. Тому на першому етапі розв'язання задачі слід використовувати підхід, що ґрунтується на об'єднанні множин та на введенні додаткових евристик.

Для оцінки загроз інформаційній безпеці традиційно застосовуються теорія надійності, математична статистика, теорія ймовірності, експертні технології прийняття рішень. У свою чергу, методи обробки експертної інформації поділяються на три основні групи: статистичні методи, методи шкалювання та алгебраїчні методи [5].

Для розв'язання задачі, яку сформульовано у цій роботі, найбільш прийнятним є алгебраїчний підхід. Суть алгебраїчних методів полягає в тому, що на множині допустимих оцінок задається відстань і результуюча оцінка визначається як така, відстань якої до оцінок експертів за певним вибраним критерієм є мінімальною.

При такій постановці задачі проблематика полягає у тому, що у задачах колективного ранжування об'єктів, як правило, розглядаються ситуації, коли експерти задають індивідуальні ранжування у фіксованому просторі із заданою кількістю об'єктів. При класичному алгебраїчному підході кожен експерт має задати свої переваги на множині усіх без виключення об'єктів, а результуюче (компромісне, колективне, групове, агреговане, інтегроване, інтегральне тощо) ранжування відшукується на множині усіх можливих ранжувань заданої множини об'єктів. Знайдене таким чином результуюче ранжування вважається узагальненою думкою групи експертів, які беруть участь у розв'язанні задачі лінійного (повного) упорядкування заданої множини об'єктів.

Для розробки комплексної системи взаємодії між елементами системи будемо застосовувати формалізм бінарних відношень, який зручно виражати орієнтованим графом або у вигляді матриці попарних порівнянь. При наявності значної кількості об'єктів порівняння, у нашому випадку – заходів безпеки, побудова такої матриці є трудомістким процесом. Тому на практиці для зменшення трудових витрат експертів нерідко крім процедури попарних порівнянь об'єктів застосовуються множинні порівняння. Результати множинних порівнянь складаються, як правило, з повних відношень, заданих між 3-5 об'єктами. Але досягнення повноти, непротирічливості, транзитивності, узгодженості тощо та задоволення вимог щодо цих показників повною мірою вимагає великих зусиль та витрат часу від експертів. Для

розв'язання задачі створення комплексу заходів безпеки будемо використовувати неповні ранжування – такі упорядкування об'єктів, кількість яких є більшою п'яти, але залишається меншою від усієї множини об'єктів, які слід упорядкувати.

Неповним ранжуванням $R^{(н)}$ множини A назвемо таке ранжування, у якому не для усіх n об'єктів множини A визначено відношення порядку. При цьому $A_i \subset A, \forall i = 1,...,k,$ відношення включення є строгим. Тобто в множині A існує хоча б один об'єкт, якого немає у підмножині $A_i, i = 1,...,k.$

На практиці часто виникають ситуації, коли експерти не в змозі здійснити достовірне ранжування об'єктів з усієї множини. Вимагати від них виконання цієї задачі означає свідомо створювати заздалегідь недостовірні початкові дані. Тому у цій роботі пропонується розглядати ситуації, коли кожен з k експертів має можливість установити частковий порядок на тій підмножині $A_i, i \in I = \{1,...,k\},$ множини об'єктів $A,\ A_i \subset A,\ i \in I,$ де він є компетентним, та які входять до його сфери впливу чи відповідальності. Тобто, задача починається з того, що k експертів задають часткові порядки $R^i = \{a_{j_1} \succ a_{j_2} \succ ... \succ a_{j_{ki}}\},$ $i \in I = \{1,...,k\},\quad k_i < k, i \in I,$ на вибраних ними підмножинах об'єктів $A_i \subset A, i \in I.$ Необхідно знайти деяке результуюче (агреговане, колективне) упорядкування n задач $R^* = (a_{i_1},...,a_{i_n}),$ $i_j \in I = \{1,...,n\},\ j \in I,$ яке побудоване за логікою, що характеризує процеси функціонування та забезпечення захисту деякої інформаційної системи. Упорядкування побудоване на основі індивідуальних упорядкувань задач, які виконуються k елементами системи $R^i = (a_{1_i},...,a_{n_i}),\ i \in J = \{1,...,k\},$ де $n_i -$ кількість задач в індивідуальному упорядкуванні, які виконуються $i -$ м елементом системи, $i \in J.$

На першому етапі розв'язання задачі здійснюється об'єднання підмножин заданих експертами заходів безпеки у єдину множину $A,$ до якої входять усі запропоновані експертами заходи $a_i \in A, i = 1,...,n,$

$$A = \bigcup_{i=1}^{k} A_i. \qquad (1)$$

Зрозуміло, що при цьому допускаються різні варіанти співвідношень між підмножинами: $A_{i_1} \cup A_{i_2} = \varnothing,\quad A_{i_1} \cup A_{i_2} \neq \varnothing,\quad A_{i_1} = A_{i_2},$ $i_1, i_2 \in J.$

Множина усіх заданих експертами заходів безпеки виду (1) є областю допустимих розв'язків для визначення результуючого ранжування заходів безпеки організації $R^*.$

Для визначення відстаней між ранжуваннями об'єктів використовують:

– метрику Кука неспівпадання рангів об'єктів у індивідуальних ранжуваннях

$$d(R^j, R^l) = \sum_{i \in I} \left| r_i^j - r_i^l \right|, \qquad (2)$$

де r_i^l – ранг i-го об'єкта у ранжуванні l-го експерта $R^l, l \in L,\ 1 \le r_i^l \le n,$

– метрику Хемінга

$$d(B^j, B^l) = 0{,}5 \sum_{i \in I} \sum_{s \in I} \left| b_{is}^j - b_{is}^l \right|, \qquad (3)$$

де $B^l = (b_{is}^l),\ l \in L,\ i, s \in I, -$ матриці попарних порівнянь (МПП), що відповідають ранжуванням $R^l, l \in L.$

Інколи для задач визначення результуючого ранжування використовується евклідова метрика.

Найпоширенішим методом знаходження результуючого ранжування об'єктів вважається обчислення медіани заданих ранжувань. Позначимо множину усіх можливих ранжувань n об'єктів через Ω^R, а множину МПП, які відповідають усім можливим ранжуванням n об'єктів – через Ω^B. Множину заданих експертами ранжувань будемо позначати через R^A, а множину відповідних їм МПП – через R^B. Для випадку, який розглядається у цій роботі, потужність множин R^A та R^B однакова: $\left|R^A\right| = \left|R^B\right| = n,\qquad R^l \in R^A,\ B^l \in R^B, l \in L.$ Зрозуміло, що $R^A \subset \Omega^R,\ R^B \subset \Omega^B.$ У загальному випадку потужність множини $\left|\Omega^B\right| = 2^{n(n-1)/2}.$ Але для методу, який описується у цій роботі, будемо вважати, що $\left|\Omega^R\right| = \left|\Omega^B\right| = n!,$ оскільки нас не цікавлять нетранзитивні елементи простору розв'язків $\Omega^B.$

Для метрики Кука (метрика неспівпадання рангів

об’єктів) виду (2) при використанні утилітарного критерію обчислюється медіана Кука-Сейфорда [4, 5]:

$$R^{CS} \in \Omega^{CS} = Arg \min_{R \in \Omega^R} \sum_{l \in L} d\left(R, R^l\right). \quad (4)$$

При використанні егалітарного критерію обчислюється ГВ-медіана (компроміс) [5]:

$$R^{ГВ} \in \Omega^{ГВ} = Arg \min_{R \in \Omega^R} \max_{l \in L} d\left(R, R^l\right). \quad (5)$$

Для метрики Хемінга (3) при використанні утилітарного критерію обчислюється медіана Кемені-Снелла:

$$R^{KC} \in \Omega^{KC} = Arg \min_{B \in \Omega^B} \sum_{l \in L} d\left(B, B^l\right). \quad (6)$$

При використанні егалітарного критерію обчислюється ВГ-медіана (компроміс) [4]:

$$R^{ВГ} \in \Omega^{ВГ} = Arg \min_{B \in \Omega^B} \max_{l \in L} d\left(B, B^l\right). \quad (7)$$

Крім того, в таких задачах можуть враховуватися коефіцієнти компетентності експертів: $\rho_1, ..., \rho_k$ (у нашому випадку: керівники підрозділів, відповідальних за впровадження заходів інформаційної безпеки, кожен – у зоні своєї відповідальності).

Методи визначення медіан виду (4)-(7) та їх особливості розглядаються у монографії [3].

Критеріальні функції, які застосовуються для визначення медіан (4)-(7), пов’язані з відстанями від заданих експертами ранжувань до обчислених розв’язків задачі. Тому ознакою ефективності одержаного розв’язку задачі слугують мінімальні значення критеріїв (4)-(7) для відповідних постановок задачі. Очевидно, що розв’язки, які не доставляють мінімумів зазначеним критеріям, є неефективними – вони домінуються розв’язками, які доставляють мінімум критеріям виду (4)-(7).

Для формалізації задачі побудови системи заходів щодо підвищення якості кібербезпеки можуть бути застосовані різні класи задач. Зокрема, перспективним є формалізація задачі розробки системи комплексної кібербезпеки у класі задач визначення нестрогого колективного ранжування: тобто, пошук результуючого ранжування може здійснюватися також у просторі нестрогих ранжувань. Такий підхід обґрунтовується тим, що деякі заходи безпеки мають виконуватися паралельно або навіть відбуватися одночасно. Тому логічно формалізувати поставлену задачу у класі обчислення розв’язку як нестрогого колективного ранжування [4, 6] (досконалі квазіпорядки [7], упорядкування [8], квазісерії [9], ранжування зі зв’язками [10], квазіпорядки [11], кластеризовані ранжування [12]).

Класичні методи теорії вибору (Кондорсе, Борда, Сімпсона, Копленда, Нансона, альтернативних голосів, відносної більшості тощо) також можуть бути обмежено застосовані при визначенні результуючого ранжування. Таке застосування описано, наприклад, в монографії [4]. Але при цьому слід враховувати недостатню обґрунтованість таких методів з точки зору теорії вимірювання. Для описаних вище задач більш прийнятними є алгебраїчні методи [4, 5].

Для обчислення медіан виду (4)-(7) може бути запропонована така схема.

Перш за все застосовується процедура об’єднання усіх заданих експертами об’єктів у єдину множину.

В результаті для кожного експерта маємо множину об’єктів, яка характеризується таким чином:

n_i – заданих ним у ранжуванні $R^i, i = 1,...,k,$ які будуть складати визначену частину відстаней;

$(n - n_i) = \nu_i$ – незаданих експертом у ранжуванні $R^i, i = 1,...,k,$ які складають ймовірнісну частину відстаней.

Після цього здійснюється генерування опорного ранжування $R^{*(0)}$ випадковим чином, тобто генеруванням перестановки n елементів.

На наступному етапі виникає необхідність визначити відстані від заданих експертами ранжувань до опорного ранжування за такими правилами:

– відстань від заданих експертами ранжувань $R^i, i = 1,...,k,$ до опорного ранжування $R^{*(0)}$ складається з двох складових: визначеної частини та ймовірнісної;

– ймовірнісна частина від заданого експертом ранжування $R^i, i = 1,...,k,$ до будь-якого опорного ранжування завжди дорівнює $\nu_i, i = 1,...,k,$ (для метрики Кука) (або $\nu_i \cdot (\nu_i - 1)/2, i = 1,...,k,$ – для метрики Хемінга);

– визначена частина дорівнює кількості інверсій між визначеними підмножинами перестановок: згенерованої опорної та заданої експертом $R^i, i = 1,...,k.$

Тобто в опорному ранжуванні кожного разу виділяємо n_i заданих і-м експертом об'єктів у ранжуванні R^i, і визначаємо кількість інверсій у перестановці об'єктів, відстань Хемінга (3) та відстань Кука (2).

Зрозуміло, що відстань від опорного ранжування $R^{*(0)}$ до кожного заданого експертом $R^i, i=1,...,k,$ дорівнює сумі ймовірнісної та визначеної частини.

Зважаючи на зазначений підхід, можна знайти результуючі ранжування виду (4)-(7), застосовуючи, наприклад, генетичний алгоритм [13] або евристичний алгоритм [14], розроблені для зазначених класів задач.

5. ВИСНОВКИ

В роботі запропоновано модель комплексного підходу до побудови системи кіберзахисту деякої складної організаційної системи. Також описано підхід до знаходження результуючого ранжування пріоритетності заходів як розв'язку задачі багатокритеріальної оптимізації. Послідовність виконання заходів може передбачати взаємодію виконавців і вимагати регламентування послідовності дій усіх елементів та підсистем організаційної системи.

Такий підхід дозволяє визначити результуючий комплекс заходів у вигляді медіани Кемені-Снелла, ГВ-медіани, медіани Кука-Сейфорда або ВГ-медіани. Запропонований підхід дозволяє:

– об'єднати різні за складом та пріоритетністю заходи інформаційної безпеки, запропоновані експертами різних підрозділів;

– знаходити компромісне рішення для різнорідної групи експертів;

– не порушувати задані переваги жодного експерта при обчисленні компромісного ранжування заходів кібербезпеки.

6. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Запропонований у роботі підхід може:

– бути корисним при розробці заходів забезпечення належного рівня кібербезпеки;

– сприятливим при розробці та впровадженні процедур швидкого реагування на загрози;

– бути використаний для розв'язання задачі доповнення даних;

– за необхідністю, містити елементи навчання, узгодження, координації, комплексності учасників експертної групи, які є керівниками підрозділів єдиної організаційної системи;

– відігравати позитивну роль при формалізації та подальшій оптимізації бізнес-процесів організації при аналізі результатів розв'язання описаної задачі;

– бути незамінним при комплексній побудові системи безпеки організації або її удосконаленні.

7. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] А.Г. Оксиюк, Я.В. Шестак, "Анализ современных методик и методов проведения оценки защищенности информационных систем", *Наукові записки Українського науково-дослідного інституту зв'язку*, №4, с.17-23, 2015.

[2] А.Г. Оксиюк, Я.В. Шестак, О.Д. Огбу, "Построение безопасной ифраструктуры как необходимость выживания", *Вісник Національного технічного університету «ХПІ». Збірник наукових праць. Серія: Механіко-технологічні системи та комплекси*, №50, с.112-117, 2016.

[3] А.Г. Додонов, Д.В. Ландэ, В.В. Цыганок, О.В. Андрейчук, С.В. Каденко, А.Н. Гайворонская, *Распознавание информационных операций*, Киев: ООО «Инжиниринг», 2017, 282 с.

[4] Г.М. Гнатієнко, В.Є. Снитюк, *Експертні технології прийняття рішень: монографія*, К.: ТОВ «Маклаут», 2008, 444с.

[5] О.Ф. Волошин, С.О. Мащенко, *Теорія прийняття рішень: навчальний посібник*, К.: Видавничо-поліграфічний центр "Київський університет", 2006, 304 с.

[6] И. М. Макаров, Т. М. Виноградская, А. А. Рубчинский, В. В. Соколов, *Теория выбора и принятия решений,* М.: Наука, 1982, 330 с.

[7] Ю.А. Шрейдер, *Равенство, сходство, порядок,* М.: Наука, 1971, 256 с.

[8] Дж. Кемени, Дж. Снелл, *Кибернетическое моделирование: некоторые приложения*, М.: Советское радио, 1972, 192 с.

[9] Б.Г. Миркин, *Проблема группового выбора*, М.: Наука, 1973, 256 с.

[10] М. Холлендер, Д. Вулф, *Непараметрические методы статистики*, М.: Финансы и статистика, 1983, 518 с.

[11] О.Ф. Волошин, Г.М. Гнатієнко, В.І. Кудін, *Послідовний аналіз варіантів. Технології та застосування*, К.: Стилос, 2013, 304 с.

[12] А.И. Орлов, *Методы принятия управленческих решений: учебник*, М.: КНОРУС, 2018, 286 с.

[13] Hnatiienko H.M., Kruglov A.I. "Definition of a compromise ranking on the set of individual rankings using the genetic

algorithm", *Міжнародний науковий симпозіум «ІНТЕЛЕКТУАЛЬНІ РІШЕННЯ». Обчислювальний інтелект (результати, проблеми, перспективи): праці міжнар. наук.-практ. конф.*, Ужгород, Україна, 15-20 квітня, 2019, с.63-64.

[14] Hnatiienko H.M., Hnatiienko V.H. "Heuristic algorithm for determining compromise rankings on a set of individual expert rankings", *INTELLIGENT SOLUTIONS. Decision Making Theory: Proceedings of the International School-Seminar*, Uzhhorod, Ukraine, April 15-20, 2019, pp.53-54.

Гнатієнко Григорій Миколайович, *кандидат технічних наук, здобув вищу освіту в Київському державному університеті ім. Т.Г. Шевченка в 1984 році, працює заступником декана факультету інформаційних технологій Київського національного університету імені Тараса Шевченка. Сфера наукових інтересів – обробка експертної інформації, інформаційні технології.*

Тмєнова Наталія Пилипівна*, кандидат фізико-математичних наук, здобула вищу освіту в Київському університеті імені Тараса Шевченка в 1999 році, працює доцентом кафедри інтелектуальних технологій факультету інформаційних технологій Київського національного університету імені Тараса Шевченка. Сфера наукових інтересів – технології обробки природномовної інформації, інформаційні технології.*

DOI
УДК 004.056.5

РОЗРОБКА СТЕГАНОГРАФІЧНОГО МЕТОДУ, СТІЙКОГО ДО АТАК ПРОТИ ВБУДОВАНОГО ПОВІДОМЛЕННЯ

Кобозєва Алла
orcid.org/0000-0001-7888-0499

Бобок Іван
orcid.org/0000-0003-4548-0709

м. Одеса, Одеський національний політехнічний університет, alla_kobozeva@ukr.net
м. Одеса, Одеський національний політехнічний університет, onu_metal@ukr.net

Історія статті:
Надійшла до редакції 15.02.2020
Прийнято 22.02.2020

Ключові слова:
стеганографічний метод;
цифрове зображення;
стійкість до атак;
сингулярне число;
блок матриці

Анотація. Особливості сучасних мережевих комунікацій приводять до необхідності використання при організації прихованого каналу зв'язку стеганографічних алгоритмів, стійких до стиску з втратами, залишаючи актуальними задачі розробки нових ефективних стеганометодів. В роботі на основі загального підходу до аналізу стану й технології функціонування інформаційних систем, який базується на теорії збурень та матричного аналізу, розроблено новий блоковий поліноміальний ступеня 2 стеганографічний метод, стійкий до атак проти вбудованого повідомлення, в тому числі значних, з одночасним збереженням надійності сприйняття формованого стеганоповідомлення, що забезпечується використаним математичним базисом. В якості контейнера розглядається цифрове зображення. Пропускна спроможність прихованого каналу зв'язку, що будується за допомогою розробленого методу, дорівнює n^{-2} біт/піксель при будь-якій алгоритмічній реалізації за умови використання всіх $n\times n$-блоків матриці контейнера, отриманих в результаті стандартної розбивки його матриці, при стеганоперетворенні. Вбудова додаткової інформації, що представляє бінарну послідовність і є результатом попереднього кодування інформації, що приховується, робиться шляхом використання нечутливих до збурних дій формальних параметрів матриці контейнера – сингулярних чисел її блоків малого розміру ($n\leq 8$). Забезпечення стійкості методу до збурних дій та надійності сприйняття формованого стеганоповідомлення досягається шляхом збільшення максимального сингулярного числа блоку, задіяного в стеганоперетворенні, при цьому величина збільшення визначається з використанням значень, отриманих шляхом піднесення сингулярних чисел блоку у натуральний ступінь k. Алгоритмічна реалізація методу вимагає додаткових досліджень для встановлення залежності значення параметру k від властивостей зображення-контейнеру.

Abstract. Features of modern network communications make it necessary to use in the organization of the hidden channel communication of steganographic algorithms that are resistant to loss compression, and leaving the tasks of developing new effective steganographic methods are relevant. The paper develops a new block steganographic method, which is resistant to attacks against the built-in message, including strong attacks. This method preserves the reliability of the perception of the formed quilting due to the mathematical basis used. It is based on a general approach to the analysis of the state and technology of information systems functioning, matrix analysis, perturbation theory. A digital image is treated as a container. The bandwidth of a hidden link that is built using the developed method is equal to n^{-2} bpp, $n\times n$ is the size of the blocks of the container that are obtained by the standard breakdown of its matrix. Such bandwidth is achieved with any algorithmic implementation of the method. Additional information is a binary sequence, it is the result of pre-coding of the information that is hidden. The embedding of additional information is done by using formal container matrix parameters that are insensitive to perturbation. These are singular values of its small blocks ($n\leq 8$). Increasing the maximum singular value of the block, which occurs when embedding additional information, leads to the stability of the method to the perturbing action and to ensure the reliability of the perception of the hip. The magnitude of the increase in the maximum singular value is determined using the values obtained by

raising the singular values of the block to a natural degree *k*. Algorithmic implementation of the method requires additional studies to determine the parameter *k*.

1. ВСТУП

На сьогоднішній день одною з важливих складових комплексного захисту інформації є стеганографічний захист [1-3]. У сучасному інформаційному суспільстві дослідження й розробки в області стеганографії стають усе більш популярними, і для цього є декілька причин:

- зростання актуальності проблеми захисту прав власності на інформацію, яка представлена в цифровому виді, у зв'язку з неперервним зростанням її цінності;
- обмеження на використання крипто-засобів у ряді країн світу, у тому числі, в Україні;
- для розв'язку задач інформаційної безпеки часто недостатньо зробити інформацію недоступною для порушника, потрібно приховати сам факт її передачі або зберігання.

В процесі стеганографування інформація, яка приховується, піддається попередньому кодуванню, в результаті якого отримується, як правило, бінарна послідовність, що далі називається додатковою інформацією (ДІ), яка вбудовується в деякий інформаційний контент – контейнер, що не привертає уваги, в результаті чого отримується стеганоповідомлення [3]. Стеганоповідомлення може передаватися каналами загального користування або зберігатися в отриманому вигляді. На сьогодні найчастіше як контейнери використовуються цифрові зображення (ЦЗ), які далі розглядаються в роботі.

Особливості сучасної комунікації, стрімке зростання обсягів інформації, що передається по каналах зв’язку, приводять до необхідності використання при організації прихованого каналу зв’язку стеганографічних алгоритмів, стійких до атак проти вбудованого повідомлення [3], зокрема до стиску з втратами. І хоча для стеганоперетворення можуть використовуватися як просторова область ЦЗ-контейнера, так і область перетворення [4,5], при організації стеганографування, коли мова йде про стійкість розроблюваних методів до атак, використовуються, як правило, області перетворення ЦЗ: частотна, області різних розкладань матриці (матриць) ЦЗ-контейнера. Розробці таких стеганоалгоритмів присвячено багато зусиль сучасних фахівців в галузі інформаційної безпеки [6-10], але, враховуючи можливу критичну важливість обсягу правильно відновленої інформації при організації прихованого каналу зв’язку в різних умовах його використання, зокрема, в умовах значних збурних дій (ЗД), питання підвищення ефективності таких алгоритмів залишається *актуальним*.

2. МАТЕРІАЛИ І МЕТОДИ

Існування значної кількості стійких до атак проти вбудованого повідомлення стеганографічних методів ще нещодавно ніяк не забезпечувалося наявністю математичних достатніх умов такої стійкості. Більшість з таких алгоритмів була розрахована на конкретні атаки, враховуючи їх специфіку, при цьому, хоча деякі з існуючих стеганографічних методів і позиціонувалися як стійкі до атак, незалежно від їх природи (та сили), але на практиці, в силу відсутності «універсального» математичного базису, їх реалізації все одно були розраховані на конкретні збурні дії (частіше, незначні) [11-14]. Більше того, деякі з фахівців-науковців обгрунтовано відкидали саму можливість створення таких методів. Наприклад, автори [15] стверджують, що стеганоалгоритм може бути стійким до стиску ЦЗ тільки тоді, коли він буде враховувати особливості алгоритму перспективного стиску, тобто алгоритм, стійкий до стиску, заснованому на вейвлет-перетворенні (JPEG2000), не може бути стійким до стиску, заснованому на дискретному косинусному перетворенні (JPEG). З урахуванням того, що стиск є лише одною з ЗД на стеганоповідомлення, специфікою якої є обнуління високочастотних складових сигналу, така думка обмежує можливості для побудови універсальних стеганоалгоритмів, стійких до атак проти вбудованого повідомлення, відволікаючи авторів розробок на врахування специфіки збурної дії.

В [16,17] вперше були запропоновані формальні умови забезпечення стійкості стеганоалгоритму до атак проти вбудованого повідомлення на основі загального підходу до аналізу стану й технології функціонування інформаційних систем [17]. Саме на основі цих достатніх умов в [8, 18] були запропоновані ефективні стеганографічні методи, алгоритмічні реалізації яких є стійкими до атак, зокрема значних, що використовують область сингулярного розкладання відповідної матриці [19], засновані на залученні до стегано-перетворення сингулярних чисел (СНЧ) [18] та сингулярних векторів (СНВ) [8] блоків матриці ЦЗ-контейнера, отриманих шляхом стандартної розбивки [20], ефективність яких в умовах значних ЗД перевищує ефективність сучасних

аналогів завдяки використаному математичному апарату. Однак для запропонованих методів, як і для більшості існуючих стеганометодів, стійких в умовах значних збурних дій, можливим є порушення надійності сприйняття стеганоповідомлення [3, 21], що є значним недоліком при організації прихованого каналу зв'язку.

3. МЕТА СТАТТІ ТА ПОСТАНОВКА ЗАДАЧ ДОСЛІДЖЕННЯ

Враховуючи спроможність та унікальність для забезпечення стійкості стеганометодів до атак проти вбудованого повідомлення математичного базису, запропонованого в [16,17], заснованого на теорії збурень та матричному аналізі, який вже добре зарекомендував себе в області стеганографії, здається доцільним його застосування для розробок нових стійких стеганометодів та алгоритмів.

Метою роботи є розробка нового стеганографічного методу, стійкого до атак проти вбудованого повідомлення, в тому числі значних, з одночасним збереженням надійності сприйняття формованого стеганоповідомлення і забезпеченням прийнятної пропускної спроможності прихованого каналу зв'язку.

Для досягнення поставленої мети в роботі розв'язуються наступні *задачі*:

1. Вибір формальних параметрів матриці ЦЗ-контейнера, нечутливих до збурних дій, які доцільно використовувати для організації процесу стійкого стеганоперетворення;
2. Збільшення нечутливості обраних формальних параметрів до збурних дій шляхом їх обґрунтованого перетворення;
3. Забезпечення надійності сприйняття формованного стеганоповідомлення.

4. ОСНОВНА ЧАСТИНА

Для простоти викладу наступного матеріалу, не обмежуючи спільності міркувань, як формальне представлення ЦЗ розглядається одна двовимірна матриця *F*, яка піддається стандартній розбивці на *n*×*n*-блоки, довільний з яких далі позначається *B*. Оскільки стан (зміна стану) будь-якої інформаційної системи згідно з загальним підходом до аналізу стану й технології функціонування інформаційних систем формально може бути представлений у вигляді сукупності збурень СНЧ і СНВ відповідної матриці (блоків матриці) [17], а процес стеганоперетворення ЦЗ-контейнера є процесом збурення зображення, розглянемо сукупність СНЧ і СНВ *B*, побудувавши для нього нормальне сингулярне розкладання [22]:

$$B = U\Sigma V^T, \quad (1)$$

де U, V – ортогональні $n\times n$-матриці лівих лексикографічно додатних і правих СНВ B відповідно, які представлені стовпцями u_i матриці U і v_i матриці V, $i = 1\ldots n$, $\Sigma = diag(\sigma_1,...,\sigma_n)$ - діагональна матриця СНЧ, $\sigma_1 \geq \ldots \sigma_n \geq 0$. Якщо СНЧ попарно різні, то розкладання (1) визначається однозначно [22]. СНЧ матриці (блоку матриці) ЦЗ є добре обумовленими, або нечутливими до збурних дій [19], чого не можна в загальному випадку сказати про СНВ [19], тому саме СНЧ розглядаються далі як формальні параметри ЦЗ-контейнера, які доцільно використовувати для організації процесу стеганоперетворення шляхом їх збурень. Відповідні СНЧ різних блоків навіть одного ЦЗ можуть значно відрізнятися одне від одного. Це ускладнює процес організації і аналізу їх збурень в процесі стеганоперетворення чи іншої збурної дії на стеганоповідомлення. Для полегшення цих процесів має сенс розглядати не самі СНЧ (в вигляді вектору $\sigma = (\sigma_1,\sigma_2,...,\sigma_n)^T$), а нормовані СНЧ, отримані шляхом нормування σ, результатом чого є вектор $\bar{\sigma} = \frac{\sigma}{\|\sigma\|}$, $\bar{\sigma} = (\bar{\sigma}_1,\bar{\sigma}_2,...,\bar{\sigma}_n)^T$, де $\|\sigma\|$ - норма вектора σ.

Для більшості блоків оригінального ЦЗ має місце співвідношення [17]:

$$\sigma_1 >> \sigma_2 \geq ... \geq \sigma_n \geq 0. \quad (2)$$

Враховуючи (2), в [23] отримано, що для таких блоків:

$$\angle(e_1,\bar{\sigma}) \approx 0, \quad (3)$$

де $\angle(e_1,\bar{\sigma})$ - кут між векторами $\bar{\sigma}$ і $e_1 = (1,0,...0) \in R^n$ - першим вектором стандартного базису простору R^n.

Оскільки вектор σ є стійким до збурних дій незалежно від природи і виду цієї дії, то такі ж властивості притаманні і вектору $\bar{\sigma}$.

Враховуючи, що для блоків оригінального ЦЗ в більшості з них точна рівність

$$\sigma_2 = ... = \sigma_n = 0 \quad (4)$$

не виконується, тобто в точності рівність (3) не досягається, то процес вбудови ДІ (біта ДІ) можна реалізувати шляхом збільшення в блоці,

що використовується при стеганоперетворенні, σ_1 на $\Delta > 0$ так, щоб в порівнянні з $\sigma_1 + \Delta$ всі інші СНЧ $\sigma_2,...,\sigma_n$ були такими, ненульовими значеннями яких можна було б знехтувати в умовах задачі, що розглядається, а кут $\angle\left(e_1,\bar{\sigma}\right)$, де вектор $\bar{\sigma}$ відповідає збільшеному першому СНЧ блоку ($\sigma_1 + \Delta$), практично дорівнював 0. Дійсно, оскільки

$$\left(e_1,\bar{\sigma}\right) = \|e_1\|\left\|\bar{\sigma}\right\| cos\left(\angle\left(e_1,\bar{\sigma}\right)\right) = cos\left(\angle\left(e_1,\bar{\sigma}\right)\right) = \\ = \frac{\sigma_1 + \Delta}{\sqrt{\left(\sigma_1 + \Delta\right)^2 + {\sigma_2}^2 + ... + {\sigma_n}^2}}, \quad (5)$$

де $(\cdot,\cdot)$ - скалярний добуток векторів-аргументів, в якості векторної норми для визначеності розглядається евклідова, то

$$\angle\left(e_1,\bar{\sigma}\right) = \arccos\frac{\sigma_1 + \Delta}{\sqrt{\left(\sigma_1 + \Delta\right)^2 + {\sigma_2}^2 + ... + {\sigma_n}^2}}. \quad (6)$$

Зрозуміло, що

$$\angle\left(e_1,\bar{\sigma}\right) \underset{\Delta\to\infty}{\to} 0, \quad (7)$$

але будь-які зміни СНЧ збурять матрицю ЦЗ. Враховуючи те, що для збереження надійності сприйняття стеганоповідомлення збурення матриці контейнера при стеганоперетворенні повинно бути незначним (чим менше збурення контейнера, тим більша ймовірність збереження надійності сприйняття стеганоповідомлення [17]), збурення блоку $\Delta > 0$ повинно бути якнайменше, щоб досягти такої близькості в (3), яка влаштує в умовах організації процесу стеганоперетворення. Таким чином, значення $\angle\left(e_1,\bar{\sigma}\right)$ хоча і повинно прямувати до нуля, але не за рахунок того, що $\Delta \to \infty$. З огляду на це розглянемо можливість застосування запропонованої ідеї зменшення кута $\angle\left(e_1,\bar{\sigma}\right)$, але не безпосередньо для вектору $\sigma = \left(\sigma_1,\sigma_2,...,\sigma_n\right)^T$, а для його перетворення.

В [24] доведено, що

$$\angle\left(e_1,\bar{\bar{\sigma}}\right) < \angle\left(e_1,\bar{\sigma}\right), \quad (8)$$

де $\bar{\bar{\sigma}} = \frac{\left({\sigma_1}^2,{\sigma_2}^2,...,{\sigma_n}^2\right)^T}{\left\|\left({\sigma_1}^2,{\sigma_2}^2,...,{\sigma_n}^2\right)^T\right\|} = \left(\bar{\bar{\sigma}}_1,\bar{\bar{\sigma}}_2,...,\bar{\bar{\sigma}}_n\right)^T$, $\angle\left(e_1,\bar{\bar{\sigma}}\right)$ - кут між векторами $\bar{\bar{\sigma}}$ і e_1, при цьому вектор $\bar{\bar{\sigma}}$ має меншу чутливість до збурних дій, ніж $\bar{\sigma}$, що принципово дає йому перевагу в порівнянні з $\bar{\sigma}$ в випадку залучення його до процесу стеганоперетворення, яке повинно бути стійким до атак проти вбудованого повідомлення.

Ступінь близькості будь-якого нормованого вектора, елементи якого задовольняють (2), до e_1 визначається відстанню між першою і другою компонентами (відповідно з (6), (7) для вектора СНЧ): чим більше ця відстань, тим ближче нормований вектор до e_1. Нехай СНЧ σ_1 збільшили на деяку величину Δ, тоді відстань $d = \sigma_1 - \sigma_2$ між σ_1 і σ_2 збільшиться на Δ і буде дорівнювати: $\bar{d} = d + \Delta$. В той же час відстань між ${\sigma_1}^2$ і ${\sigma_2}^2$ теж зміниться і буде дорівнювати:

$$\left(\sigma_1 + \Delta\right)^2 - {\sigma_2}^2 = \bar{d}\left(\sigma_1 + \sigma_2 + \Delta\right). \quad (9)$$

З врахуванням (2) і того, що $\Delta > 0$, з (9) випливає, що відстань між ${\sigma_1}^2$ і ${\sigma_2}^2$ збільшиться більше ніж між σ_1 і σ_2, що приведе до того, що після збурення σ_1 на $\Delta > 0$ відповідний збурений вектор $\bar{\bar{\sigma}}$ буде ближче до e_1, чим збурений вектор $\bar{\sigma}$; чи інакше: для того, щоб досягти однакової близькості з e_1 векторам $\bar{\sigma}$ і $\bar{\bar{\sigma}}$, першу компоненту вектору $\left({\sigma_1}^2,{\sigma_2}^2,...,{\sigma_n}^2\right)^T$ треба збільшити менше, ніж першу компоненту $\sigma = \left(\sigma_1,\sigma_2,...,\sigma_n\right)^T$. Це означає, що для організації стеганоперетворення за рахунок наближення вектора СНЧ (перетворених СНЧ) до вектора e_1 з більшою ймовірністю надійність сприйняття буде зберігатися у випадку вектора $\left({\sigma_1}^2,{\sigma_2}^2,...,{\sigma_n}^2\right)^T$.

Ефект (8), (9) (зростання відстані між першою та другою компонентами відповідного вектора) збільшиться ще більше, якщо розглянути вектор, елементами якого є СНЧ в ступені, більшому за 2: $\left({\sigma_1}^k,{\sigma_2}^k,...,{\sigma_n}^k\right)^T$. Взагалі, чим більше буде ступінь, тим до більшого збільшення відстані

між першим і другим елементом будемо приходити при збільшенні σ_1:

$$\left(\sigma_1+\Delta\right)^k-{\sigma_2}^k=\overline{d}\sum_{i=1}^{k}\left(\sigma_1+\Delta\right)^{k-i}{\sigma_2}^{i-1}. \quad (10)$$

тобто зменшення кута $\angle\left(e_1,\overline{\sigma}\right)$ до нуля (до значення, порівнянного з нулем) можна забезпечити не тільки за рахунок $\Delta\to\infty$ (7), що приведе до порушення надійності сприйняття стеганоповідомлення, а за рахунок збільшення ступеня СНЧ блоку ${\sigma_i}^k, i=\overline{1,n}$, при безпосередньому стеганоперетворенні при фіксованому (невеликому) значенні Δ:

$$\lim_{k\to\infty}\arccos\frac{\left(\sigma_1+\Delta\right)^k}{\sqrt{\left(\sigma_1+\Delta\right)^{2k}+{\sigma_2}^{2k}+\ldots+{\sigma_l}^{2k}}}=1. \quad (11)$$

З урахуванням (11) основні кроки метода, що пропонується, стійкого до атак, в тому числі, значних, проти вбудованого повідомлення, що зберігає надійність сприйняття стеганоповідомлення, наступні.

Вбудова ДІ

Крок 1. Матриця *F* ЦЗ-контейнера розбивається стандартним чином на $2l\times 2l$-блоки малого розміру ($n=2l$).

Крок 2. Черговий $2l\times 2l$-блок *B* контейнера, що використовується для вбудови 1 біта p_m ДІ і визначається згідно з секретним ключем, розбивається на чотири $l\times l$-блоки $B^{(i)}$, $i=\overline{1,4}$, що не перетинаються.

Крок 3. Для кожного з 4-х отриманих $l\times l$-блоків $B^{(i)}$, $i=\overline{1,4}$, знайти СНЧ $\sigma_1^{(i)}$ і $\sigma_2^{(i)}$ і відстані між їх *k*-ими ступенями: $d_k^{(i)}=\left(\sigma_1^{(i)}\right)^k-\left(\sigma_2^{(i)}\right)^k$. Обчислити $d_{max}=\max\limits_{1\le i\le 4}d_k^{(i)}$.

Крок 4. Згідно з секретним ключем залежно від значення p_m визначається конкретний блок $B^{(j)}$ в межах *B*, в який буде відбуватися безпосередньо вбудова p_m. В цьому блоці в результаті вбудови ДІ потрібно забезпечити

$$\left(\sigma_1^{(j)}\right)^k\ge\left(\sigma_2^{(j)}\right)^k+d_{max}+\Delta, \quad (12)$$

де $\Delta>0$. Це робиться шляхом збільшення СНЧ $\sigma_1^{(j)}$, результатом чого є $\sigma_{1,new}^{(j)}$:

$$\sigma_{1,new}^{(j)}\ge\sqrt[k]{\left(\sigma_2^{(j)}\right)^k+d_{max}+\Delta}. \quad (13)$$

Крок 5. Відновити $B_{new}^{(j)}$ з урахуванням нового значення першого СНЧ і незмінених інших СНЧ і СНВ.

Крок 6. Побудувати $2l\times 2l$-блок стеганоповідомлення, що відповідає блоку *B* контейнера, враховуючи незмінність всіх $B^{(i)}$, $i=\overline{1,4}, i\ne j$, і $B_{new}^{(j)}$.

Крок 7. Перехід на крок 2 до наступного блоку контейнера, задіяного в стеганоперетворенні, при наявності такого.

Крок 8. Закінчення формування стеганоповідомлення; результат – ЦЗ з матрицею $\overline{F}$.

Декодування ДІ

Крок 1. Матриця $\overline{F}$ ЦЗ-стеганоповідомлення розбивається стандартним чином на $2l\times 2l$-блоки.

Крок 2. Черговий $2l\times 2l$-блок $\overline{B}$ стеганоповідомлення, в якому міститься ДІ, що визначається згідно з секретним ключем, розбивається на чотири $l\times l$-блоки $\overline{B}^{(i)}$, $i=\overline{1,4}$, що не перетинаються.

Крок 3. Для кожного з 4-х отриманих $l\times l$-блоків $\overline{B}^{(i)}$, $i=\overline{1,4}$, знайти:

3.1. СНЧ: $\sigma_{1_S}^{(i)}$, $\sigma_{2_S}^{(i)}$,…, $\sigma_{l_S}^{(i)}$, $i=\overline{1,4}$;

3.2. Кути $\sigma_S^{(i)}$ між векторами $\left(\left(\sigma_{1_S}^{(i)}\right)^k,\left(\sigma_{2_S}^{(i)}\right)^k,\ldots,\left(\sigma_{l_S}^{(i)}\right)^k\right)^T$ і e_1, $i=\overline{1,4}$.

Крок 4. Визначити:

$$\sigma_S^{(j)}=\min_{1\le i\le 4}\sigma_S^{(i)}. \quad (14)$$

Декодування чергового біту $\overline{p}_m$ ДІ відбувається з $\overline{B}^{(j)}$. Значення $\overline{p}_m$ залежить від того, з якого саме $l\times l$-блоку $\overline{B}^{(j)}$ відбувається декодування (з врахуванням секретного ключа).

Крок 5. Перехід на крок 2 до наступного блоку стеганоповідомлення, задіяного в стеганоперетворенні, при наявності такого.

Крок 6. Закінчення декодування ДІ.

Зауваження 1. Враховуючу наявну кореляцію між значеннями сусідніх пікселів оригінального

ЦЗ, можна стверджувати, що значення $\sigma_1^{(i)}$, $i=\overline{1,4}$, не значно відрізняються одне від іншого [25], тому вибір конкретного блоку $B^{(j)}$ на кроці 4, який відбувається згідно з секретним ключем без врахування наступного спотворення блоку при стеганоперетворенні, не може привести до таких впливів на надійність сприйняття стеганоповідомлення, які будуть суттєво відрізнятися при різних значеннях *j*.

Зауваження 2. Головним при побудові алгоритмічної реалізації запропонованого методу є визначення параметру *k*. Це визначення повинно досягатися за допомогою компромісу між двома вимогами. По-перше, забезпечення близькості вектора $\left(\left(\sigma_{1,new}{}^{(j)}\right)^k,\left(\sigma_2{}^{(j)}\right)^k,...,\left(\sigma_n{}^{(j)}\right)^k\right)^T$ блоку $B_{new}{}^{(j)}$ до вектора e_1 при вбудові ДІ буде тим краще, чим більше буде значення *k*. Але, по-друге, при дуже великому *k* значення $\sigma_S^{(i)}$, $i=\overline{1,4}$, враховуючи особливості машинної арифметики, можуть практично не відрізнятися одне від одного, що може привести до помилок при визначенні блоку $\overline{B}^{(j)}$ для декодування чергового біта ДІ, тобто помилок при декодуванні.

Зауваження 3. Визначення параметру Δ, що використовується в процесі вбудови ДІ на 4 кроці, потребує аналізу спотворення ЦЗ в результаті стеганоперетворення. Параметр Δ повинен визначатися таким чином, щоб значення пікового відношення «сигнал-шум» (*PSNR*) для побудованого стеганоповідомлення було більшим за 35-40 dB, що вважається прийнятним при організації прихованого каналу зв'язку [15].

Зауваження 4. Запропонований метод є блоковим, тому його обчислювальна складність для ЦЗ розміром $t\times t$ визначається кількістю блоків, які не перетинаються, отриманих шляхом стандартної розбивки, тобто складає $\underline{O}\left(t^2\right)$ операцій.

Зауваження 5. При алгоритмічній реалізації запропонованого методу доцільно використовуватися блоки малого розміру: $n\leq 8$, оскільки для оригінальних ЦЗ співвідношення (4) при $n>8$ практично ніколи не виконується [25], а це означає, що його штучне досягнення в процесі вбудови ДІ в такому випадку може привести до порушення надійності сприйняття формованного стеганоповідомлення.

Зауваження 6. При вбудові ДІ на кроці 4 секретний ключ може визначати номер *j* блоку $B^{(j)}$ залежно від значення p_m, враховуючи, наприклад, його парність/непарність, тобто, якщо $p_m=0$, то $j=2\vee j=4$, інакше $j=1\vee j=3$. В такому випадку парність/непарність *j* для $\overline{B}^{(j)}$ на кроці 4 декодування визначить $\overline{p}_m$.

Пропускна спроможність прихованого каналу зв'язку, що будується за допомогою розробленого методу, буде дорівнювати n^{-2} біт/піксель при будь-якій алгоритмічній реалізації за умови використання всіх блоків матриці ЦЗ-контейнера, отриманих в результаті стандартної розбивки. Така пропускна спроможність є прийнятною для сучасних стеганометодів [15].

5. ВИСНОВКИ

В роботі на основі теорії збурень та матричного аналізу розроблено новий блоковий стеганографічний метод, стійкий до атак проти вбудованого повідомлення, в тому числі значних, з одночасним збереженням надійності сприйняття формованого стеганоповідомлення, що забезпечується використаним математичним базисом.

Вбудова ДІ робиться шляхом використання нечутливих до збурних дій формальних параметрів матриці ЦЗ-контейнера – СНЧ її блоків малого розміру, отриманих шляхом її стандартної розбивки. Збільшення стійкості методу до збурних дій та забезпечення надійності сприйняття формованного стеганоповідомлення досягається шляхом збільшення максимального СНЧ блоку, задіяного в стеганоперетворенні, при цьому збурення СНЧ визначається за допомогою піднесення СНЧ блоків у натуральний ступінь *k*.

Алгоритмічна реалізація методу вимагає додаткових досліджень для встановлення значень параметрів *k* і Δ, над чим зараз працюють автори.

6. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] M.E. Saleh, A.A. Aly, F.A. Omara, "Data security using cryptography and steganography techniques," *International Journal of Advanced Computer Science and Applications*, Vol. 7, Issue 6, pp. 390-397, 2016.

[2] S. Malalla, F.R. Shareef, "Novel approach for Arabic text steganography based on the "BloodGroup" text hiding method," *Engineering, Technology & Applied Science Research*, Vol. 7, Issue 2, pp. 1482-1485, 2017.

[3] В.Г. Грибунин, И.Н. Оков, И.В. Туринцев, *Цифровая стеганография*, М.: СОЛОН-Пресс, 2009. 272 с.

[4] D. Subhashini, P. Nalini, G. Chandrasekhar, "Comparison analysis of spatial Domain and compressed Domain steganographic techniques," *International Journal of Engineering Research and Technology*, Vol. 1, Issue 4, pp. 1-6, 2012.
[5] B. Li, "A survey on image steganography and steganalysis," *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 2, Issue 2, pp. 142-172, 2011.
[6] S. Singh, T.J. Siddiqui, "A security enhanced robust steganography algorithm for data hiding," *International Journal of Computer Science Issues*, Vol. 9, Issue 3, pp. 131-139, 2012.
[7] S. Singh, T.J. Siddiqui, "Robust image steganography using complex wavelet transform," in *Proceedings of 2013 International Conference on Multimedia, Signal Processing and Communication Technologies*, Aligarh, India, 23-25 Nov. 2013, pp.56 – 60.
[8] М.А. Мельник, «Sign-нечувствительность сингулярных векторов матрицы изображения как основа стеганоалгоритма, устойчивого к сжатию», *Інформатика та математичні методи в моделюванні*, Том 3, № 2, С. 116-126, 2013.
[9] Z. Bao, X. Luo, Y. Zhang, C. Yang, F. Liu, "A robust image steganography on resisting JPEG compression with no side information," *IETE Technical Review*, Vol. 35, Issue 1, pp. 4-13, 2018.
[10] J. Tao, S. Li, X. Zhang, Z. Wang, "Towards robust image steganography," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 29, Issue 2, pp. 594-600, 2019.
[11] Г.Ф. Конахович, А.Ю. Пузыренко, *Компьютерная стеганография: теория и практика*, К.: МК-Пресс, 2006. 288 с.
[12] А.А. Кобозева, М.А. Мельник, «Нечувствительность стеганосообщения к сжатию и формальные достаточные условия ее обеспечения», *Збірник наукових праць Військового інституту КНУ імені Тараса Шевченка*, № 38, С. 193-203, 2012.
[13] А.А. Кобозева, В.А. Хорошко, *Анализ информационной безопасности*, К.: ГУИКТ, 2009. 251 с.
[14] М.А. Мельник, «Стеганоалгоритм, устойчивый к сжатию», *Інформаційна безпека*, № 2(8), с. 99-106, 2012.
[15] Д. Деммель, *Вычислительная линейная алгебра: теория и приложения*, М.: Мир, 2001. 430 с.
[16] Р. Гонсалес, Р. Вудс, *Цифровая обработка изображений*, М.: Техносфера, 2006. 1070 с.
[17] C. Bergman, J. Davidson, "Unitary embedding for data hiding with the SVD," in *Proceedings of Security, Steganography, and Watermarking of Multimedia Contents VII*, Vol. 5681, pp. 619-630, 2005.
[18] A.A. Kobozeva, I.I. Bobok, A.I. Garbuz, "General principles of integrity checking of digital images and application for steganalysis," *Transport and Telecommunication*, Vol. 17, Issue 2, pp. 128-137, 2016.
[19] И.И. Бобок, «Теоретическое развитие общего подхода к проблеме выявления нарушений целостности цифровых контентов, основанного на анализе полного набора формальных параметров», *Інформатика та математичні методи в моделюванні*, Том 7, № 3, с. 170-177, 2017.
[20] І.І. Бобок, «Теоретичні основи методу виявлення порушення цілісності цифрового зображення в результаті накладання шуму», *Збірник наукових праць Військового інституту КНУ імені Тараса Шевченка*, № 63, С. 73-84, 2019.
[21] M.O. Kozina, "Steganographic method for solving a triple task", Матеріали *VII міжнародна науково-практична конференція «Проблеми та перспективи розвитку IT-індустрії»*, Харків, Україна, 17-18 квітня 2015 р., с. 32.
[22] T.H. The, T.L. Tien, "An efficient blind watermarking method based on significant difference of wavelet tree quantization using adaptive threshold," *International Journal of Electronics and Electrical Engineering*, Vol. 1, Issue 2, pp. 98-103, 2013.
[23] R.S. Run, S.J. Horng, W.H. Lin, T.W. Kao, and P. Fan, "An efficient wavelet-tree-based watermarking method," Expert Systems with Applications., Vol. 38, pp. 14357-14366, 2011.
[24] J.C. Patra, A.K. Kishore, C. Bornand, "Improved CRT-based DCT domain watermarking technique with robustness against JPEG compression for digital media authentication," in *Proceedings of 2011 IEEE International Conference on Systems, Man, and Cybernetics*, Anchorage, AK, USA, 9-12 Oct. 2011, pp. 2940–2945.
[25] Y.R. Wang, W.H. Lin, L. Yang, "An intelligent watermarking method based on particle swarm optimization," *Expert Systems with Applications*, Vol. 38, pp. 8024-8029, 2011.

***Кобозєва Алла Анатоліївна**, доктор технічних наук, професор, завідувач кафедри Інформатики і управління захистом інформаційних систем Одеського національного політехнічного університету.*

***Бобок Іван Ігорович**, кандидат технічних наук, доцент кафедри Інформаційних технологій проектування в електроніці й телекомунікаціях Одеського національного політехнічного університету.*

DOI:
УДК 004.056.55

НОВИЙ ПІДХІД ДО ПОБУДОВИ ПОСТ-КВАНТОВОЇ СХЕМИ ЕЛЕКТРОННОГО ЦИФРОВОГО ПІДПИСУ

Олександр Кузнецов [1]
0000-0003-2331-6326

Анастасія Кіян [1]
0000-0003-2110-010X

Андрій Пушкарьов [2]
0000-0002-2429-0394

Тетяна Кузнецова [1]
0000-0001-6154-7139

[1] Харківський національний університет імені В.Н.Каразіна, майдан Свободи 4, 61022, Харків, Україна, kuznetsov@karain.ua, nastyak931@gmail.com, kuznetsova.tatiana17@gmail.com, https://www.univer.kharkov.ua
[2] Адміністрація Державної служби спеціального зв'язку та захисту інформації України, вул. Солом'янська, 13, 03110, м. Київ, Україна, http://www.dsszzi.gov.ua

Історія статті:
Надійшла до редакції 14.02.2020
Прийнято 28.02.2020

Ключові слова: кодові криптосистеми, електронний цифровий підпис, пост-квантова криптографія, квантова стійкість

Анотація. Методи криптографічного захисту інформації мають важливе значення в розбудові сучасної інфраструктури кібербезпеки. Останнім часом з'явилися нові виклики та загрози криптографічним перетворенням. Зокрема, поява та стрімкий розвиток новітніх технологій квантових обчислень зумовлює негайну потребу в розробці та дослідженні нових методів пост-квантових криптографічних перетворень, тобто таких, які будуть стійкими навіть за умови можливого застосування квантового криптоаналізу. Стаття присвячена аналізу можливостей реалізації схем електронного цифрового підпису з використанням кодів, що виправляють помилки. Подібний підхід дозволяє побудувати схеми, стійкі як до класичного криптоаналізу, так і в умовах криптоаналізу з використанням квантових обчислень. У рамках статті описано принципи функціонування класичної кодової схеми електронного цифрового підпису CFS, що побудована з використанням перетворення Нідеррайтера, а також запропоновано новий підхід, що дозволяє реалізувати підпис згідно перетворень схеми Мак-Еліса. Такий підхід зберігає переваги свого попередника і надає додатковий захист від спеціального виду атак. Також у статті здійснено порівняльний аналіз та характеристику розглянутих схем згідно критеріїв стійкості до класичного та квантового криптоаналізу, складності виконання необхідних перетворень та довжині формованих підписів. Отримані результати дозволяють стверджувати про можливість побудови надійних та безпечних криптографічних перетворень, зокрема, алгоритмів електронного цифрового підпису, які базуються на застосуванні кодів та є безпечними навіть в умовах можливого застосування квантового криптоаналізу. Втім, варто зазначити, що недоліком схем підписів на основі коду є великий обсяг ключових даних, необхідних алгоритму, а також складність у створенні підпису через необхідність багаторазового розшифрування синдрому, що залишається актуальною темою і потребує подальших досліджень.

Key words: Code-based cryptosystems, electronic digital signature, post-quantum cryptography, quantum resistance

Abstract. Cryptographic information security techniques are essential in building a modern cybersecurity infrastructure. Recently, there have been new challenges and threats to cryptographic transformation. In particular, the emergence and rapid development of the latest quantum computing technologies necessitates the urgent need for the development and research of new methods of post-quantum cryptographic transformations, that is, those that will be sustainable even if quantum cryptanalysis is possible. This article is devoted to the analysis of possibilities of implementation of digital signature schemes based on using error-correcting codes. This approach allows cryptographers to build schemes that are resistant to both classic cryptanalysis and cryptanalysis which uses quantum computing. The article describes the principles of the classic digital signature scheme which is named CFS and built using a Niederreiter-like transform, and also we propose a new approach that enables an implementation of signature according to the McEliece transformations. This approach preserves the advantages of its predecessor and provides additional protection against special attacks. Also, a comparative analysis and characterization of the considered schemes according to the criteria of resistance to classic and quantum cryptanalysis, complexity of necessary transformations and length of generated signatures are made. The results show that reliable and secure cryptographic transformations can be built, in particular, electronic

digital signature algorithms that are code-based and secure even in the case of quantum cryptanalysis. However, it should be noted that the drawback of code-based signature schemes is the large amount of key data required by the algorithm, as well as the difficulty in creating a signature due to the need for multiple decryption of the syndrome, which remains a topical topic and needs further research.

1. ВСТУП

Використовуючи закони квантової механіки, можна створити принципово новий тип обчислювальних машин, які дозволять вирішувати деякі завдання, недоступні навіть самим потужним сучасним суперкомп'ютерам. Різко зросте швидкість багатьох складних обчислень; повідомлення, надіслані по лініях квантового зв'язку, неможливо буде ні перехопити, ні скопіювати. Сьогодні вже створені прототипи цих квантових комп'ютерів майбутнього. Активні розробки у сфері квантового комп'ютера у 2016 році логічно призвели до реакції криптографічної спільноти у вигляді анонсування конкурсу пост-квантової стандартизації Національним інститутом стандартів та технологій. Розробники представили свої проекти у трьох напрямках: схеми електронного цифрового підпису, направлене шифрування та інкапсуляції ключів [1-3].

Статистика першого та другого туру конкурсу, що завершився у 2019 році, продемонструвала, що популярними напрямки розробки є криптографія, що базується на решітках, кодова криптографія, мультиваріативна криптографія та криптографія, заснована на геш-функціях. При цьому перші два напрямки охоплюють ¾ з усіх представлених проектів, що підтверджує актуальність всебічного розгляду цих напрямків у контексті реалізації криптографічних перетворень.

Сьогодні ведуться дебати щодо доцільності використання у реальних системах схеми електронних цифрових підписів, побудованих з використанням кодів. Сумніви виникають через великі об'єми ключових даних, що потребуються подібного роду схемам, проблеми швидкої підробки та інше. Слід зауважити, що, на жаль, ні одна кодова схема цифрового підпису не пройшла до другого туру конкурсу стандартизації NIST, а, отже, залишається актуальним питання розробки нових схем, що усували б недоліки своїх попередників [4].

У випадку розгляду кодових схем електронного підпису класичною є схема CFS, що названа згідно ініціалів її розробників Courtois, Finiasz та Sendrier, які вперше застосували підхід, заснований на алгебраїчних блокових кодах по відношенню до цифрового підпису. Вона побудована на використанні перетворень згідно криптосистеми Нідеррайтера. У статті розглянуто принципи функціонування CFS, а також запропоновано нову схему електронного цифрового підпису, що базується на використанні криптосистеми Мак-Еліса. Запропонований підхід дозволяє не тільки реалізувати пост-квантово стійку схему, але і надає додаткові переваги у вигляді захищеності від особливого типу атак. Аналіз такого роду атак по відношенню до схеми, а також порівняльне дослідження запропонованої схеми та схеми CFS розглянуто надалі.

2. CFS ЯК КЛАСИЧНА КОДОВА СХЕМА ЕЦП

CFS передбачає використання алгебраїчного (n, k, d) коду з класу незвідних кодів Гоппи. Формування ключових даних для функціонування схеми проходить аналогічно алгоритму, описаному для формування ключів схеми Нідеррайтера [5]. У якості вхідних даних схеми CFS використовуються:

- функція гешування h;
- швидкий алгоритм декодування алгебраїчного коду;
- безпосередньо повідомлення (відкритий текст).

Швидкий алгоритм декодування алгебраїчного коду, тобто той, що має поліноміальну складність, застосовується для декодування синдромної послідовності $s = (s_0, s_1, ..., s_{n-k-1})$ (криптограми у схемі Нідеррайтера) [6]. У випадку застосування алгоритму можливо виконання однієї з ситуації:

- Якщо декодування успішне, буде виведено знайдений вектор помилок $e = (e_0, e_1, ..., e_{n-1})$, що відповідає синдрому.
- Якщо декодування невдале, буде виведено повідомлення про помилку знаходження вектору помилок.

Алгоритм формування підпису полягає у поступовому виконанні декількох кроків (рис.1) [7].

1. Гешування відкритого тексту M;
2. Присвоєння лічильнику i значення i=1.
3. Геш-значення повідомлення $h(M)$ та лічильника i представляють як бітові послідовності, від конкатенації яких, обчислюють нове геш - значення $h(h(M)\|i)$.
4. $h(h(M)\|i)$ інтерпретується як синдромна послідовність $s_X = (s_0, s_1, ..., s_{n-k-1})$, що обчислена для деякого довільного кодового слова та вектору помилок $e = (e_0, e_1, ..., e_{n-1})$.
5. Формування вектору:

$$s_X^{*T} = X^{-1} \cdot s_X{}^T = X^{-1} \cdot H_X \cdot e^T = \\ = H \cdot P \cdot e^T = H \cdot \bar{e}^T \quad (1)$$

6. Застосування алгоритму швидкого декодування для знаходження вектору $\bar{e}^T = P \cdot e^T$.

7. Якщо декодування невдале, то необхідно інкременувати значення лічильника i, і повернутися на крок 3. Алгоритм застосовується до тих пір, пока не буде виведено знайдений вектор $\bar{e}^T = P \cdot e^T$, що відповідає вектору s_X^*.

8. Коли вектор s_X^* знайдено, обчислюємо

$$e^T = P^{-1} \cdot \bar{e}^T = P^{-1} \cdot P \cdot e^T \quad (2)$$

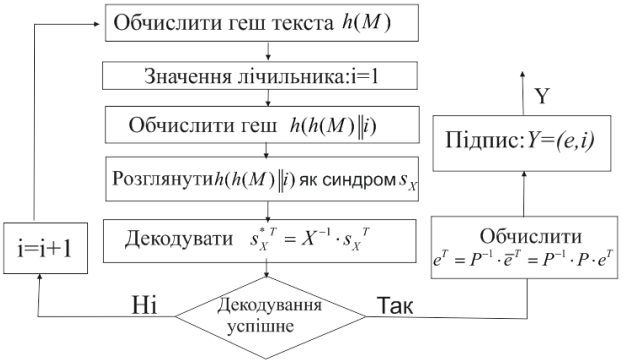


Рис.1 - Процес формування підпису згідно схеми CFS

Результатом виконання послідовності кроків є формування кінцевого підпису повідомлення M, що складається з двох частин: значення лічильника та вектору e, $Y = (e,i)$. Формально можна записати сформований підпис як

$$Y = (e,i) : H_X \cdot e^T = \left(h(h(M)\|i)\right)^T \quad (3)$$

Для того, щоб перевірити справжність підпису у якості вхідних даних необхідно мати відкритий ключ, що складається з матриці H_X, функцію хешування h, сам підпис $Y = (e,i)$ та повідомлення M. Для того, щоб верифікувати підпис потрібно обчислити значення двох векторів:

$$(s'_X)^T = H_X \cdot e^T,\ (s''_X)^T = h(h(M)\|i).$$

Цифровий підпис визнається правильним тільки за умови, що ці два вектори будуть однаковими.

Таким чином сутність функціонування схеми CFS можна визначити як багаторазове гешування повідомлення, конкатенованого з випадковим значенням лічильна для виділення коректної синдромної послідовності.

3. ОСОБЛИВОСТІ ЗАПРОПОНОВАНОЇ СХЕМИ ПІДПИСУ

Ми пропонуємо побудови схему електронного цифрового підпису, що базується на використанні односторонньої функції Мак-Еліса [8]. Формування ключових даних проходить аналогічно схемі CFS. Отже секретними ключами схеми є матриці X та P (у випадку недвійкових кодів додається ще матриця D), які є невиродженою матрицею $k \times k$ та матрицею перестановки $n \times n$ відповідно, а також швидкий алгоритм декодування алгебраїчного коду. Відкритим ключем у свою чергу виступає матриця G_X, що формується згідно правила $G_X = X \cdot H \cdot P \cdot D$, де G - це породжувальна матриця алгебраїчного коду [9].

При формуванні підпису аналогічно схемі CFS використовується функція гешування h, що була докладно описана раніше, тому на ній не буде зосереджуватися увага. Алгоритм декодування полягає у можливості знаходження вектору помилок $e = (e_0, e_1, ..., e_{n-1})$ та вектору $I = (I_0, I_1, ..., I_{k-1})$ згідно початковому кодовому слову з помилками $c_X^* = (c*_0, c*_1, ..., c*_{n-1})$, враховуючи співвідношення $c_X^* = I \cdot G_X + e$.

Алгоритм формування підпису запропонованої схеми проходить згідно таких кроків (рис.5).

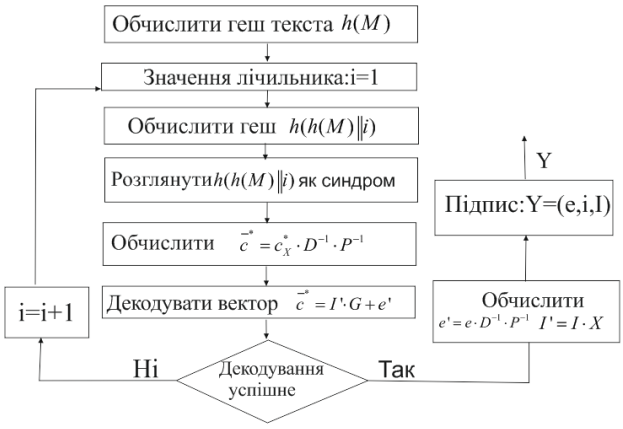


Рис.2 - Процес формування підпису згідно запропонованої схеми

1. Знаходження геш-коду від повідомлення, яке необхідно підписати
2. Присвоюємо значення лічильна i рівним 1.
3. Конкатенуємо геш-значення повідомлення та лічильника з подальшим хешуванням утвореної послідовності: $h(h(M)\|i)$
4. Інтерпретуємо $h(h(M)\|i)$ як кодове слово з помилками $c_X^* = (c*_0, c*_1, ..., c*_{n-1})$, що обчислене для певних векторів $e = (e_0, e_1, ..., e_{n-1})$ та $I = (I_0, I_1, ..., I_{k-1})$.

5. Знаходимо значення вектора $\bar{c}^{*} = c_X^{*} \cdot D^{-1} \cdot P^{-1}$. Цей вектор представляє викревлене не більше, ніж в t позиціях кодове слово, а отже, його можна декодувати з використанням алгоритму поліноміальної складності.

6. Висуваємо припущення:

$$\begin{aligned} \bar{c}^{*} &= c_X^{*} \cdot D^{-1} \cdot P^{-1} = (I \cdot G_X + e) \cdot D^{-1} \cdot P^{-1} = \\ &= (I \cdot X \cdot H \cdot P \cdot D + e) \cdot D^{-1} \cdot P^{-1} = \\ &= I \cdot X \cdot H + e \cdot D^{-1} \cdot P^{-1} \end{aligned} \quad (4)$$

7. Застосовуємо алгоритм поліноміальної складності для декодування кодового слова $\bar{c}^{*} = I' \cdot G + e'$, $e' = e \cdot D^{-1} \cdot P^{-1}$ і отримання вектору $I' = I \cdot X$.

8. Якщо внаслідок декодування не отримано вірний результат, збільшуємо значення лічильника на 1 та повертаємося на крок 3.

9. Якщо декодування успішне, отримуємо значення векторів e' та I'.

10. Обчислюємо значення векторів $e = e' \cdot D \cdot P$ та $I = I' X^{-1}$.

Результатом виконання алгоритму є формування підпису повідомлення, що складається зі значення лічильника, для якого вдало застосовано алгоритм декодування, вектору помилок та інформаційного вектору:

$$Y = (I, e, i) : IG_X + e = \left(h(h(M) \| i)\right) \quad (5)$$

Для запропонованої схеми на вхід декодеру поступає кодове слово з помилками, що має довжину n біт, а звідси загальна кількість n-бітних векторів визначається як 2^n.

При цьому відомо, що складність обчислення векторів I та e по відомому геш-значенню $h(h(M) \| i)$ для криптоаналітика визначається як NP-повна задача.

Для верифікації підпису є необхідним обчислення двох векторів: $c_X^{*}{}' = IG_X + e$ та $c_X^{*}{}'' = h(h(M) \| i)$. Якщо $c_X^{*}{}' = c_X^{*}{}''$ і вага Хеммінга вектора e не перевищує виправляючої здатності коду $w(e) \le t$, то підпис вважається справжнім. У іншому випадку, при невиконанні однієї з умов, робиться висновок, що підпис було змінено.

Таким чином згідно запропонованої схеми електронного цифрового підпису процедура верифікації відрізняється від схеми CFS шляхом додавання ще однієї умови, яка забезпечує альтернативну схему перевагами над CFS, що будуть розглянуті надалі.

Варто зазначити, що оцінка складності перевірки згідно схеми CFS справедлива по відношенню і до альтернативної схеми. При цьому в оцінках обох схем не враховуються обчислювальні затрати на роботу геш-функції та зняття дії маскуючих матриць.

4. ПОРІВНЯННЯ ЕФЕКТИВНОСТІ КОДОВИХ СХЕМ ЕЦП

Порівняльну характеристику схеми CFS та запропонованої схеми доцільно буде провести з використанням декількох критеріїв:

- складність формування та перевірки ЕЦП;
- стійкість ЕЦП;
- довжина підпису;
- обсяги ключових даних.

Проведемо аналіз поступово за всіма пунктами.

Для обох схем найбільш затратним є етап формування підпису, який полягає в успішному декодуванні синдромної послідовності, тому проведемо оцінку кількості спроб декодування для кожної зі схем.

Процес декодування тісно пов'язаний з поняттям синдромної послідовності. У випадку використання CFS ця послідовність має довжину $n-k$. Синдромна послідовність є бітовою послідовністю, тому може складатися лише зі значень 0 та 1. При чому кожне наступне значення не залежить від попереднього. Отже, загальна кількість можливих синдромних послідовностей дорівнює 2^{n-k}.

Відомо, що виправляюча здатність алгебраїчного двійкового блокового (n,k,d) коду визначається як $t = \left\lfloor \frac{d-1}{2} \right\rfloor$, якщо під час передачі відбудеться кількість помилок, що не перевищує значення *t,* усі ці помилки будуть гарантовано виправлені. Отже, кількість синдромних послідовностей, для яких декодування гарантовано пройде успішно можна визначити як:

$$N = \sum_{i=0}^{t} C_n^i . \quad (6)$$

Припустимо, що геш-значення $h(h(M) \| i)$ формуються рівно ймовірно, тоді ймовірність успішного декодування під час виконання алгоритму можна визначити [10-11] :

$$P_{y.\partial.} = \frac{\sum_{i=0}^{t} C_n^i}{2^{n-k}} . \quad (7)$$

У випадку використання двійкових сепарабельних кодів Гоппи, що задовольняють умовам $n = 2^m$, $k = n - mt$, $t = deg\ G(x)$, $d \ge 2t + 1$, можна скористатися апроксимацією:

$$P_{y.\partial.} = \frac{\sum_{i=0}^{t} C_n^i}{2^{n-k}} \approx \frac{\frac{n^t}{t!}}{n^t} = \frac{1}{t!} . \quad (8)$$

Звідси маємо, що здійснивши в середньому $t!$ спроб, декодування буде успішним. Кожна подібна спроба потребує $t^2 \cdot m^3$ двійкових операцій, але цей результат є приблизним, оскільки не береться до

уваги витрати на формування геш-кодів та зняття дій маскуючи матриць. Звідси можна оцінити середню кількість бітових операцій, яких вимагає виконання алгоритму формування підпису згідно схеми CFS:

$$N_{ф.п.} = t^2 \cdot m^3 \cdot t!. \quad (8)$$

У випадку використання альтеративної схеми на вході декодеру опиняється кодове слово з помилками $\overline{c}^{*} = I' \cdot G + e'$, що має довжину n біт. Загальна кількість n-бітних векторів визначається як 2^n. Кодове слово складається з двох компонент: k-бітного вектору I' (він може приймати одне з 2^k значень) та вектору e', вага Хеммінга якого не перевищує виправляючої здатності коду, тому, за аналогією з CFS, цей вектор набуває одного з значень. Значення I' та e' не залежить одне від одного. З вищесказаного можна зробити висновок, що кодове слово може набувати одне з $2^k \cdot N$ значень. Отже, ймовірність успішного декодування одиничного випадку можна визначити як:

$$P_{у.д.} = \frac{2^k \cdot \sum_{i=0}^{t} C_n^i}{2^n} = \frac{\sum_{i=0}^{t} C_n^i}{2^{n-k}}. \quad (9)$$

Очевидно, що дана оцінка збігається з оцінкою успішного декодування у схемі CFS. Звідси можна здійснити апроксимацію з урахуванням кодових відношень кодів Гоппа і отримати оцінку:

$$N_{ф.п.} = t^2 \cdot m^3 \cdot t!. \quad (10)$$

Наступним кроком оцінимо складність перевірки підпису згідно обох схем. При використанні схеми CFS необхідно обчислити геш-значення $h(h(M)\|i)$ і порівняти його з добутком $H_X \cdot e^T$. За умови, що складність хешування не враховується, як і в оцінці формування підпису, складність перевірки ЕЦП залежить тільки від кількості бітових операцій складання та множення при обчисленні добутку, тобто складність визначається як:

$$N_{п.п.} = (n-k) \cdot n = m \cdot t \cdot 2^m. \quad (11)$$

Ця оцінка справедлива по відношенню і до запропонованої схеми.

Отже, оцінки складності формування та перевірки підпису згідно схеми CFS та нової схеми є тотожними. При цьому в оцінках обох схем не враховуються обчислювальні затрати на роботу геш-функції та зняття дій маскуючи матриць [14, 25].

Як зазначалося раніше, схема CFS базується на використанні односторонньої функції з криптосистеми Нідеррайтера. Стійкість цієї функції до атаки, заснованої на переставному декодуванні, визначається як кількість кровельних множин, при яких можливо виправити всі комбінації з t помилок без знання секретного ключа [12]:

$$N \geq \frac{C_n^t}{C_{n-k}^t} = \frac{\frac{n!}{t!(n-t)!}}{\frac{(n-k)!}{t!(n-k-t)!}} = \frac{n!(n-k-t)!}{(n-t)!(n-k)!}. \quad (12)$$

З метою формування підпису $Y' = (e', i')$ для зміненого повідомлення M' зловмиснику потрібно реалізувати декодування випадкового коду в середньому $t!$ разів. Враховуючи цей факт, оцінку стійкості ЕЦП за схемою CFS можна визначити як:

$$N_c \geq t! \frac{C_n^t}{C_{n-k}^t} = t! \frac{n!(n-k-t)!}{(n-t)!(n-k)!} = \\ = t! \frac{2^m!(mt-t)!}{(2^m-t)!(mt)!}. \quad (13)$$

В ряді робіт виконано дослідження, що демонструють еквівалентність стійкостей криптосистем Мак-Еліса та Нідеррайтера. Звідси можна застосувати припущення, що стійкості схем CFS та запропонованої схеми, також еквівалентні.

У випадку використання квантового криптоаналізу оцінки стійкості обох схем набувають іншого характеру. Використовуючи один з найпопулярніших квантових алгоритмів, алгоритм Гровера, можна визначити кількість ітерацій для декодування випадкового коду, які потрібно виконати *t!* разів [13]:

$$C^{\frac{n}{2\log n}},\ C = \frac{1}{(1-R)^{1-R}}. \quad (14)$$

Припустимо, що квантовий алгоритм може бути застосовано для пошуку значення лічильника i, шляхом перебору значень (атака груба сила), що потребує в середньому $\frac{\pi}{4}\sqrt{t!}$ спроб. Отже, стійкість схем ЕЦП в умовах застосування квантових комп'ютерів можна визначити [14-15]:

$$N_{с.к.} \geq \frac{\pi}{4}\sqrt{t!}\left(\frac{1}{(1-R)^{1-R}}\right)^{\frac{n}{2\log n}} = \\ = \frac{\pi}{4}\sqrt{t!}\left(\left(1-\frac{k}{n}\right)^{\frac{k}{n}-1}\right)^{\frac{n}{2\log n}} = \\ = \frac{\pi}{4}\sqrt{t!}\left(1-\frac{k}{n}\right)^{\frac{k-n}{2\log n}} = \frac{\pi}{4}\sqrt{t!}\left(\frac{m \cdot t}{2^m}\right)^{t/2}. \quad (15)$$

Згідно схеми CFS підпис $Y = (e, i)$ містить дві складові: двійковий вектору *e*, що має довжину *n*, і ціле число i. Останнє може набувати значень в діапазоні $0,1,..,2^{n-k}-1$. Звідси маємо, що бітова довжина підпису визначається згідно виразу:

$$l_{ЕЦП} = 2 \cdot n - k = 2^m + m \cdot t. \quad (16)$$

Як зазначалося під час розгляду складності формування підпису, вектор e здатен набувати обмеженої кількості значень. Обмеження накладається згідно виправляючої здатності використовуваного коду. Кількість можливих векторів e визначається як [13]:

$$N_{w(e)\le t} = \sum_{i=0}^{t} C_n^i. \quad (17)$$

Через те, що вектор e відповідає умові вище, його можливо перетворити у беззбиткову послідовність $e*$ з довжиною $\lceil \log_2(N_{w(e)\le t}) \rceil$ біт. Тоді маємо:

$$l_{ЕЦП}* = \left\lceil \log_2 \left(\sum_{i=0}^{t} C_n^i \right) \right\rceil + n - k = \\ = \left\lceil \log_2 \left(\sum_{i=0}^{t} C_{2^m}^i \right) \right\rceil + m \cdot t. \quad (18)$$

Використовуючи вираз для верхньої границі Хеммінга, вираз можна перетворити:

$$l_{ЭЦП}* \le \left\lceil \log_2 \left(2^{n-k} \right) \right\rceil + n - k = 2 \cdot m \cdot t. \quad (19)$$

У випадку використання запропонованої схеми складових підпису $Y = (I, e, i)$ стає більше: вектор I (довжина k біт), вектор e та цілого числа i, довжина яких визначається так же, як і у схемі CFS. Звідси маємо, що довжина підпису $Y = (I, e, i)$ визначається згідно виразу:

$$l_{ЕЦП} = 2 \cdot n = 2^{m+1}. \quad (20)$$

Якщо здійснити беззбиткове перетворення вектора e, тоді цю оцінку можна переписати як:

$$l_{ЕЦП}* = \left\lceil \log_2 \left(\sum_{i=0}^{t} C_n^i \right) \right\rceil + n = \\ = \left\lceil \log_2 \left(\sum_{i=0}^{t} C_{2^m}^i \right) \right\rceil + 2^m. \quad (21)$$

Аналогічно розгляду CFS, використовуючи верхню границю Хеммінга маємо[16]:

$$l_{ЕЦП}* \le \left\lceil \log_2 \left(2^{n-k} \right) \right\rceil + n = m \cdot t + 2^m. \quad (22)$$

Схема CFS та її альтернатива побудовані на двох різних підходах: перша полягає у використанні функції зі схеми Нідеррайтера, друга-зі схеми Мак-Еліса, від яких прямо залежать обсяги ключових даних схем підпису. Відкритий ключ:

1. Довжина відкритого ключа CFS визначається кількістю комірок матриці $H_X = X \cdot H \cdot P$:

$$l_{в.к.} = (n-k) \cdot n = n^2 - kn = m \cdot t \cdot 2^m. \quad (23)$$

2. Довжина відкритого ключа альтернативної схеми визначається кількістю комірок матриці $G_X = X \cdot G \cdot P$:

$$l_{в.к.} = k \cdot n = (2^m - m \cdot t) \cdot 2^m. \quad (24)$$

Секретний ключ:

1. Довжина секретного ключа CFS визначається сумою кількості двійкових комірок матриці X (має розмір) та довжини n цілих чисел в діапазоні $0,1,..,n-1$ для визначення матриці P:

$$l_{с.к.} = (n-k)^2 + n \cdot \lceil \log_2 n \rceil = \\ = (m \cdot t)^2 + 2^m \cdot m. \quad (25)$$

2. Довжина секретного ключа альтернативної схеми визначається сумою кількості двійкових комірок матриці X (має розмір $k \times k$) та довжини n цілих чисел в діапазоні $0,1,..,n-1$ для визначення матриці P:

$$l_{с.к.} = k^2 + n \cdot \lceil \log_2 n \rceil = (2^m - m \cdot t)^2 + 2^m \cdot m. \quad (26)$$

Одразу слід зауважити, що в попередньому підрозділі було описано, що стійкість схем CFS та альтернативної схеми еквівалентні, що слідує з еквівалентності оцінок криптосистем Мак-Еліса та Нідеррайтера, на яких базуються вказані вище схеми підпису. Незважаючи на це, стійкість CFS та її альтернативи від деяких видів атак різниться. Цей факт докладніше буде розглянуто надалі. На сьогодні, на схему підпису CFS існує два типи атак: відновлення секретного ключа, та атака підробки підпису, яка полягає у намаганні створити дійсний підпис для повідомлення без знання секретного ключа. Атаки відновлення ключа проти схеми Мак-Еліса традиційно вважаються менш ефективними, ніж атаки дешифрування, і в даний момент не знайдено ефективної атаки цього типу у випадку використання кодів Гоппа. Через це в подальшому буде розглянуто атаки в контексті атак на алгоритми формування цифрового підпису-атаку підробки.

Для реалізації атаки підробки потрібно вирішити проблему декодування синдрому: знаходження вектору помилки найменшої ваги відповідно до наданого синдрому. Однак, на противагу стандартній проблемі декодування синдрому, тільки одна може бути вирішена як одна з багатьох [17].

Проблема декодування одного з багатьох: присутні параметри n,k,t та N, двійкова $k \times n$ матриця H і набір з N k-бітних двійкових векторів s_i. Потрібно знайти двійковий вектор помилок e з вагою Хеммінга t або менше таких, що задовольняють умові [24]:

$$\exists i \in [1,N] : e \cdot H^T = s_i . \qquad (27)$$

Існують дві найпопулярніші атаки на кодові криптосистеми: декодування множини даних (ISD) та загальний алгоритм, що базується на парадоксі «днів народжень» (GBA) [18]. ISD атака в умовах застосування принципу «декодування один з багатьох» здатна зменшити реальні затрати на виконання майже в $\sqrt{N}$ разів, де N - кількість доступних варіантів, порівняно з випадком, коли потрібно знайти один конкретний синдром. Атаки типу GBA, на відміну від атак ISD, здатні працювати лише у випадку наявності багатьох рішень. Ідею використання парадоксу «про дні народження» у контексті моделювання атак на схеми цифрового підпису належить Даніелю Блейхенбахеру. Він запропонував створювати список бажаних синдромів і використовувати цей список як один з початкових списків GBA. Порівняно з атакою ISD це дозволяє значно скоротити витрати на підробку підпису. Складність даної атаки відносно CFS з лічильником можна оцінити за наступною формулою:

$$C_{GBA} = L \log(L), \qquad (28)$$

де

$$L = \min\left(\frac{2^k}{C_n^{t-\lfloor t/3 \rfloor}}, \sqrt{\frac{2^k}{C_n^{\lfloor t/3 \rfloor}}}\right),$$

L - це довжина найбільшого списку, що використовується в алгоритмі. У випадку CFS з повним декодуванням, бажаним варіантом є знаходження слова з вагою $t+\delta$ і оскільки $C_n^{t+\delta} > 2^{mt}$, список, сформований із застосуванням операції XOR стовпчиків матриці H, є достатньо великим і розмір L найбільшого списку близький до значення $2^{\frac{mt}{3}}$.

Аналізуючи альтернативну схему, варто зауважити, що вона володіє вагомою перевагою перед CFS, оскільки здатна забезпечити захист від швидкої підробки підпису на основі додавання довільного кодового слова. Атака цього типу по відношенню до CFS може бути організована завдяки виконанню наступних дій:

– Обираємо довільне кодове слово c використовуваного (n,k,d) коду, перевірочна матриця якого позначається як H_x. В такому випадку справедлива рівність $H_x \cdot c^T = 0$. Маємо сформований підпис $Y = (e,i)$.

– Виконуємо додавання кодового слова:

$$\begin{aligned} Y &= (e+c,i) : H_x \cdot (e+c)^T = \\ &= H_x \cdot e^T + H_x \cdot c^T = \\ &= H_x \cdot e^T = (h(h(M) \| \mathrm{i}))^T . \end{aligned} \qquad (29)$$

Змінюючи останній вираз стосовно альтернативної схеми, отримаємо

$$Y = (I, e+\hat{c}, i) : IG_X + e + \hat{c} \neq \left(h(h(M) \| i)\right). \qquad (30)$$

тобто швидка підробка підпису у такому випадку неможлива. Також ця властивість посилена завдяки додатковій перевірці ваги Хеммінга вектору помилок під час процедури верифікації підпису. Це дозволяє захиститися також від інших гіпотетичних атак таких, як одночасна підробка двох елементів підпису і т.п.

5. ВИСНОВКИ

Стаття присвячена дослідженням способів цифрового підпису на основі кодів, що виправляють помилки. Стійкість таких схем ґрунтується на складності розв'язання теоретично складної проблеми розшифровки синдрому, що належить до класу задач NP-складних. З останнього факту випливає стійкість схем як до класичного криптоаналізу, так і до криптоаналізу з використанням квантових обчислень. У статті пропонується новий підхід до формування схеми цифрового підпису, що докорінно відрізняється від відомої схеми CFS. Однією із відмінністю запропонованої схеми є додавання нового елемента до підпису, це рішення дозволяє забезпечити захист від спеціальних видів атак, таких як одночасна підробка. Звичайно, додавання нового елемента збільшує остаточний розмір підпису, але таке збільшення не є критичним. Слід зазначити, що, незважаючи на відмінності в побудові підписів, обидві схеми забезпечують рівноцінний рівень захисту від класичного та квантового криптоаналізу. Однак варто зазначити, що недоліком схем підписів на основі коду є великий обсяг ключових даних, необхідних алгоритму, а також складність у створенні підпису через необхідність багаторазового розшифрування синдрому, що залишається актуальною темою і потребує подальших досліджень.

6. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] D. Moody, “Post-Quntum Cryptography: NIST’s Plan for the Future,” *The Seventh International Conference on PostQuntum Cryptography*, Japan, 2016.
[2] R. Overbeck, N. Sendrier, *Code-based cryptography*. in: Daniel J. Bernstein, et al. (eds), First International Workshop on Post-quantum Cryptography, PQ Crypto 2006, Leuven, The Netherland, May 23-26, 2006. Selected papers, pp. 95-145.
[3] *Post - quantum cryptography*, [Online]. Available: https://csrc.nist.gov/Projects/Post-Quantum-Cryptography
[4] D. Bernstein, J. Buchmann and E. Dahmen, *Post-Quantum Cryptography*, Springer-Verlag, Berlin-Heidleberg, 2009, 245 p.
[5] H. Niederreiter, “Knapsack-type cryptosystems and algebraic coding theory,” *Problem Control and Inform Theory*, v. 15, pp. 19-34, 1986.
[6] N. Courtois, M. Finiasz and N. Sendrier, “How to achieve a McEliece-based digital signature scheme”, *in Advances in Cryptology - ASIACRYPT 2001*, volume 2248, pp. 157-174.
[7] M. Finiasz, *Parallel-CFS: Strengthening the CFS McEliece-based signature scheme*, in Biryukov, A., Gong, G., Stinson, D., eds.: Selected Areas in Cryptography. Volume 6544 of LNCS., Springer, 2010, pp. 159-170.
[8] R. J. McEliece, *A public-key cryptosystem based on algebraic coding theory*, DSN Progress Report 42-44, Jet Propulsion Lab., Pasadena, CA, January-February, 1978, pp. 114-116.
[9] Yu.V. Stasev, A.A. Kuznetsov, “Asymmetric code-theoretical schemes constructed with the use of algebraic geometric codes,” *Kibernetika i Sistemnyi Analiz*, No. 3, pp. 47-57, May-June 2005.
[10] В. Д. Гоппа, “Новый класс линейных корректирующих кодов,” *Проблемы передачи информации*, т. 6, вып. 3, с. 24-30, 1970.
[11] В. Д. Гоппа, “На неприводимых кодах достигается пропускная способность ДСК,” *Проблемы передачи информации*, т. 10, вып. 1, с. 111-112, 1974.
[12] A. Kuznetsov, R. Serhiienko and D. Prokopovych-Tkachenko, "Construction of cascade codes in the frequency domain," *2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkov, 2017, pp. 131-136.
[13] A. Kuznetsov, I. Svatovskij, N. Kiyan and A. Pushkar'ov, "Code-based public-key cryptosystems for the post-quantum period," *2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkov, 2017, pp. 125-130.
[14] L. Grover, “A fast quantum mechanical algorithm for database search,” *Proceedings of the 28th annual ACM symposium on the theory of computing (STOC, 96)*, ACM Press, New York, 1996, pp. 212-219.
[15] L. Grover, “A framework for fast quantum mechanical algorithms,” *Proceedings of the 13th annual ACM symposium on theory of computing*, ACM Press, New York, 1998, pp. 53-62.
[16] Y. X. Li, R.H. Deng, X.M. Wang, *On the equivalence of McEliece’s and Niederreiter’s public-key cryptosystems*. [Online]. Available: https://ieeexplore.ieee.org/document/272496
[17] J. Stern, *A method for finding codewords of small weight*, in Cohen, G., Wolfmann, J., eds.: Coding theory and applications, Volume 388 of LNCS., Springer, 1989, pp. 106-113.
[18] N. Sendrier, “Decoding one out of many,” *in Yang, B.Y., ed.: PQCrypto 2011*. Volume 7071 of LNCS. Springer, 2011, pp. 51-67.

***Кузнецов Олександр Олександрович**, професор кафедри безпеки інформаційних систем і технологій Харківського національного університету імені В.Н.Каразіна, м. Харків, Україна. Сфера наукових інтересів: криптографія, теорія алгебраїчного кодування, пост-квантові криптоперетворення.*

***Кіян Анастасія Сергіївна**, здобувач кафедри безпеки інформаційних систем і технологій Харківського національного університету імені В.Н.Каразіна, м. Харків, Україна. Сфера наукових інтересів: криптографія, теорія алгебраїчного кодування, пост-квантові криптоперетворення.*

***Пушкарьов Андрій Іванович,** директор департаменту захисту інформації Державної служби спеціального зв’язку та захисту інформації України, м. Київ, Україна. Сфера наукових інтересів: інформаційна та кібербезпека, пост-квантові криптоперетворення.*

***Кузнецова Тетяна Юріївна,** науковий співробітник науково-дослідної частини Харківського національного університету імені В.Н.Каразіна, м. Харків, Україна. Сфера наукових інтересів: криптографія та теорія алгебраїчного кодування.*

DOI:
УДК 681.3.06

ПРОБЛЕМИ ЗАХИСТУ КРИТИЧНО ВАЖЛИВИХ ОБ'ЄКТІВ ІНФРАСТРУКТУРИ

Лукова-Чуйко Наталія[1]
orcid.org/0000-0003-3224-40614

Наконечний Володимир[2]
orcid.org/0000-0002-0247-5400

Толюпа Сергій[3]
orcid.org/0000-0002-1919-9174

Зюбіна Руслана[4]
orcid.org/0000-0002-8654-6981

м. Київ, Київський національний університет імені Тараса Шевченко
e-mai: [1]lukova@ukr.net, [2]nvc2006@i.ua, [3]tolupa@i.ua, [4]ziubina@knu.ua

Історія статті:
Надійшла до редакції 21.07.2019
Прийнято 29.07.2019

Ключові слова:
національна стратегія;
кібербезпека;
кібертероризм;
кіберзагроза;
критично важливий об'єкт інфраструктури

Анотація: *У роботі висвітлені актуальні питання, пов'язані із захистом критично важливих об'єктів інфраструктури, від функціонування яких залежить виживання людської спільноти. Показано, що стрімкий прогрес в сфері інформаційних технологій, з одного боку дозволяє сучасним економікам багатьох країн стати нерозривно взаємопов'язаними, з іншого боку цей процес піддає людство цілому ряду безпрецедентних загроз, які прагнуть дестабілізувати суспільне життя, втручаючись в роботу критично важливих об'єктів інфраструктури. Проведено аналіз того, що є істинними причинами такої пильної уваги до подібних об'єктів з боку кібертерористів. Представлені топ 10 основних загроз для промислових систем управління. Вказані кроки визначення відповідної критичності об'єктів інфраструктури. Наведено список 11 критично важливих секторів і 37 відповідних підсекторів, визначених Європейським Союзом. Запропоновано відповідь на питання, з якою метою країнам слід розробляти загальнонаціональні стратегії захисту критично важливих об'єктів інфраструктури. Запропонована стратегія кількісної оцінки рівня захищеності критично важливих об'єктів інфраструктури від ризику стороннього кібернетичного впливу. Обгрунтовано необхідність об'єднання державами різних елементів захисту критично важливих об'єктів інфраструктури. Зазначено, що політики в області кібербезпеки повинні займати центральне місце в захисті критично важливих об'єктів інфраструктури. Відзначено, що не всі національні стратегії кібербезпеки забезпечують однакове місце і значимість для критично важливих об'єктів інфраструктури. Зазначено, що при розробці національної стратегія щодо захисту критично важливих об'єктів інфраструктури, важливо скласти повний перелік всіх національних політик, що мають до неї відношення. Зроблено висновок про те, що є важливим інструментом для захисту критично важливих об'єктів інфраструктури, на сьогоднішній день.*

Abstract: *The paper deals with topical issues related to the protection of critical infrastructure, which depend on the survival of the human community. It is shown that rapid progress in the field of information technology, on the one hand, allows the modern economies of many countries to become inextricably interconnected; interfering with the work of critical infrastructure. The analysis of what are the true causes of such close attention to such objects by cyber terrorists. Top 10 major threats to industrial control systems are presented. The following steps determine the criticality of the infrastructure objects. The following is a list of 11 critical sectors and 37 relevant sub-sectors identified by the European Union. The answer is given to the question with which target countries should develop national strategies for the protection of critical infrastructure objects. A strategy for quantifying the level of security of critical infrastructure against the risk of third-party cybernetic exposure is proposed. The necessity to unite different elements of protection of critical infrastructure objects by states is substantiated. It is important that cybersecurity policies should be central to the protection of critical infrastructure. It is noted that not all national cybersecurity strategies provide the same place and relevance for critical infrastructure. It is noted that when developing a national strategy for the protection of critical infrastructure, it is important to compile a comprehensive*

list of all relevant national policies. It is concluded that today it is an important tool for the protection of critical infrastructure.

1. ВСТУП

Завдяки стрімкому прогресу в сфері інформаційних технологій (ІТ) сучасні економіки багатьох країн стали нерозривно взаємопов'язаними, що підвергає людство низці безпрецедентних загроз і чинників уразливості. Багато з них походять від терористичних груп і хакерів які прагнуть дестабілізувати суспільне життя, втручаючись в активи і процеси, від яких залежить виживання людської спільноти. Такі активи і процеси є центральними вузлами, відомими як критично важливі об'єкти інфраструктури (КВОІ).

Заміщення фізичного світу цифровим і настання епохи life 3.0, з одного боку, дозволяє людству здійснювати контроль інфраструктури з будь-якої точки світу, але з іншого боку, з'являються фактори уразливості до погроз. Це означає, що при високій взаємозалежності інфраструктури на рівні секторів і галузей, «вдала» атака на КВОІ з боку кібертерористів може привести до порушення або зруйнування життєво важливих систем безпосередньо в самій країні і, за ланцюговою реакцією, у всьому світі.

На думку групи урядових експертів з досягнень в області інформатизації і телекомунікації, в контексті міжнародної безпеки, «використання ІТ в терористичних цілях, крім вербування, фінансування, навчання і підбурювання, в тому числі для терористичних атак проти ІТ або ІТ-залежної інфраструктури, збільшує ймовірність того, що, якщо це залишити без уваги, може виникнути загроза загальному миру і безнебезпеці» [1].

Дана обставина є важливим фактором для кібертерористів і поряд з тим, що інфраструктура багатьох країн розвивається колосальними темпами, зброя кібертерористів також стає все більш витонченою. У зв'язку з цим, захист критично важливих об'єктів інфраструктури від навмисних кіберінцидентів зі сторони окремих осіб, організацій або країн є надзвичайно актуальною проблемою з якою зіткнулося людство XXI сторіччя.

2. ПОСТАНОВКА ПРОБЛЕМИ

Народження захисту КВОІ стало наслідком терористичних подій 11 вересня 2001 року. І хоча за своєю природою кіберзагрози відрізняються від фізичних погроз, кінцевий результат може бути однаково сумним.

Загрози щодо КВОІ можуть бути або окремими або випадковими діями, або частиною більш широкого плану по атаці на інфраструктуру в одному і тому ж секторі (наприклад, дамби, греблі, гідроелектростанції). Виявлення взаємозв'язків і закономірностей в подібних сценаріях часто вимагає наявності серйозних аналітичних інструментів і обробки інформації з великих і різнорідних джерел.

Однак, на жаль, інформація про більшість кібератак не публікується. Це зазначає ОБСЄ в «Керівництві по передовій практиці щодо захисту неядерної критично важливою енергетичної інфраструктури від терористичних атак з акцентом на загрози, які виходять з кіберпростору», щоб ще більше ускладнити ситуацію щодо енергетичного сектора, інформація про більшість кібератак не публікується, оскільки відповідні оператори не хочуть повідомляти про ці інциденти. Проте, здатність розпізнавати основну динаміку і методи якомога раніше є ключовим фактором, що дозволяє владі обмінюватися оперативною інформацією. Це підвищує здатність ефективніше реагувати на поточні атаки і запобігати неминучі атаки проти ймовірних жертв [2].

3. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Дотримання принципів безпеки дозволить сформувати низку вимог до системи кібербезпеки КВОІ, уникнути проблем пов’язаних з занадто високою вартістю реалізації цих вимог, неможливістю ефективного контролю за їх виконанням та важкістю їх засвоєння виконавцями, а також повністю нівелювати ризики усіх можливих загроз для КВОІ. Разом з цим саме контроль за виконанням вимог до кібербезпеки може привести, як відомо, до бажаного результату. Так, наприклад, якщо частина вимог не може бути реалізованою через брак коштів – саме ефективний контроль дозволить раціонально розподілити вирішення проблем, що виникають у процесі створення інформаційних систем КВОІ та її експлуатації. Відсутність ефективного контролю за виконанням вимог з боку виконавців може привести до оптимізації роботи останніх за рахунок ігнорування вимог по безпеці тощо.

У системах, де організований всеосяжний контроль усіх вимог, окрім величезної кількості інструкцій, наказів, законодавчих і підзаконних актів, у яких вони формулюються, як правило складаються списки вимог до підконтрольних структур і процесів. Наявність такого роду списків дозволить як контролерам, так і виконавцям швидше осягнути усю систему вимог, швидше їх освоїти й, відповідно, краще їх виконувати.

У деяких випадках то, що виглядає як окрема атака спрямована на відносно «неважливі» цілі, насправді, може бути частиною більш амбітних кримінальних стратегій [3].

Аналізуючи питання, що ж є причинами такої пильної уваги до КВОІ з боку кібертерористів,

можна припустити, що по-перше, вони можуть бути «привабливою метою» через свою стратегічну цінність для суспільства, особливо в високорозвинених індустріальних країнах. Втручання в функціонування КВОІ, з можливістю генерування каскадних ефектів, дозволить кібертерористам максимізувати збиток «вдалою» кібератакою і насаджувати страх до рівнів, які не могли б бути досягнуті, напавши на «рядові» об’єкти.

Інші КВОІ можуть бути «привабливими» для кібертерористів демонстрацією безсилля державних установ.

Третьою мотивацією, пов'язаною з двома попередніми, є бажання отримати більш високий рівень суспільного розголосу, що неможливо було б здійснити, зосередивши увагу на «звичайних» об'єктах.

Однак, як це не парадоксально, кібертерористи можуть домагатися контролю над КВОІ не для того, щоб завдавати шкоди або залякувати, а за зовсім іншою причиною - бажання встановити власну соціальну значущість.

В окремих випадках існує сукупність факторів, що спонукають кібертерористів здійснювати напади на КВОІ.

Остаточне рішення щодо цільової інфраструктури буде залежати від оперативних можливостей кібергрупи для запуску конкретної кібератаки. І захисні заходи, прийняті на певному КВОІ, природно, будуть, впливати на таке рішення., зосередивши увагу на «звичайних» об'єктах.

У той час як взаємопов'язані і інтегровані комп'ютеризовані системи управління істотно спростили методи роботи КВОІ і розширили можливості їх мережевої взаємодії, стало також можливим збільшення площі кібератаки і, отже, зросла вірогідність піддати КВОІ високому ризику порушення їх функціонування.

Відомі, як мінімум, 10 основних загроз для промислових систем управління, які представлені в таблиці 1 [2].

Таблиця 1. Топ 10 загроз для промислових систем управління

№	Загроза	Пояснення
1	Несанкціоноване використання точок доступу віддаленого обслуговування.	Точки доступу для технічного обслуговування - це спеціально створені зовнішні входи в мережу ІТ, які часто недостатньо захищені.
2	Мережеві атаки через офісні або корпоративні мережі.	Офісні ІТ зазвичай пов'язані з мережею декількома способами. У більшості випадків мережеві з'єднання з офісів в мережу ІТ також існують, тому зловмисники можуть отримати доступ по цьому маршруту.
3	Атаки на стандартні компоненти, що використовуються в мережі ІТ.	Стандартні ІТ-компоненти (комерційні готові продукти (COTS)), такі як системне програмне забезпечення, сервери додатків або бази даних, часто містять недоліки або уразливості, які можуть бути використані зловмисниками.
4	Атаки на мережеві компоненти.	Зловмисники можуть маніпулювати мережевими компонентами, наприклад, для проведення атак «зловмисник в середині» або для полегшення прослуховування.
5	Атаки типу "відмова в обслуговуванні".	Розподілені атаки типу "відмова в обслуговуванні" можуть послабити мережеві з'єднання і найважливіші ресурси і привести до відмови систем, наприклад, для порушення роботи ІТ.
6	Людська помилка і саботаж.	Умисні дії, вчинені як внутрішніми, так і зовнішніми зловмисниками, являють собою серйозну загрозу для будь-яких цілей захисту. Недбалість і людська помилка також є більшою загрозою, особливо щодо конфіденційності та доступності цілей захисту.
7	Впровадження шкідливого ПО через знімні носії і зовнішнє обладнання.	Використання знімних носіїв і мобільних ІТ-компонентів зовнішнього персоналу завжди тягне за собою великий ризик зараження шкідливим ПЗ.
8	Читання і опублікування новин в мережі ІТ.	Більшість компонентів управління в даний час використовують протоколи з відкритим текстом, тому зв'язок не захищений. Це дозволяє відносно легко читати і вводити команди управління.
9	Несанкціонований доступ до ресурсів.	Внутрішні зловмисники і наступні атаки після початкового зовнішнього проникнення роблять це особливо простим, якщо служби та компоненти в мережі процесів не використовують методи аутентифікації і авторизації або якщо ці методи небезпечні.
10	Форс-мажорні обставини або технічні несправності.	Перебої, викликані екстремальною погодою або технічними несправностями, можуть відбутися в будь-який час - ризик і потенційний збиток можуть бути зведені до мінімуму тільки в момент виникнення таких випадків.

У процесі захисту КВОІ надзвичайно важливе розуміння того, які об'єкти інфраструктури є критично важливими. Так в Резолюції Ради Безпеки 2341 (2017) прямо говориться про те, що «кожна держава визначає, що становить його критично важливі об'єкти інфраструктури».

Проте, вона не рекомендує будь-який конкретний метод вибору для виділення КВОІ серед безлічі об'єктів інфраструктури, розташованих на їх територіях.

Інші міжнародні документи також не дають ніяких вказівок. Таким чином, країнам надається більша свобода дій при виборі критеріїв для визначення того, які об'єкти інфраструктури, що діють на їх території, відповідають порогу «критичності». Завдання досить складне.

Різниця між значущими об'єктами інфраструктури і тими, які повинні отримати статус «критично важливих», є ключовим фактором, що дозволяє розставити пріоритети щодо обмежених ресурсів для захисту величезних активів, систем і процесів.

З одного боку, включення занадто великої кількості інфраструктур в категорію «критично важливих» є фінансово нестійким і може стати некерованим завданням. З іншого боку, надто обмежувальний підхід може залишити ряд ключових активів і процесів незахищеними з потенційно катастрофічними наслідками в разі аварії.

Існує тенденція урядів розширювати досить вузькі національні списки КВОІ тому, що «занадто мало осіб, які приймають рішення, готові прийняти політичний ризик, який може виникнути при видаленні елемента з «критично важливого» списку, і виникає спокуса постійно розширювати коло об'єктів, які вважаються критично важливими.

Двозначність марнотратна, оскільки ресурси не направляються туди, де вони можуть надати найбільший вплив [4]

На сьогоднішній день Україна також зіткнулася з проблемою відбору своїх КВОІ. При цьому, в жодному із законодавчих документів України не прописано процедуру процесу ідентифікації КВОІ, тому, для відповіді на це питання можна звернутися до світового досвіду. Отже спробуємо розібратися в цьому процесі.

Першим кроком в процесах ідентифікації КВОІ зазвичай є прийняття всеосяжного визначення того, що мається на увазі під КВОІ. Це корисно для створення фундаменту, на якому будуть розроблятися подальші політичні та нормативні рамки.

В цілому, в той час як визначення деяких країн підкреслюють призначення інфраструктури (тобто критичність пов'язана з виконанням основних соціальних функцій), інші роблять упор на наслідки руйнування або збою (тобто критичність обумовлена конкретними наслідками переривання обслуговування).

КВОІ можуть бути визначені, серед іншого, з урахуванням тієї ролі, яку вони грають в просуванні і захисті прав людини, а також вплив на права людини, яке може статися через пошкодження, порушень або руйнувань об'єктів інфраструктури.

Такий підхід відповідає існуючим визначенням. Наприклад, Європейський Союз (ЄС) визначає КВОІ як «актив, систему або її частину», яка «необхідна для підтримання життєво важливих функцій суспільства, здоров'я, безпеки, збереження, економічного або соціального благополуччя людей», порушення або руйнування яких матиме значний вплив «в результаті недотримання цих функцій» [5].

Цей крок в Україні вже майже зроблений. В проекті закону України “Про критичну інфраструктуру та її захист” в пунктах 10 та 11 статті 1 цього проекту дається визначення критичній інфраструктурі та об’єкту критичної інфраструктури відповідно.

Критична інфраструктура - об’єкти, які є стратегічно важливими для економіки і національної безпеки, порушення функціонування яких може завдати шкоди життєво важливим національним інтересам.

Об’єкт критичної інфраструктури - визначений у встановленому законодавством порядку складовий елемент критичної інфраструктури, функціональність, безперервність, цілісність і стійкість якої забезпечують реалізацію життєво важливих національних інтересів.

Одним із суб’єктів, що забезпечують захист КВОІ є діяльність РНБО України, яка повинна координувати роботу правоохоронних та інших органів у сфері кібербезпеки та, наприклад, її координаційний центр ніяк не може запрацювати на повну тільки тому, що у самому РНБО ніяк не вдається скоординувати блоки захисту, інформаційні та силові блоки в один вузол, щоб швидше реагувати на кіберзагрози.

Також є проблеми й в такому важливому процесі, як співробітництво спеціальних структур з приватним бізнесом, який має власний досвід і напрацювання протидії кіберзагрозам.

Силовики нерідко зловживають своїми повноваженнями створюючи перешкоди для бізнес-партнерів. Від цього довіра один до одного тільки втрачається і вже ні про яке співробітництво не може йти мова. А звідси – і про посилення кібербезпеки держави.

В свою чергу, на державному рівні питання кіберзахисту повинне набути системного характеру. Можна довго говорити про важливість і актуальність посилення регулювання на національному та міжнародному рівнях діяльності в кіберпросторі і зростання ролі в цьому приватного сектора; встановлення контролю над кіберзброєю, а також посилення охорони критичної інфраструктури України; впровадження інновацій в сфері кібербезпеки та активізації дискусій щодо пошуку нових джерел фінансування засобів кібербезпеки; вдосконалення освітніх напрямів підготовки фахівців даної сфери діяльності; активізацію на національному та міждержавно-

му рівнях обміну інформацією про кібератаки, тощо.

Однак без системного та комплексного підходу всі зазначені кроки не дозволять вивести рівень кібербезпеки, а звідси – і національної безпеки України загалом, на новий якісний рівень.

Другий етап в ідентифікації КВОІ є найбільш складним, з огляду на те, що саме тут відбувається «розстановка пріоритетів». Цей етап, як правило, спрямований на виявлення секторів, які розглядаються як критично важливі.

Початковий підхід міг би полягати в розгляді інших країн, які мають схожі соціальні, географічні особливості, а також порівняний рівень технічного і економічного розвитку.

Ряд секторів, швидше за все, будуть розглядатися як «критично важливі» у всіх країнах. При цьому важливо відзначити, що певний сектор або підсектор може мати вирішальне значення для однієї конкретної сфери, але не для іншої. Розмір і особливості певної національної економіки можуть визначити, що є критичним, а що менш критичним. Крім того, той факт, що певний сектор позначений як критично важливий, не повинен автоматично означати, що всі базові служби є критично важливими. Беручи до уваги ці відмінності, країни в значній мірі приходять до аналогічних висновків.

У таблиці 2 наведено список з 11 секторів і 37 відповідних підсекторів, визначених ЄС [6].

Таблиця 2. Орієнтовний список секторів і підсекторів, визначених ЄС

1	Енергетика	1 Видобуток, переробка, обробка та зберігання нафти і газу, включаючи трубопроводи. 2 Виробництво електроенергії. 3 Передача електроенергії, газу та нафти. 4 Розподіл електроенергії, газу та нафти.
2	Інформаційні, комунікаційні технології.	5 Інформаційна система і захист мережі. 6 Інструменти автоматизації і управління системами. 7 Інтернет. 8 Надання фіксованих телекомунікаційних послуг. 9 Надання мобільного зв'язку. 10 Радіозв'язок і навігація. 11 Супутниковий зв'язок. 12 Телерадіомовлення.
3	Вода.	13 Забезпечення питною водою. 14 Контроль якості води. 15 Розробка свердловин і контроль кількості води.
4	Їжа.	16 Забезпечення продовольством і продовольчої безпеки.
5	Охорона здоров'я.	17 Медична і лікарняна допомога. 18 Ліки, сироватки, вакцини та фармацевтика. 19 Біолабораторії й біоагенти.
6	Фінансування.	20 Платіжні послуги / платіжні структури (приватні). 21 Державне фінансове асигнування.
7	Громадський та правовий порядок і безпека.	22 Підтримка громадського та правового порядку, охорони і безпеки. 23 Виконання правосуддя і утримання під вартою.
8	Громадська адміністрація.	24 Урядові функції. 25 Збройні сили. 26 Послуги цивільної адміністрації. 27 Служби порятунку. 28 Поштові та кур'єрські послуги.
9	Транспорт	29 Дорожній транспорт. 30 Залізничний транспорт. 31 Повітряне сполучення. 32 Водний транспорт у внутрішньому сполученні. 33 Морський та каботажний транспорт.
10	Хімічна і атомна промисловість.	34 Виробництво і зберігання/переробка хімічних і атомних речовин. 35 Трубопроводи небезпечних вантажів (хімічні речовини).
11	Космос	36 Космос. 37 Дослідження.

Третій крок пов'язує раніше встановлені сектори і підсектори зі списком окремих інфраструктурних активів, систем і процесів. Країни розробили безліч наборів показників для визначен-

ня певних інфраструктур як «критично важливих». Ці індикатори зазвичай прагнуть «виміряти» наслідки руйнування об'єктів інфраструктури або функціонального збою і включають комбінацію з географічного охоплення впливу, його тривалості, а також тяжкості потенційних наслідків з точки зору:

- економічних наслідків (вплив на ВВП, прямі і непрямі економічні втрати, чисельність зайнятого персоналу, податкові надходження);
- кількості жертв і масштаби евакуації населення;
- порушення державного управління; шкоди навколишньому середовищу.

Для обґрунтування третього кроку застосовують різноманітні методології. Так, наприклад, консорціум на чолі з голландською дослідницькою організацією, спробував згрупувати їх за трьома основними типами [7]:

1) підхід, заснований на послугах (наприклад, Швейцарія), де уряд ідентифікує критично важливі активи на основі галузевих критеріїв, що визначають порогові значення / кількісне вироблення рівня обслуговування активів, наприклад, кількість доставлених мегават;

2) підхід, заснований на операторі (наприклад, Франція), де завдання визначення того, які активи або послуги є критично важливими, залишається за окремими операторами КВОІ;

3) підхід, заснований на активах або гібридах (наприклад, у Великобританії), в якому використовуються елементи підходів, орієнтовані як на послуги, так на оператора.

Завдання на предмет оцінювання рівня захищеності КВОІ від ризику стороннього кібервпливу належать, як відомо, до класу багатокритеріальних [8]. Для їх колегіального рішення в умовах невизначеності і конфлікту серед існуючих методів математичного моделювання, методів формування та дослідження узагальнених показників якості з використанням графоаналітичного і ним подібних підходів, експертних методів вирішення складних завдань оцінювання та вибору будь-яких об’єктів, в тому числі спеціального призначення, а також аналізу та прогнозування ситуацій з великою кількістю значимих факторів, найбільш раціональними і визначальними є саме експертні методи. Вони дають можливість більш глибоко вивчити явища, які істотно впливають на рівень захищеності як держави в цілому, так і окремих об’єктів її інформаційної та кіберінфраструктури від впливу внутрішніх і зовнішніх кібернетичних втручань та загроз, виявити найбільш важливе та істотне у цих процесах, не опускаючи тих деталей і взаємозв’язків, без яких не може бути побудована модель досліджуваної проблеми.

Метою такої моделі є оцінювання готовності об’єктів інформаційної і кіберінфраструктури до безпечного функціонування в умовах стороннього кібервпливу та встановлення на підставі так званого “індексу кіберпотужності” вимог до власних систем кібербезпеки [9].

Індекс кіберпотужності ($G_{рів.захищ.}$) з точки зору експертів може бути обчислений за формулою:

$$G_{рів.захищ} = (\sum_{i=1}^{n}(g_i \times G)) \times 100\%$$

де g_i – вагові коефіцієнти категорій другого рівня ієрархії $G_{iфакт}$; n – число категорій.

Прийняття рішення щодо здатності держави протистояти кібератакам має здійснюватися за 100-бальною шкалою на підставі такого правила:

- якщо $90 < G_{рів.захищ} < 100$, то рівень захищеності держави від ризику стороннього кібервпливу вважається достатньо високим для підтримки безпечного функціонування об’єктів її інформаційної і кіберінфраструктури;
- якщо $45 < G_{рів.захищ} < 90$, то рівень захищеності держави від ризику стороннього кібервпливу вважається допустимим для підтримки безпечного функціонування об’єктів її інформаційної і кіберінфраструктури;
- якщо $G_{рів.захищ} < 45$, то рівень захищеності держави від ризику стороннього кібервпливу вважається недостатнім.

Таким чином, запропонована стратегія оцінки дає можливість одержати кількісну оцінку рівня захищеності КВОІ від ризику стороннього кібернетичного впливу, встановити вимоги до формування ними власних систем кібернетичної безпеки та розробити заходи спрямовані на підвищення їх результативності.

Підставою таким діям може слугувати виявлення відхилень від штатного режиму функціонування державних інформаційних ресурсів, ІТ систем і мереж, а також відповідних програмних і апаратних засобів [10].

Для захисту КВОІ країни повинні мати відповідну загальнонаціональну стратегію. Досвід показує, що більшість країн забезпечували заходи безпеки для своїх КВОІ задовго до того, як їх захист затвердився в якості самостійного політичного поля.

Відповідні захисні заходи, як правило, приймалися поетапно в формі нормативних актів, що охоплюють конкретні загрози. В окремих випадках державна політика досягала значного рівня складності, відповідаючи високим міжнародним стандартам. Наприклад, в атомній енергетиці Україна розробила сучасну і ефективну систему фізичного захисту ядерних об'єктів і матеріалів.

Виникає питання, для чого таким країнам слід розробляти загальнонаціональні стратегії захисту КВОІ, коли в них уже введені і функціонують подібні нормативні акти, політика і практика, що охоплюють практично всі, критично важливі сектора.

Відповідаючи на це питання, можна припустити, що основна причина цього полягає в тому, що в сучасних суспільствах безпека КВОІ стає все більш складним завданням. Взаємозалежність між секторами з потенціалом для каскадних ефектів в разі кібератак, аварій природного або техногенного походження вимагає необхідності контролювати загальну картину для ефективної координації дій щодо запобігання, реагування і відновлення між секторами.

Така всеосяжна стратегія спрямована, перш за все, на більш ефективний розподіл фінансових і людських ресурсів, раціоналізацію робочих потоків навколо заздалегідь визначених цілей. Це не означає, що загальнодержавні стратегії по захисту КВОІ повинні автоматично замінювати існуючі галузеві заходи захисту, особливо коли ці заходи виявилися успішними чи відповідають обов'язковим міжнародним нормативним рамкам.

Необхідно, щоб держави об'єднували різні елементи захисту КВОІ під загальною егідою і зробили їх частиною єдиної системи управління.

Оскільки захист КВОІ пов'язаний з декількома різними сферами діяльності, основними цілями загальнодержавної стратегії є:

- визначення організаційних структур;
- встановлення вимірюваних цілей і термінів;
- закладка основи для ефективного запобігання та управління інцидентами шляхом гармонізації завдання між різними областями політики.

Більшість країн, у тому числі ті, які не мають спеціальних стратегій захисту критично важливих об'єктів інфраструктури, розглядають питання пов'язані із захистом КВОІ в різних політичних нормативних документах, введених різними урядовими установами. Ці документи зазвичай включають в себе національні, в тому числі і кіберстратегії в області боротьби з тероризмом.

Хоча такі різні політики могли бути прийняті в різний час і різними державними установами, вкрай важливо, щоб вони стали частинами єдиного підходу до захисту КВОІ. Це вимагає визначення країнами:

- взаємодії між цими іншими політиками і спеціальною стратегією захисту КВОІ;
- ступінь, в якій ці інші політики або сама стратегія безпеки КВОІ були скориговані і оптимізовані для запобігання конфліктам та забезпечення загальної координації політики на національному рівні.

Політики в області кібербезпеки займають центральне місце в системі безпеки критично важливих об'єктів інфраструктури, оскільки забезпечують основу, в якій країни визначають цілі і засоби захисту КВОІ і критично важливих інформаційних інфраструктур.

Ряд регіональних інструментів пов'язують концепції кібербезпеки з КВОІ. Стратегія кібербезпеки ЄС 2013 року, відповідно до якої Європейська комісія взяла на себе зобов'язання «продовжувати свою діяльність, що проводиться спільним дослідницьким центром в тісній координації з владою держав-членів і власниками і операторами критично важливої інфраструктури, щодо виявлення уразливості європейських критично важливих об'єктів інфраструктури та стимулювання розвитку відмовостійких систем» [11].

Відповідно до директиви Європейського Союзу 2016/1148 про заходи щодо забезпечення високого загального рівня безпеки мережевих та інформаційних систем по всьому Союзу, держави-члени ЄС повинні призначити операторів послуг життєзабезпечення (OES) і ввести нові вимоги до безпеки і звітності для таких організацій.

З огляду на це, не всі національні стратегії кібербезпеки забезпечують однакове місце і «вагу» для КВОІ. Між країнами існують значні відмінності. Деякі стратегії були написані тільки з точки зору кіберзлочинності і, як правило, не беруть до уваги національні явища дестабілізації, міжгалузеві впливи і антикризове управління для КВОІ. Стратегії, написані з точки зору кібербезпеки, засновані на національній оцінці ризику, візьмуть ширшу перспективу, яка дасть місце для захисту КВОІ ... » [12].

Важливим інструментом, запропонованим міжнародним союзом електрозв'язку (МСЕ), є сховище національних стратегій з кібербезпеки (NSC). Воно включає в себе велику колекцію національних стратегій з кібербезпеки, чи то у формі одного або декількох документів, або як частина більш широких стратегій в сфері ІТ або національної безпеки [13].

З огляду на розмаїття підходів між різними існуючими стратегіями щодо захисту КВОІ і кібербезпеки, МСЕ в даний час очолює зусилля з різними глобальними учасниками по створенню загального довідкового керівництва національних стратегій з кібербезпеки.

Цей документ покликаний дати країнам чітке уявлення про мету і зміст національної стратегії з кібербезпеки, показати в загальних рисах існуючі моделі і ресурси, спрямовувати країни в процесі розробки своїх стратегій та їх оцінки [13].

4. ВИСНОВКИ

Проведений в даній роботі аналіз дозволяє зробити висновок про актуальність проблеми захисту КВОІ. При цьому показано, що зі збільшенням прогресу в сфері ІТ з'явилася можли-

вість сучасним економікам багатьох країн стати нерозривно взаємопов'язаними. Проте цей процес, в свою чергу, ставить перед людством цілу низку безпрецедентних загроз, які впливаючи на роботу КВОІ дестабілізують суспільне життя.

Проведений в роботі аналіз виявив справжні причини пильної уваги до КВОІ з боку кібертерористів. При цьому, виявлені як мінімум три основні причини такої «привабливості». В роботі представлені топ 10 основних загроз для промислових систем управління. Запропоновані кроки визначення відповідної критичності об'єктів інфраструктури і наведено список 11 секторів і 37 відповідних критично важливих підсекторів, визначених ЄС.

Запропонована стратегія кількісної оцінки рівня захищеності критично важливих об’єктів інфраструктури від ризику стороннього кібернетичного впливу.

Дана відповідь на питання, з якою метою країнам слід розробляти загальнонаціональні стратегії захисту критично важливих об'єктів інфраструктури та обґрунтовано необхідність об'єднання державами різних елементів захисту КВОІ під загальною егідою з метою зробити їх частиною єдиної системи управління.

Показано, що політики в області кібербезпеки повинні займати центральне місце в захисті КВОІ. При цьому зазначено, що не всі національні стратегії кібербезпеки забезпечують однакове місце і значимість для КВОІ.

Показано, що при розробці національної стратегія щодо захисту КВОІ, важливо скласти повний перелік всіх національних політик, що мають до неї відношення.

Зроблено висновок про те, що важливим інструментом для захисту КВОІ, на сьогоднішній день, є сховище національних стратегій з кібербезпеки, запропонований міжнародним союзом електрозв'язку.

5. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] GGE 2015, Report of the Group of Governmental Experts on Information and Telecommunication Events in the Context of International Security, General Assembly, [Online].: https://ccdcoe.org/sites/default/files/documents/UN-150722- GGEReport2015.pdf.

[2] OSCE 2013, Good Practices Guide on Non-Nuclear Critical Energy Infrastructure Protection (NNCEIP) from Terrorist Attacks Focusing on Threats Emanating from Cyberspace, , 2013, [Online], : www.osce.org/atu/103500?download=true.

[3] Ackerman 2007, Assessing Terrorist Motivation for Attacking Critical Infrastructures, Nonproliferation Research Center, Monterey Institute for International Studies, [Online]. https://e-reports-ext.llnl.gov/pdf/341566.pdf.

[4] Clemente, 2013, Cybersecurity and Global Interdependence: What is Critical? Chatham House, February 2013, [Online]. www.chathamhouse.org/sites/files/chathamhouse/pulic/Research/International%20Security/0213prcvber.pdf.

[5] Council Directive 2008/114 / EC, Life Safety, General Secretariat for Defense and National Security). Article 2.

[6] European Commission 2005, Green Paper on the European Program for the Protection of Critical Infrastructure, COM (2005) 576 total.

[7] RECIPE 2011, Good Practice Guide for PKI Policies and for Politicians in Europe, [Online]. https://Users/SM/Downloads/RECIPE_manual%20(1).pdf.

[8] Бурячок В.Л., Толюпа С.В., Толубко В.Б., Хорошко В.О. «Інфо-рмаційна та кібербезпека: соціотехнічний аспект» // Навчальний посібник. – К.: Наш формат, 2015. – 288с.

[9] Serhii Toliupa, Hanna Shvedova and Ivan Parkhomenko. Security and Regulatory Aspects of the Critical Infrastructure Objects Functioning and Cyberpower Level Assesment. 3rd IEEE International Conference On Advanced Information and Communication Technologies (AICT) – 2019. Lviv, Ukraine. Scopus/

[10]. Бурячок В.Л. Основи формування державної системи кібернетичної безпеки: Монографія. – К.: НАУ, 2013. – 432 с.

[11] European Commission 2013 bis. Working Paper on a New Approach to the European Critical Infrastructure Protection Program - Building a Safer European Critical Infrastructure, SWD (2013) 318 total/ [Online]. https://ec.europa.eu/energy/sites/ener/files/documents/20130828_epcip_commission_staff_working_document.pdf

[12] The GFCE-MERIDIAN Good Practice Guide on Critical Information Infrastructure Protection for governmental policy-makers, [Online]. https://www.meridianprocess.org/siteassets/meridian/gfce-meridian-gpg-to-ciip.pdf

[13] National Cybersecurity Strategies Repository, [Online]. https://www.itu.int/en/ITU-D/Cybersecurity/Pages/National-Strategies-repository.aspx.

Лукова-Чуйко Наталія Вікторівна
Доктор технічних наук , доцент, професор кафедри кібербезпеки та захисту інформації Київського національного університету імені Тараса Шевченка.
Інтереси: інформаційні технології, криптографія, інформаційна та кібербезпека.
Наука: автор більше 60 наукових праць.

Наконечний Володимир Сергійович
Доктор технічних наук, старший науковий співробітник, професор кафедри кібербезпеки та захисту інформації Київського національного університету імені Тараса Шевченка.
Інтереси: системи розпізнавання багатопозиційного базування, системи технічного захисту інформації, кібербезпека та кіберзахист.
Наука: автор більше 130 наукових праць.

Толюпа Сергій Васильович
Доктор технічних наук, професор, професор кафедри кібербезпеки та захисту інформації Київського національного університету імені Тараса Шевченка.
Інтереси: інтелектуальні системи управління, напрямки підвищення ефективності функціонування інформаційних технологій, системи технічного захисту інформації, кібербезпека та кіберзахист..
Наука: автор більше 160 наукових праць.

Зюбіна Руслана Віталіївна,
Вищу освіту здобула у Національному авіаційному університеті у 2013 році. , Зараз працює асистентом кафедри кібербезпеки та захисту інформації факультету інформаційних технологій Київського національного університету імені Тараса Шевченко. Наукові інтереси: інформаційна безпека, кібербезпека, інформаційні технології.

DOI:
УДК 004.056.5

ВИМОГИ ДО МЕХАНІЗМІВ ЗАХИЩЕНОСТІ ОС В РАМКАХ КЛАСУ ВИКОРИСТАННЯ

Бичков Олексій [1]
orcid.org/0000-0002-9378-9535

Шестак Яніна [2]
orcid.org/0000-0002-1703-0316

[1] Україна, м. Київ, Київський національний університет імені Тараса Шевченка,
e-mail: bos.knu@gmail.com[1], lucenko.y@ukr.net.gna5@ukr.net[2]

Історія статті:
Надійшла до редакції
Прийнято

Ключові слова:
захищеність;
операційна;
система;
класифікація;
вимоги;
рівень;
захищенність

Key words:
security;
operating room;
system;
classification;
requirements;
level;
security

Анотація: Одним з компонент сучасних інформаційних систем є комп'ютерна техніка загального призначення, на якій встановлене спеціалізоване програмне забезпечення, або з якої ведуть роботу із спеціалізованим програмним забезпеченням. Для роботи цієї комп'ютерної техніки необхідна операційна система загального призначення. Для захисту інформаційної системи необхідно захистити усі її ланки, в тому числі операційну систему. Зокрема, у випадку клієнт-серверних технологій слід приділити увагу як захисту серверної компоненти, так і захисту клієнтської компоненти. Захист операційної системи вимагає розуміння, в рамках якої інформаційної системи буде працювати ОС, які умови роботи та обмеження накладаються на захищене середовище, захист від яких загроз буде потрібен захищеному середовищу, якими механізмами можна забезпечити потрібну нам захищеність середовища та яка буде «ціна» їх застосування чи незастосування та інше. Тому безпека програмного забезпечення має спиратися на політику безпеки операційної системи, розширюючи та уточнюючи її та наскільки це є можливим. При розробці програмного забезпечення слід спиратися на механізми безпеки, які передбачені операційною системою або інформаційної системою. Це потрібно для уніфікації та спрощення системи безпеки, полегшення її обслуговування, за рахунок зменшення кількості механізмів, які створені для розв'язання однієї і тієї ж задачі. Також слід застосовувати типові та загальновідомі компоненти та стандарти, по можливості уникати компонентів з закритим вихідним кодом, або компонентів, які не підтримують, або некоректно підтримують типові стандарти.

Очевидно, що неправильний захист операційної системи може привести до провалу системи безпеки в цілому, бо робота спеціалізованого програмного забезпечення та робота з периферійними пристроями відбувається саме під контролем операційної системи.

В статті детально розглянуті пропозиції щодо вимог до механізмів захисту операційних систем. Ці вимоги орієнтовані не на універсальні потреби, а враховують різні варіанти використання комп'ютерних пристроїв.

Abstract: One of the components of modern information systems is general-purpose computer equipment on which specialized software is installed or on which specialized software is working. This computer equipment requires a general-purpose operating system. To protect the information system it is necessary to protect all its links, including the operating system. In particular, in case of client-server technologies it is necessary to pay attention both to protection of server component and to protection of a client component. Protection of an operating system demands understanding, within what limits of what information system the OS will work, what working conditions and restrictions are imposed on the protected environment, what threats to the protected environment is required, what mechanisms is it possible to provide with protection of the environment necessary to us and what "price" of their application or nonapplication and other will be. Therefore, software safety should rely on a policy of safety of an operating system, expanding and specifying it and as much as possible. It is necessary to rely on safety mechanisms which are provided by an operating system or information system at software working out. This is necessary for unification and simplification of system of safety, simplification of its service, at the expense of reduction of quantity of the mechanisms created for the decision of the same problem. It is also necessary to apply typical and well-known components and standards, to avoid components with the closed source code as much as possible, or components which do not support, or incorrectly support typical standards.

Obviously, incorrect protection of the operating system can lead to the failure of the security system as a whole, because the operation of specialized software and work with peripherals is under the control of the operating system.

In article in detail offers concerning requirements to mechanisms of protection of operating systems are considered. These requirements are focused not on universal requirements, and consider various variants of use of computer devices-

1. ВСТУП

Сучасний світ - це світ інформаційних технологій, які увібрали в себе всю суть нашого буття. Інформаційна безпека є наріжним каменем загальної безпеки держави. Як відомо, хто володіє інформацією, той володіє світом.

Є багато думок про те, з чого починати будувати захист свого інформаційного простору, робочого місця, відділу, підприємства, держави [2-7]. Всіма інформаційно-телекомунікаційними системами керують комп'ютери. Комп'ютери працюють під управлінням операційних систем (ОС). Отже, фундаментом інформаційної безпеки є забезпечення безпеки на рівні операційної системи.

Авторами в [1] розроблена класифікація варіантів використання захищених операційних систем в інформаційно-телекомунікаційних системах.

Наведемо деякі визначення з [1].

Будемо вважати операційну систему захищено., якщо вона забезпечує збереження інформації і цілісність даних власними засобами без використання допоміжного програмного забезпечення [8].

2. МЕТА СТАТТІ

Метою статті є розроблення пропозицій щодо вимог до механізмів захисту операційних систем. Дана стаття є продовженням досліджень, що були запропановані в [1].

3. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

В [1] наведено таку класифікацію операційних систем за варіантами використання:

1. ОС для використання на робочій станції, не підключеній до мережі передачі даних.
2. ОС для використання на робочій станції, підключеній до внутрішньої мережі інтранет.
3. ОС для використання на робочій станції, підключеній до зовнішньої мережі інтернет.
4. ОС для використання на серверах, підключених у внутрішній мережі інтранет.
5. ОС для використання на серверах, підключених в зовнішній мережі інтернет.

У "Помаранчевій книзі" пропонується така класифікація рівнів захищеності операційних систем:

- високий (A), - середній (B), - умовний (C), - неприпустимий (D).

У відповідності до класифікації операційних систем за варіантами використання, сертифікат класу B1 буде позначати, що ОС використовується на робочих станціях, що не підключений до Інтернет і інтранет і має рівень захищеності «середній».

Будемо вважати рівень захищеності ОС «неприпустимим», якщо ОС не задовольняє хоча б одному пункту рівня захищеності С.

3.1. ПАРАМЕТРИ, ЗА ЯКИМИ МОЖНА ПРИСВАЮВАТИ СТУПІНЬ ЗАХИЩЕНОСТІ

Основною задачею захищеної операційної системи є запобігання несанкціонованому доступу до даних із зовнішніх (по відношенню до неї) джерел. Тому ми будемо класифікувати операційні системи, по особливостях алгоритмів управління ресурсами – локальні і мережеві ОС [9-11].

Локальні ОС управляють ресурсами окремого комп'ютера. Мережеві ОС беруть участь в управлінні ресурсами мережі.

Наведемо параметри, за якими необхідно надавати рівні захищеності.

Для класу використання 1.

Умовним рівнем захищеності буде називатися ОС, при наявності таких критеріїв:

- отримати права адміністратора безпеки можна тільки після виконання явних, що протоколюються дій;
- аутентифікаційна інформація повинна бути захищена від несанкціонованого доступу
- повинен бути реалізований механізм управління процесами в системі
- користувачі повинні пройти ідентифікацію і аутентифікацію перш ніж виконувати будь - які дії в системі
- управління системними сервісами і драйверами можливо тільки адміністратором або групою довірених осіб
- можливість локально заблокувати доступ до системи на час відсутності користувача
- процеси і об'єкти ядра не повинні мати можливість за власною ініціативою змінювати дані інших процесів або об'єктів ядра.
- повинно бути задано явне і недвозначне перерахування допустимих типів доступу (читання, запис, виконання, виконання з більш вищими привілеями) для кожного суб'єкта або групи по відношенню до об'єкта (ресурсу).
- повинен проводиться і записуватися моніторинг ключових подій в системі - контроль доступу повинен бути застосовний до кожного об'єкту і кожному суб'єкту.
- можливість змінювати правила доступу має надаватися виділеним суб'єктам (адміністраторам, довіреною групам).

Середнім рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного рівня):
- можливість відновлення системи після збою, ушкодження системних файлів, можливість створення контрольної точки відновлення
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання)
- вхід в систему під single-user повинен вимагати авторизації
- кожне реєстроване дія повинна асоціюватися з конкретним користувачем
- повинна здійснюватися очищення тимчасових системних файлів
- вся інформація про успішні і неуспішних спробах авторизації повинні записуватися в файл
- всі події в системі, які мають відношення до роботи системи

безпосередньо, до безпеки і до призначених для користувача додатків повинні фіксуватися і записуватися
- в системі повинні бути інтегрована можливість здійснювати криптографічно стійкий захист даних від несанкціонованого доступу

Високим рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного і середнього рівня):
- повинна здійснюватися очищення (обнулення, знеособлення) звільняються областей оперативної пам'яті ЕОМ і зовнішніх накопичувачів. Очищення здійснюється довільної записом в звільнену область пам'яті, раніше

використану для зберігання даних, що захищаються (файлів).
- архівування результатів моніторингу з метою економії займаного місця
- система повинна використовувати інтегроване шифрування файлової системи, при якому неможливо вважати дані при знятті накопичувача і підключення його до іншого комп'ютера
- система повинна використовувати інтегроване шифрування файлової системи, при якому неможливо вважати дані при знятті накопичувача і підключення його до іншого комп'ютера
- наявність механізму, який би на рівні ядра забороняв виконання коду з адресного сегмента даних
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання)
- при виявленні несанкціонованої модифікації системних файлів (при завданні відповідної опції адміністратором) проводиться автоматичне відновлення файлів з резервної копії.
- механізм встановлення паролів не повинен допускати установки слабких паролів менше 8 символів і мають осмислене значення.

Для класу використання 2.

Умовним рівнем захищеності буде називатися ОС, при наявності таких критеріїв:

- отримати права адміністратора безпеки можна тільки після виконання явних, що протоколюються дій;
- аутентифікаційних інформація повинна бути захищена від несанкціонованого доступу
- повинен бути реалізований механізм управління процесами в системі
- користувачі повинні пройти ідентифікацію і аутентифікацію перш ніж виконувати будь-які дії в системі
- управління системними сервісами і драйверами можливо тільки адміністратору або групі довірених осіб
- можливість локально заблокувати доступ до системи на час відсутності користувача
- процеси і об'єкти ядра не повинні мати можливість за власною ініціативою змінювати дані інших процесів або об'єктів ядра.
- повинно бути задано явне і недвозначне перерахування допустимих типів доступу (читання, запис, виконання, виконання з більш вищими привілеями) для кожного суб'єкта або групи по відношенню до об'єкта (Ресурсу).
- повинен проводиться і записуватися моніторинг ключових подій в системі
- контроль доступу повинен бути застосовний до кожного об'єкту і кожному суб'єкту.
- можливість змінювати правила доступу має надаватися виділеним суб'єктам (адміністраторам, довіреною групам).

Середнім рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного рівня):

- можливість відновлення системи після збою, ушкодження системних файлів, можливість створення контрольної точки відновлення
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання)
- вхід в систему під single-user повинен вимагати авторизації
- кожне реєстроване дія повинна асоціюватися з конкретним користувачем
- повинна здійснюватися очищення тимчасових системних файлів
- можливість управляти настроюванням підключення до мережі і параметрами підключення тільки адміністратору або групі довірених осіб
- вся інформація про успішні і неуспішних спробах авторизації повинна записуватися в файл
- всі події в системі, які мають відношення до роботи системи безпосередньо, до безпеки і до призначених для користувача додатків повинні фіксуватися і записуватися
- в системі повинні бути інтегрована можливість здійснювати криптографічно стійкий захист даних від несанкціонованого доступу

Високим рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного і середнього рівня):

- повинна здійснюватися очищення (обнулення, знеособлення) звільняються областей оперативної пам'яті ЕОМ і зовнішніх накопичувачів. очищення здійснюється довільної записом в звільнену область пам'яті, раніше використану для зберігання даних, що захищаються (файлів).
- архівування результатів моніторингу з метою економії займаного місця
- система повинна використовувати інтегроване шифрування файлової системи, при якому неможливо вважати дані при знятті накопичувача і підключення його до іншого комп'ютера
- система повинна використовувати інтегроване шифрування файлової системи, при якому неможливо вважати дані при знятті накопичувача і підключення його до іншого комп'ютера
- наявність механізму, який би на рівні ядра забороняв виконання коду з адресного сегмента даних
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання)

- можливість управління правами доступу до мережі для різних користувачів
- при виявленні несанкціонованої модифікації системних файлів (при завданні відповідної опції адміністратором) проводиться автоматичне відновлення файлів з резервної копії;
- наявність пакетного фільтра з можливістю створення правил для вхідного і вихідного мережевого трафіку з можливістю запису пакетів і переглядом мережевих з'єднань, а також правилами, що допускають вільне безперешкодне переміщення внутрішніх пакетів на локальні інтерфейси і адреса 127.0.0.1
- механізм встановлення паролів не повинен допускати установки слабких паролів менше 8 символів і мають осмислене значення.

Для класу використання 3

(ОС для використання на робочої станції, підключеної до зовнішньої мережі інтернет)

Умовним рівнем захищеності буде називатися ОС, при наявності таких критеріїв:

- отримати права адміністратора безпеки можна тільки після виконання явних, що протоколюються дій;
- аутентифікаційна інформація повинна бути захищена від несанкціонованого доступу;
- повинен бути реалізований механізм управління процесами в системі;
- користувачі повинні пройти ідентифікацію і аутентифікацію перш ніж виконувати будь-які дії в системі;
- управління системними сервісами і драйверами можливо тільки адміністратору або групі довірених осіб;
- можливість локально заблокувати доступ до системи на час відсутності користувача;
- процеси і об'єкти ядра не повинні мати можливість за власною ініціативою змінювати дані інших процесів або об'єктів ядра;
- повинно бути задано явне і недвозначне перелічення допустимих типів доступу (читання, запис, виконання, виконання з більш вищими привілеями) для кожного суб'єкта або групи по відношенню до об'єкта (ресурсу);
- повинен проводиться і записуватися моніторинг ключових подій в системі;
- контроль доступу повинен бути застосовний до кожного об'єкту і кожного суб'єкту;
- можливість змінювати правила доступу має надаватися виділеним суб'єктам (адміністраторам, довіреною групам).

Середнім рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного рівня):

- можливість відновлення системи після збою, ушкодження системних файлів;
- можливість створення контрольної точки відновлення;
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання);
- вхід в систему під single-user повинен вимагати авторизації;
- кожна реєстрована дія повинна асоціюватися з конкретним користувачем;
- повинно здійснюватися очищення тимчасових системних файлів;
- можливість управляти настроюванням підключення до мережі і параметрами підключення надається тільки адміністратору або групі довірених осіб;
- вся інформація про успішні і неуспішні спроби авторизації повинні записуватися в файл;
- всі події в системі, які мають відношення до роботи системи безпосередньо, до безпеки і до призначених для користувача додатків повинні фіксуватися і записуватися;
- в системі повинна бути інтегрована можливість здійснювати криптографічно стійкий захист даних від несанкціонованого доступу.

Високим рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного і середнього рівня):

- повинна здійснюватися очищення (обнулення, знеособлення) областей оперативної пам'яті ЕОМ і зовнішніх накопичувачів, що звільняються. Очищення здійснюється довільним записом в звільнену область пам'яті;
- архівування результатів моніторингу з метою економії займаного місця;

- система повинна використовувати інтегроване шифрування файлової системи, при якому неможливо зчитати дані при знятті накопичувача і підключення його до іншого комп'ютера;
- наявність механізму, який би на рівні ядра забороняв виконання коду з адресного сегмента даних;
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання);
- можливість управління правами доступу до мережі для різних користувачів;
- при виявленні несанкціонованої модифікації системних файлів (при завданні відповідної опції адміністратором) проводиться автоматичне відновлення файлів з резервної копії;
- наявність пакетного фільтру з можливістю створення правил для вхідного і вихідного мережевого трафіку з можливістю запису пакетів і переглядом мережевих з'єднань, а також правилами, що допускають вільне безперешкодне переміщення внутрішніх пакетів на локальні інтерфейси і адреси 127.0.0.1;
- підтримка стабільної роботи системного моніторингу протягом 15 хвилин атаки з повним фіксуванням дій атакуючого і паралельної відправкою на сервер резервних копій;
- система не повинна давати інформацію про своє найменування, версії і версій працюючих сервісів неавторизованих користувачів;
- механізм встановлення паролів не повинен допускати установки слабких паролів.

Для класу використання 4

Умовним рівнем захищеності буде називатися ОС, при наявності таких критеріїв:

- отримати права адміністратора безпеки можна тільки після виконання явних, що протоколюються дій;
- аутентифікаційна інформація повинна бути захищена від несанкціонованого доступу;
- повинен бути реалізований механізм управління процесами в системі;
- користувачі повинні пройти ідентифікацію і аутентифікацію перш ніж виконувати будь-які дії в системі;
- управління системними сервісами і драйверами можливо тільки адміністратору або групі довірених осіб;
- наявність пакетного фільтра з можливістю створення правил для вхідного і вихідного мережевого трафіку;
- можливість локально заблокувати доступ до системи на час відсутності користувача;
- процеси і об'єкти ядра не повинні мати можливість за власною ініціативою змінювати дані інших процесів або об'єктів ядра;
- повинно бути задано явне і недвозначне перерахування допустимих типів доступу (читання, запис, виконання, виконання з більш вищими привілеями) для кожного суб'єкта або групи по відношенню до (ресурсу);
- повинен проводиться і записуватися моніторинг ключових подій в системі;
- контроль доступу повинен бути застосовний до кожного об'єкту і кожного суб'єкту.
- можливість змінювати правила доступ має надаватися виділеним суб'єктам (адміністраторам, довіреним групам).

Середнім рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного рівня):

- можливість відновлення системи після збою, ушкодження системних файлів, можливість створення контрольної точки відновлення;
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання);
- вхід в систему під single-user повинен вимагати авторизації;
- кожне реєстроване дію має асоціюватися з конкретним користувачем;
- повинно здійснюватися очищення тимчасових системних файлів;
- можливість управляти настроюванням підключення до мережі і параметрами підключення тільки адміністратору або групі довірених осіб;
- в систему повинен бути вбудований механізм, що забезпечує передачу даних на

віддалений сервер резервного копіювання;
- механізм встановлення паролів не повинен допускати установки слабких паролів менше 8 символів і мати осмислене значення;
- вся інформація про успішні і неуспішні спроби авторизації повинні записуватися в файл;
- всі події в системі, які мають відношення до роботи системи безпосередньо, до безпеки і до призначених для користувача додатків повинні фіксуватися і записуватися;
- в системі повинні бути інтегрована можливість здійснювати криптографічно стійкий захист даних від несанкціонованого доступу;
- наявність пакетного фільтра з можливістю створення правил для вхідного і вихідного мережевого трафіку з можливістю запису пакетів і переглядом мережевих з'єднань, а також правилами, що допускають вільне безперешкодне переміщення внутрішніх пакетів на локальні інтерфейси і адресу 127.0.0.1.

Високим рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного і середнього рівня):
- повинно здійснюватися очищення (обнулення, знеособлення)
областей оперативної пам'яті ЕОМ і зовнішніх накопичувачів, що звільняються. Очищення здійснюється довільної записом в звільнену область пам'яті;
- архівування результатів моніторингу з метою економії займаного місця;
- система повинна використовувати інтегроване шифрування файлової системи, при якому неможливо зчитати дані при знятті накопичувача і підключення його до іншого комп'ютера
- наявність механізму, який би на рівні ядра забороняв виконання коду з адресного сегмента даних;
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання);
- можливість управління правами доступу до мережі для різних користувачів;
- при виявленні несанкціонованої модифікації системних файлів (при завданні відповідної опції адміністратором) проводиться автоматичне відновлення файлів з резервної копії;
- підтримання стабільної роботи системного моніторингу протягом 15 хвилин атаки з повним фіксуванням дій атакуючого і паралельної відправкою на сервер резервних копій;
- система не повинна давати інформацію про своє найменування, версії і версії працюючих сервісів неавторизованих користувачів.

Для класу використання 5.

Умовним рівнем захищеності буде називатися ОС, при наявності таких критеріїв:
- отримати права адміністратора безпеки можна тільки після виконання явних, що протоколюються дій;
- аутентифікаційна інформація повинна бути захищена від несанкціонованого доступу;
- повинен бути реалізований механізм управління процесами в системі;
- користувачі повинні пройти ідентифікацію і аутентифікацію перш ніж виконувати будь-які дії в системі;
- управління системними сервісами і драйверами можливо тільки адміністратору або групі довірених осіб;
- наявність пакетного фільтра з можливістю створення правил для вхідного і вихідного мережевого трафіку;
- можливість локально заблокувати доступ до системи на час відсутності користувача;
- процеси і об'єкти ядра не повинні мати можливість за власною ініціативою змінювати дані інших процесів або об'єктів ядра;
- повинно бути задано явне і недвозначне перелічення допустимих типів доступу (читання, запис, виконання, виконання з більш вищими привілеями) для кожного суб'єкта або групи по відношенню до об'єкта (ресурсу);
- повинен проводиться і записуватися моніторинг ключових подій в системі;
- контроль доступу повинен бути застосовний до кожного об'єкту і кожного суб'єкту.
- можливість змінювати правила доступу має надаватися виділеним суб'єктам (адміністраторам, довіреним групам).

Середнім рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного рівня):

- можливість відновлення системи після збою, ушкодження системних файлів, можливість створення контрольної точки відновлення;
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання)
- вхід в систему під single-user повинен вимагати авторизації;
- кожна реєстрована дія має асоціюватися з конкретним користувачем;
- механізм встановлення паролів не повинен допускати установки слабких паролів менше 8 символів і мати осмислене значення;
- повинно здійснюватися очищення тимчасових системних файлів;
- можливість управляти настроюванням підключення до мережі і параметрами підключення тільки адміністратору або групі довірених осіб;
- вся інформація про успішні і неуспішні спроби авторизації повинні записуватися в файл;
- всі події в системі, які мають відношення до роботи системи безпосередньо, до безпеки і до призначених для користувача додатків повинні фіксуватися і записуватися;
- в системі повинен бути вбудований механізм, що забезпечує передачу даних на віддалений сервер резервного копіювання;
- в системі повинна бути інтегрована можливість здійснювати криптографічно стійкий захист даних від несанкціонованого доступу;
- наявність пакетного фільтра з можливістю створення правил для вхідного і вихідного мережевого трафіку з можливістю запису пакетів і переглядом мережевих з'єднань, можливістю створення окремих правил для всіх мережевих інтерфейсів.

Високим рівнем захищеності буде називатися ОС, при наявності таких критеріїв (в сукупності з критеріями умовного і середнього рівня):

- повинно здійснюватися очищення (обнулення, знеособлення) областей оперативної пам'яті ЕОМ і зовнішніх накопичувачів, що звільняються. Очищення здійснюється довільним записом в звільнену область пам'яті;
- архівування результатів моніторингу з метою економії займаного місця;
- система повинна використовувати інтегроване шифрування файлової системи, при якому неможливо вважати дані при знятті накопичувача і підключення його до іншого комп'ютера;
- наявність механізму, який би на рівні ядра забороняв виконання коду з адресного сегмента даних;
- наявність системних таблиць з описом прав доступу для підключення пристрою для кожного користувача з встановленими дозволами використання (читання, запис, виконання);
- можливість управління правами доступу до мережі для різних користувачів;
- при виявленні несанкціонованої модифікації системних файлів (при завданні відповідної опції адміністратором) проводиться автоматичне відновлення файлів з резервної копії;
- підтримка стабільної роботи системного моніторингу протягом 15 хвилин атаки з повним фіксуванням дій атакуючого і паралельної відправкою на сервер резервних копій;
- система не повинна давати інформацію про своє найменування, версії і версії працюючих сервісів неавторизованих користувачів;
- наявність пакетного фільтра з можливістю створення правил для вхідного і вихідного мережевого трафіку з можливістю запису пакетів і переглядом мережевих з'єднань, правилами, що допускають вільне безперешкодне переміщення внутрішніх пакетів на на локальні інтерфейси і адресу 127.0.0.1.
- можливість створення правила «Що явно не дозволено - заборонено». Створення правил, для яких більш часто виконується відповідність, повинні бути написані вище за правила, для яких відповідність виконується рідше.
- останнє правило в розділі повинно блокувати і заносити в лог всі пакети для даного інтерфейсу і напрямку;

- при блокуванні небажаних пакетів на них не повинена відправлятися відповідь, тобто щоб атакуючий не знав чи досягли його пакети мети;
- можливість створення динамічних правил для відбиття атак типу «відмова в обслуговуванні».

4. ВИСНОВОК

Захист операційної системи починається з відповіді на питання:
- в рамках якої інформаційної системи працює ОС ?
- який зміст вкладається в термін «захищене середовище» ?
- які умови роботи та обмеження, які ми накладаємо на захищене середовище ?
- які задачі стоять перед захищеним середовищем?
- захист від яких загроз буде потрібен захищеному середовищу?
- чи виконано всі типові дії (industry standard security actions) по забезпеченню належного рівня захищеності середовища?
- якими механізмами можна забезпечити потрібну нам захищеність середовища та яка буде «ціна» їх застосування чи незастосування?

Відповіді на ці питання є різними у різних проектах та задачах, тому необхідно для кожного конкретного випадку розробляти документ, який містить відповіді на ці питання, які відображають спільну точку зору керівництва та технічного персоналу. Цей документ носить назву «політика безпеки». Рекомендується, щоб цей документ мав ієрархічну будову, тобто розглядав загальні поняття, поступово їх уточнюючи (переходячи від задач, які виконує інформаційна система чи її уповноважені користувачі – до загроз реалізації цих задач, і від цих загроз - до механізмів активного, реактивного, проактивного, пасивного захисту інформаційної системи).

В статті запропоновано класифікацію операційних систем за варіантами використання, критерії захищеності, вимоги до механізмів захищеності ОС в рамках класу використання Для всіх класів операційних систем розроблені також критерії, при наявності яких можна присвоїти операційній системі відповідний рівень захищеності: умовний, середній або високий**.**

7. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] О.С. Бичков. Про концепцію захищеної операційної системи реального часу // Безпека інформаційних систем і технологій. №1(1). – 2019. – с. 42-56.

[2] Multilevel Secure Operating Systems // Journal Of Information Science And Engineering 15, 91-106 (1999).

[3] Bell E., LaPaduIa L. Secure Computer Systems: A Mathematical Model // MITRE Technical Report 2547, Volume II, 1973.

[4] Л.И. Фроимсон, С.В. Кутепов, О.В. Тараканов, А.В. Шереметов. Основные принципы построения защищенной операционной системы для мобильных устройств. Спецтехника и связь. – 2013.- №1.-С.43-47.

[5] P. Colp, M. Nanavati, J. Zhu, W. Aiello, G. Coker, T. Deegan, P. Loscocco, and A. Warfield. Breaking Up is Hard to Do: Security and Functionality in a Commodity Hypervisor. In SOSP'11: 23rd Symposium on Operating Systems Principles, pages 189–202, Cascais, October 2011. 14, 26, 27.

[6] N. Ferguson and B. Schneier. A Cryptographic Evaluation of IPsec. Counterpane Internet Security, Inc, December 2003. 14, 135.

[7] T. Garfinkel and M. Rosenblum. A Virtual Machine Introspection Based Architecture for Intrusion Detection. In NDSS'03: 10th Network and Distributed System Security Symposium, pages 191–206, San Diego, February 2003. 19, 70.

[8] E. I. Brycova, "Basic principles of construction of secure operating systems", Vestn. Astrakhan State Technical Univ. Ser. Management, Computer Sciences and Informatics, 2013, no. 2, 52–57.

[9] Apple. iOS Security. http://images.apple.com/ipad/business/docs/iOS_Security_Feb14.pdf, February 2014. 147.

[10] P. A. Karger and D. R. Safford. I/O for Virtual Machine Monitors: Security and

Performance Issues. IEEE Security & Privacy, 6(5):16–23, September 2008. 19.

[11] H. Härtig, M. Hohmuth, N. Feske, C. Helmuth, A. Lackorzynski, F. Mehnert, and M. Peter. The Nizza Secure-System Architecture. In CollaborateCom'05: 1st Conference on Collaborative Computing: Networking, Applications and Worksharing, San Jose, December 2005. 15, 139.

***Шестак Яніна Володимирівна**, кандидат технічних наук, асистент кафедри кібербезпеки та захисту інформації Київського національного університету імені Тараса Шевченка. Основними напрямами наукових досліджень є процес оцінювання захищеності від зовнішніх впливів інформаційних систем.*

***Бичков Олексій Сергійович**, доктор технічних наук, доцент,*

***Бичков Олексій Сергійович**, доктор технічних наук, доцент, завідувач кафедри Програмних систем і технологій Київського національного університету імені Тараса Шевченка.*

Область наукових інтересів: Математичні моделі систем захисту інформації, теорія гібридних автоматів, теорія можливості.

DOI:
УДК 004.8, 004.62, 004.93

COMPARATIVE ANALYSIS OF SYSTEM LOGS AND STREAMING DATA ANOMALY DETECTION ALGORITHMS

Andriy Lishchytovych [1]
orcid.org/0000-0002-3395-8616

Volodymyr Pavlenko [2]
orcid.org/0000-0002-3958-0415

Alexander Shmatok [3]
orcid.org/0000-0002-3351-2745

Yuriy Finenko [4]
orcid.org/0000-0003-1887-2475

[1] Kyiv, Ukraine, The Open International University of human development "Ukraine", AL@sors.me
[2] Kyiv, Ukraine, The Open International University of human development "Ukraine", Pavlenko.v@i.ua
[3] Kyiv, Ukraine, The Open International University of human development "Ukraine", sh_al_st@ukr.net
[4] Kyiv, Ukraine, The Open International University of human development "Ukraine", talaveryuriy@gmail.com

Історія статті:
Надійшла до редакції 21.07.2019
Прийнято 29.07.2019

Ключові слова:
виявлення аномалій;
системні журнали;
дерево прийняття рішень;
кластеризація;
аналіз даних;
ієрархічна часова пам'ять

Анотація: У цьому документі подано опис та порівняльний аналіз декількох загальноприйнятих підходів до аналізу системних журналів та потокових даних, що масово генеруються IT-інфраструктурою компанії, та виявленню аномалій. Важливість виявлення аномалії продиктована зростаючими витратами у випадку простою системи через події, які могли б бути передбачені на основі записів журналу з попереджувальними даними. Системи виявлення аномалій побудовані за допомогою стандартного процесу збору даних, аналізу, вилучення інформації та виявлення відхилень. Виявлення аномальної поведінки системи відіграє важливу роль у масштабних системах управління інцидентами. Своєчасне виявлення дозволяє IT-адміністраторам швидко виявити проблеми та негайно їх вирішити. Такий підхід значно скорочує час простою системи. Більшість IT-систем генерують журнали з детальною інформацією про операції. Тому журнали стають ідеальним джерелом даних рішень виявлення аномалії. Обсяг журналів унеможливлює їх аналіз вручну та вимагає автоматизованих підходів.Більша частина документа стосується кроку виявлення аномалії та таких алгоритмів, як регресія, дерево рішень, SVM, кластеризація, аналіз основних компонентів, видобуток інваріантів та ієрархічна модель тимчасової пам'яті. Алгоритми пошуку аномалії, що базуються на моделях, та ієрархічні алгоритми тимчасової пам'яті використовувались для обробки наборів даних HDFS, BGL та NAB з ~16 млн. повідомленнями журналу та ~365 тис. точками потокових даних. Дані були вручну позначені мітками, щоб дозволити навчання моделей та розрахунок точності їх роботи. Відповідно до результатів, системи контрольованого виявлення аномалій досягають високої точності, але потребують значних зусиль для тренувань моделей, тоді як алгоритм на основі HTM моделі показує найвищу точність виявлення при відсутності тренування.

Abstract: This paper provides with the description, comparative analysis of multiple commonly used approaches of the analysis of system logs, and streaming data massively generated by company IT infrastructure with an unattended anomaly detection feature. An importance of the anomaly detection is dictated by the growing costs of system downtime due to the events that would have been predicted based on the log entries with the abnormal data reported. Anomaly detection systems are built using standard workflow of the data collection, parsing, information extraction and detection steps. Most of the document is related to the anomaly detection step and algorithms like regression, decision tree, SVM, clustering, principal components analysis, invariants mining and hierarchical temporal memory model. Model-based anomaly algorithms and hierarchical temporary memory algorithms were used to process HDFS, BGL and NAB datasets with ~16m log messages and 365k data points of the streaming data. The data was manually labeled to enable the training of the models and accuracy calculation. According to the results, supervised anomaly detection systems achieve high precision but require significant training effort, while HTM-based algorithm shows the highest detection precision with zero

training. Detection of the abnormal system behavior plays an important role in large-scale incident management systems. Timely detection allows IT administrators to quickly identify issues and resolve them immediately. This approach reduces the system downtime dramatically.Most of the IT systems generate logs with the detailed information of the operations. Therefore, the logs become an ideal data source of the anomaly detection solutions. The volume of the logs makes it impossible to analyze them manually and requires automated approaches.

1. INTRODUCTION

Modern enterprises utilize large scale networks, including hybrids where cloud and on-premise items are connected into the single mesh. These systems are in use as the core part of IT, supporting different services – e-commerce, social networks, search engines and knowledge bases. IT infrastructure is designed to work 24/7 serving huge number of users globally. Any issues with this system will break down company services and lead to significant revenue loss.

Detection of the abnormal system behavior plays an important role in large-scale incident management systems. Timely detection allows IT administrators to quickly identify issues and resolve them immediately. This approach reduces the system downtime dramatically.

Most of the IT systems generate logs with the detailed information of the operations. Therefore, the logs become an ideal data source of the anomaly detection solutions. The volume of the logs makes it impossible to analyze them manually and requires automated approaches.

This paper provides the review of multiple anomaly detection algorithms with the evaluation of their efficiency.

2. ANOMALY DETECTION WORKFLOW AND ALGORITHMS

Common log-based anomaly detection systems are built using the following workflow:

Collection step provides the ways to gather logs information and put it into the common storage. Outcome of the logs parsing is structured timestamped events stream. At this stage some additional information could be also extracted in case of stable logs structure. To use the machine learning models information has to be transformed into the numerical form so called feature vectors. Every event becomes the vector and all the vectors are event matrix. Having the matrix, machine learning algorithms could be used to identify patterns and detect any anomalies.

Let's define an anomaly as a point in time where the behavior of the system is unusual and significantly different from previous, normal behavior. An anomaly may signify a negative change in the system, like a fluctuation in the turbine rotation frequency of a jet engine, possibly indicating an imminent failure. An anomaly can also be positive, like an abnormally high number of web clicks on a new product page, implying stronger than normal demand. Either way, anomalies in data identify abnormal behavior with potentially useful information. Anomalies can be spatial, where an individual data instance can be considered anomalous with respect to the rest of data, independent of where it occurs in the data stream.

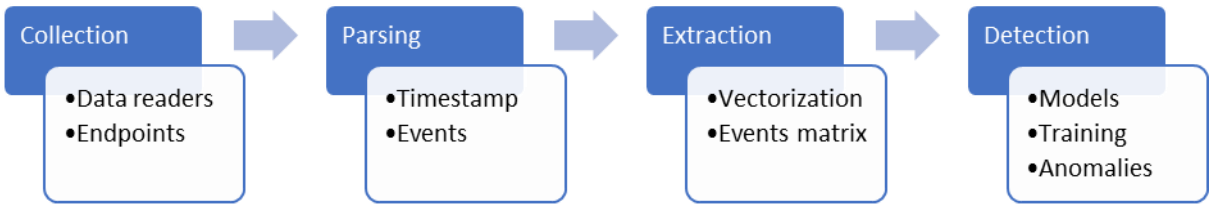


Fig. 1 - Anomaly detection system workflow

The description of the workflow steps and algorithms are described in [1].

3. LOG PARSING

Logs are plain text that consists of constant parts and variable parts, which may vary among different occurrences. For instance, given the logs of "Connection from 10.10.34.12 closed" and "Connection from 10.10.34.13 closed", the words "Connection", "from" and "closed" are considered as constant parts because they always stay the same, while the remaining parts are called variable parts as they are not fixed. Constant parts are predefined in source codes by developers, and variable parts are often generated dynamically (e.g., port number, IP address) that could not be well utilized in anomaly detection. The purpose of log parsing is to separate constant parts from variable parts and form a well-

established log event (i.e., “Connection from * closed” in the example). There are two types of log parsing methods: clustering based (e.g., LKE [5], LogSig [15]) and heuristic-based (e.g., iPLoM [11], SLCT [16]). In clustering-based log parsers, distances between logs are calculated first, and clustering techniques are often employed to group logs into different clusters in the next step. Finally, event template is generated from each cluster. For heuristic-based approaches, the occurrences of each word on each log position are counted. Next, frequent words are selected and composed as the event candidates. Finally, some candidates are chosen to be the log events.

4. FEATURE EXTRACTION

The main purpose of this step is to extract valuable features from log events that could be fed into anomaly detection models. The input of feature extraction is log events generated in the log parsing step, and the output is an event count matrix. In order to extract features, we firstly need to separate log data into various groups, where each group represents a log sequence. To do so, windowing is applied to divide a log dataset into finite chunks [3].

Fixed window. Both fixed windows and sliding windows are based on timestamp, which records the occurrence time of each log. Each fixed window has its size, which means the time span or time duration (the window size is Δt, which is a constant value, such as one hour or one day). Thus, the number of fixed windows depends on the predefined window size. Logs that happened in the same window are regarded as a log sequence.

Sliding window. Different from fixed windows, sliding windows consist of two attributes: window size and step size, e.g., hourly windows sliding every five minutes. In general, step size is smaller than window size, therefore causing the overlap of different windows. The number of sliding windows, which is often larger than fixed windows, mainly depends on both window size and step size. Logs that occurred in the same sliding window are also grouped as a log sequence, though logs may duplicate in multiple sliding windows due to the overlap.

Session window. Compared with the above two windowing types, session windows are based on identifiers instead of the timestamp. Identifiers are utilized to mark different execution paths in some log data. For instance, HDFS logs with block_id record the allocation, writing, replication, deletion of certain block. Thus, we can group logs according to the identifiers, where each session window has a unique identifier. After constructing the log sequences with windowing techniques, an event count matrix X is generated. In each log sequence, we count the occurrence number of each log event to form the event count vector. For example, if the event count vector is [0, 0, 2, 3, 0, 1, 0], it means that event 3 occurred twice and event 4 occurred three times in this log sequence. Finally, plenty of event count vectors are constructed to be an event count matrix X, where entry X(i,j) records how many times the event j occurred in the i-th log sequence.

5. SUPERVISED ANOMALY DETECTION

Supervised learning (e.g., decision tree) is defined as a machine learning task of deriving a model from labeled training data. Labeled training data, which indicate normal or anomalous state by labels, are the prerequisite of supervised anomaly detection. The more labeled the training data, the more precise the model would be. We will introduce three representative supervised methods: logistic regression, decision tree, and support vector machine (SVM) in the following.

6. LOGISTIC REGRESSION

Logistic regression is a statistical model that has been widely-used for classification. To decide the state of an instance, logistic regression estimates the probability p of all possible states (normal or anomalous). The probability p is calculated by a logistic function, which is built on labeled training data. When a new instance appears, the logistic function could compute the probability p, $p \in (0,1)$ of all possible states. After obtaining the probabilities, the states with the largest probability is the classification output. To detect anomalies, an event count vector is constructed from each log sequence, and every event count vector together with its label are called an instance. Firstly, we use training instances to establish the logistic regression model, which is actually a logistic function. After obtaining the model, we feed a testing instance X into the logistic function to compute its possibility p of anomaly, the label of X is anomalous when $p \geq 0.5$ and normal otherwise.

7. DECISION TREE

Decision Tree is a tree structure diagram that uses branches to illustrate the predicted state for each instance. The decision tree is constructed in a top-down manner using training data. Each tree node

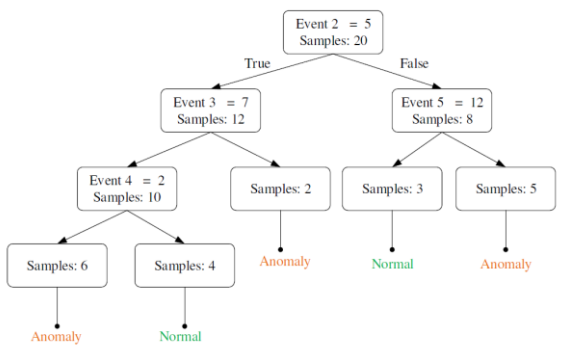


Fig. 2 -Decision Tree

is created using the current “best” attribute, which is selected by attribute’s information gain [6]. For example, the root node in Figure 2 shows that there are totally 20 instances in our dataset. When splitting the root node, the occurrence number of Event 2 is treated as the “best” attribute. Thus, the entire 20 training instances are split into two subsets according to the value of this attribute, in which one contains 12 instances and the other consists of 8 instances.

Decision Tree was first applied to failure diagnosis for web request log system in [4]. The event count vectors together with their labels described in Section III-B are utilized to build the decision tree. To detect the state of a new instance, it traverses the decision tree according to the predicates of each traversed tree node. In the end of traverse, the instance will arrive one of the leaves, which reflects the state of this instance.

8. SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a supervised learning method for classification. In SVM, a hyperplane is constructed to separate various classes of instances in high-dimension space. Finding the hyperplane is an optimization problem, which maximizes the distance between the hyperplane and the nearest data point in different classes.

In [8], Liang et al. employ SVM to detect failures and compared it with other methods. Similar to Logistic Regression and Decision Tree, the training instances are event count vectors together with their labels. In anomaly detection via SVM, if a new instance is located above the hyperplane, it would be reported as an anomaly, while marked as normal otherwise. In this paper, we only discuss linear SVM.

9. UNSUPERVISED ANOMALY DETECTION

Unlike supervised methods, unsupervised learning is another common machine learning task but its training data is unlabeled. Unsupervised methods are more applicable in real-world production environment due to the lack of labels. Common unsupervised approaches include various clustering methods, association rule mining, PCA and etc.

10. LOG CLUSTERING

In [9], Lin et al. design a clustering-based method called Log Cluster to identify online system problems. Log Cluster requires two training phases, namely knowledge base initialization phase and online learning phase. Thus, the training instances are divided into two parts for these two phases, respectively.

Knowledge base initialization phase contains three steps: log vectorization, log clustering, representative vectors extraction. Firstly, log sequences are vectorized as event count vectors, which are further revised by Inverse Document. Frequency (IDF) [14] and normalization. Secondly, Log Cluster clusters normal and abnormal event count vectors separately with agglomerative hierarchical clustering, which generates two sets of vector clusters (i.e., normal clusters and abnormal clusters) as knowledge base. Finally, we select a representative vector for each cluster by computing its centroid.

Online learning phase is used to further adjust the clusters constructed in knowledge base initialization phase. In online learning phase, event count vectors are added into the knowledge base one by one. Given an event count vector, the distances between

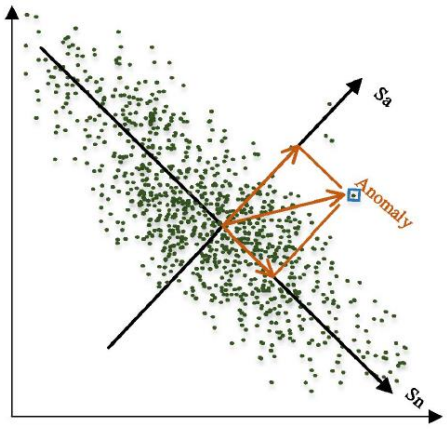


Fig. 3 -Log Cluster anomaly detection

it and existing representative vectors are computed. If the smallest distance is less than a threshold, this event count vector will be added to the nearest cluster and the representative vector of this cluster will be updated. Otherwise, Log Cluster creates a new cluster using this event count vector. After constructing the knowledge base and complete the online learning process, Log Cluster can be employed to detect anomalies. Specifically, to determine the state of a new log sequence, we compute its distance to representative vectors in knowledge base. If the smallest distance is larger than a threshold, the log sequence is reported as an anomaly. Otherwise, if the nearest cluster is a normal/an abnormal cluster, the log sequence is reported as normal/abnormal.

11. PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is a statistical method that has been widely used to conduct dimension reduction. The basic idea behind PCA is to project high-dimension data (e.g., high-

dimension points) to a new coordinate system composed of k principal components (i.e., k dimensions), where k is set to be less than the original dimension. PCA calculates the k principal components by finding components (i.e., axes) which catch the most variance among the high-dimension data. Thus, the PCA-transformed low-dimension data can preserve the major characteristics (e.g., the similarity between two points) of the original high-dimension data. For example, in Figure 3, PCA attempts to transform two-dimension points to the one-dimension points. S_n is selected as the principal component because the distance between points can be best described by mapping them to S_n

where n (∗) represents the number of logs which belong to corresponding event type ∗. Intuitively, Invariants mining could uncover the linear relationships between multiple log events that represent system normal execution behaviors. Linear relationships prevail in real-world system events. For example, normally, a file must be closed after it was opened. Thus, log with phrase “open file” and log with phrase “close file” would appear in pair. If the number of log events “open file” and that of “close file” in an instance are not equal, it will be marked abnormal because it violates the linear relationship. Invariants mining, which aims at finding invariants (i.e., linear relationships), contains three steps. The input of invariants mining is an

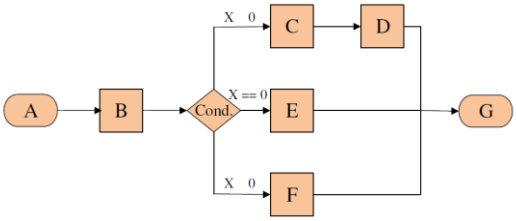


Fig. 4 - Principal Component Analysis

PCA was first applied in log-based anomaly detection [17]. In their anomaly detection method, each log sequence is vectorized as an event count vector. After that, PCA is employed to find patterns between the dimensions of event count vectors. Employing PCA, two subspaces are generated, namely normal space Sn and anomaly space Sa. Sn is constructed by the first *k* principal components and Sn is constructed by the remaining *(n-k)*, where n is the original dimension. Then, the projection $y_a = (1-PP^T)y$ of an event count vector y to S_a is calculated, where $p=[v_1, v_2, v_3, \ldots, v_n]$ is the first *k* principal components. If the length of y_a is larger than a threshold, the corresponding event count vector will be reported as an anomaly. For example, the selected point in Figure 3 is an anomaly because the length of its projection on S_a is too large

12. INVARIANTS MINING

Program Invariants are the linear relationships that always hold during system running even with various inputs and under different workloads. Invariants mining was first applied to log-based anomaly detection in [10]. Logs that have the same session id (e.g., *block_id* in HDFS) often represent the program execution flow of that session.

In this execution flow, the system generates a log message at each stage from A to G. Assuming that there are plenty of instances running in the system and they follow the program execution flow in Figure 4, the following equations would be valid:

event count matrix generated from log sequences, where each row is an event count vector. Firstly, the invariant space is estimated using singular value decomposition, which determines the amount r of invariants that need to be mined in the next step. Secondly, this method finds out the invariants by a brute force search algorithm. Finally, each mined invariant candidate is validated by comparing its support with a threshold (e.g., supported by 98% of the event count vectors). This step will continue until r independent invariants are obtained. In anomaly detection based on invariants, when a new log sequence arrives, we check whether it obey the invariants. The log sequence will be reported as an anomaly if at least one invariant is broken.

13. TOOL IMPLEMENTATION

[1] implemented six anomaly detection methods in Python with over 4,000 lines of code and packaged them as a toolkit. For supervised methods, [1] utilizes a widely-used machine learning package, scikit-learn [13], to implement the learning models of Logistic Regression, Decision Tree, and SVM.

There are plenty of parameters in SVM and logistic regression, and we manually tune these parameters to achieve the best results during training. For SVM, we tried different kernels and related parameters one by one, and we found that SVM with linear kernel obtains the better anomaly detection accuracy than other kernels. For logistic regression, different parameters are also explored,

and they are carefully tuned to achieve the best performance.

Implementing unsupervised methods, however, is not straightforward. For log clustering, [1] were not able to directly use the clustering API from scikit-learn, because it is not designed for large-scale datasets, where our data cannot fit to the memory. We implemented the clustering algorithm into an online version, whereby each data instance is grouped into a cluster one by one. There are multiple thresholds to be tuned. We also paid great efforts to implement the invariants mining method, because we built a search space for possible invariants and proposed multiple ways to prune all unnecessary invariants. It is very time-consuming to test different combination of thresholds. [1] finally implemented PCA method according to the original reference based on the use of an API from scikit-learn. PCA has only two parameters and it is easy to tune.

14. SUPERVISED METHODS EXPERIMENTAL DATASETS

Log Datasets. Publicly available production logs are scarce data because companies rarely publish them due to confidential issues. Fortunately, by exploring an abundance of literature and intensively contacting the corresponding authors, [1] has successfully obtained two log datasets, HDFS data [17] and BGL data [12], which are suitable for evaluating existing anomaly detection methods. Both datasets are collected from production systems, with a total of 15,923,592 log messages and 365,298 anomaly samples, that are manually labeled by the original domain experts. Thus, [1] took these labels (anomaly or not) as the ground truth for accuracy evaluation purposes. More statistical information of the datasets is provided in Table I. HDFS data contain 11,175,629 log messages, which were collected from Amazon EC2 platform [17]. HDFS logs record a unique block ID for each block operation such as allocation, writing, replication, deletion. Thus, the operations in logs can be more naturally captured by session windows, as introduced in III-B, because each unique block ID can be utilized to slice the logs into a set of log sequences. Then we extract feature vectors from these log sequences and generate 575,061 event count vectors. Among them, 16,838 samples are marked as anomalies.

BGL data contain 4,747,963 log messages, which were recorded by the Blue Gene supercomputer system at Lawrence Livermore National Labs (LLNL) [12]. Unlike HDFS data, BGL logs have no identifier recorded for each job execution. Thus, we have to use fixed windows or sliding windows to slice logs as log sequences, and then extract the corresponding event count vectors. But the number of windows depends on the chosen window size (and step size). In BGL data, 348,460 log messages are labeled as failures, and a log sequence is marked as an anomaly if any failure logs exist in that sequence

15. SUPERVISED METHODS ACCURACY RESULTS

To explore the accuracy of supervised methods, we use them to detect anomalies on HDFS data and BGL data. [1] uses session windows to slice HDFS data and then generate the event count matrix, while fixed windows and sliding windows are applied to BGL data separately. In order to check the validity of three supervised methods (namely Logistic Regression, Decision Tree, SVM), we first train the models on training data, and then apply them to testing data. [1] reports both training accuracy and testing accuracy in different settings, as illustrated in Tables 1 - 4. We can observe that all supervised methods achieve high training accuracy (over 0.95), which implies that normal instances and abnormal instances are well separated by using our feature representation. However, their accuracy on testing data varies with different methods and datasets. The overall accuracy on HDFS data is higher than the accuracy on BGL data with both fixed windows and sliding windows. This is mainly because HDFS system records relatively simple operations with only 29 event types, which is much less than that in BGL data, which is 385. Besides, HDFS data are grouped by session windows, thereby causing a higher correlation between events in each log sequence. Therefore, anomaly detection methods on HDFS perform better than on BGL.

16. USING HTM TO DETECT THE ANOMALIES

There are a number of other restrictions that can make methods unsuitable for real-time streaming anomaly detection, such as computational constraints that impede scalability. An example is Lytics Anomalyzer [25], which runs in O(n2), limiting its usefulness in practice where streams are arbitrarily long. Dimensionality is another factor that can make some methods restrictive. For instance, online variants of principle component analysis (PCA) such as osPCA [26] or window-based PCA [27] can only work with high-dimensional, multivariate data streams that can be projected onto a low dimensional space. Techniques that require data labels, such as supervised classification-based methods [28], are typically unsuitable for real-time anomaly detection and continuous learning. On the other hand, Hierarchical Temporal Memory models are best suited for the streaming data by the nature of its architecture [18].

Table 1. Supervised HDFS data set with session windows

	Supervised HDFS with session windows					
	Training			Testing		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Logistic	0.96	1	0.98	0.95	1	0.98
Decision Tree	1	1	1	1	0.99	1
SVM	0.96	1	0.98	0.95	1	0.98

Table 2. Supervised BGL data set with fixed windows

	Supervised BGL with fixed windows					
	Training			Testing		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Logistic	0.99	1	1	0.95	0.57	0.71
Decision Tree	1	1	1	0.95	0.57	0.72
SVM	1	1	1	0.95	0.57	0.71

Table 3. Supervised BGL data set with sliding windows

	Supervised BGL with sliding windows					
	Training			Testing		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Logistic	0.99	0.81	0.89	1	0.7	0.82
Decision Tree	1	1	1	0.92	0.63	0.74
SVM	1	1	1	0.99	0.75	0.85

Table 4. Unsupervised HDF & BGL data sets

	Unsupervised models on HDFS & BGL					
	HDFS			BGL		
	Precision	Recall	F-measure	Precision	Recall	F-measure
Log Clustering	0.87	0.74	0.8	0.42	0.87	0.57
Invariant Mining	0.88	0.95	0.91	0.83	0.99	0.91
PCA	0.98	0.67	0.79	0.5	0.61	0.55

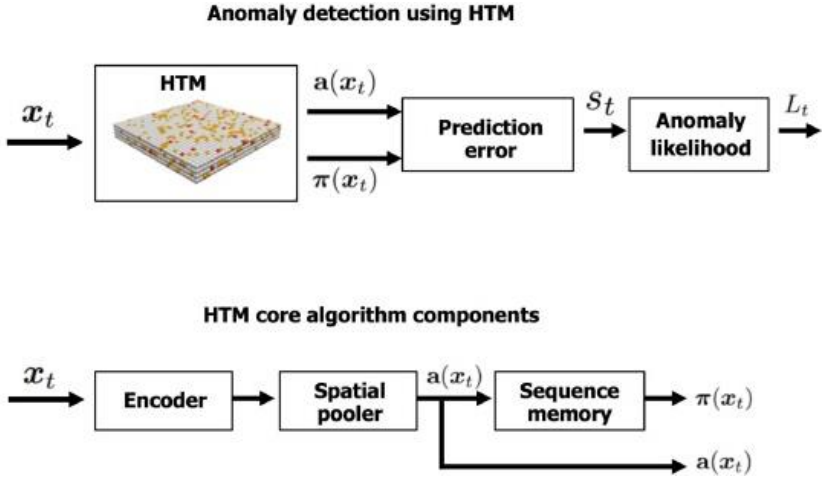


Fig. 5 - HTM-based anomaly detection workflow

Model-based approaches have been developed for specific use cases, but require explicit domain knowledge and are not generalizable. Domain-specific examples include anomaly detection in

aircraft engine measurements [19], cloud datacenter temperatures [20], and ATM fraud detection [21]. Kalman filtering is a common technique, but the parameter tuning often requires domain knowledge and choosing specific residual error models [22 - 24]. Model-based approaches are often computationally efficient but their lack of generalizability limits their applicability to general streaming applications.

17. HTM BENCHMARK DATASET

The aim of the dataset is to present algorithms with the challenges they will face in real-world scenarios, such as a mix of spatial and temporal anomalies, clean and noisy data, and data streams where the statistics evolve over time. The best way to do this is to provide data streams from real-world use cases, and from a variety of domains and applications. The data currently in the dataset represents a variety of sources, ranging from server network utilization to temperature sensors on industrial machines to social media chatter.

Dataset (NAB) version 1.0 contains 58 data streams, each with 1000–22,000 records, for a total of 365,551 data points [29]. Also included are some artificially-generated data files that test anomalous behaviors not yet represented in the corpus's real data, as well as several data files without any anomalies. All data files are labeled, either because we know the root cause for the anomalies from the provider of the data, or as a result of the well-defined NAB labeling procedure. These labels define the ground truth anomalies used in the NAB scoring process.

Traditional scoring methods, such as precision and recall, don't suffice because they don't effectively test anomaly detection algorithms for real-time use. For example, they do not incorporate time and do not reward early detection. Therefore, the standard classification metrics – true positive (TP), false positive (FP), true negative (TN), and false negative (FN) are not applicable for evaluating algorithms for the above requirements.

NAB includes three different application profiles: standard, reward low FPs, and reward low FNs. The standard profile assigns TPs, FPs, and FNs with relative weights (tied to the window size) such that random detections made 10% of the time would get a zero-final score on average. The latter two profiles accredit greater penalties for FPs and FNs, respectively. These two profiles are somewhat arbitrary but designed to be illustrative of algorithm behavior. The NAB codebase itself is designed such that the user can easily tweak the relative weights and re-score all algorithms. The application profiles thus help evaluate the sensitivity of detectors to specific applications criteria. The combination of anomaly windows, a smooth temporal scoring function (details in the next section), and the introduction of application profiles allows researchers to evaluate online anomaly detector implementations against the requirements of the ideal detector. Specifically, the overall NAB scoring system evaluates real-time performance, prefers earlier detection of anomalies, penalizes "spam" (i.e. FPs), and provides realistic costs for the standard classification evaluation metrics TP, FP, TN, and FN.

Table 5 Streaming data anomaly detection results

Detector	Standard Profile	Reward Low FP	Reward Low FN
Perfect	*100*	*100*	*100*
Numenta HTM	70.5-69.7	62.6-61.7	75.2-74.2
CAD OSE	69.9	67	73.2
earthgecko Skyline	58.2	46.2	63.9
KNN CAD	58	43.4	64.8
Relative Entropy	54.6	47.6	58.8
Random Cut Forest	51.7	38.4	59.7
Twitter ADVec v1.0.0	47.1	33.6	53.5
Windowed Gaussian	39.6	20.9	47.4
Etsy Skyline	35.7	27.1	44.5
Bayesian Changepoint	17.7	3.2	32.2
EXPoSE	16.4	3.2	26.9
Random	11	1.2	19.5
Null	*0*	*0*	*0*

18. UNSUPERVISED METHODS ACCURACY RESULTS

The scores are normalized such that the maximum possible is 100.0 (i.e. the perfect detector), and a baseline of 0.0 is determined by the "null" detector (which makes no detections) [29].

19. CONCLUSION

Decision tree is more interpretable than the other two methods, as developers can detect anomalies with meaningful explanations (i.e., predicates in tree nodes). Logistic regression cannot solve linearly non-separable problems, which can be solved by SVM using kernels. However, parameters of SVM are hard to tune (e.g., penalty parameter), so it often requires much manual effort to establish a model. Unsupervised methods are more practical and meaningful due to the lack of labels. Log clustering uses the idea of online learning. Therefore, it is suitable for processing large volume of log data. Invariants mining not only can detect anomalies with a high accuracy, but also can provide meaningful and intuitive interpretation for each detected anomaly. However, the invariants mining process is time consuming. PCA is not easy to understand and is sensitive to the data. Thus, its anomaly detection accuracy varies over different datasets.

The following findings have been made:

- Supervised anomaly detection methods achieve high precision, while the recall varies over different datasets and window settings.
- Anomaly detection with sliding windows can achieve higher accuracy than that of fixed windows.
- Unsupervised methods generally achieve inferior performance against supervised methods. But invariants mining manifests as a promising method with stable, high performance.
- The settings of window size and step size have different effects on supervised methods and unsupervised methods.
- Most anomaly detection methods scale linearly with log size, but the methods of Log Clustering and Invariants Mining need further optimizations for speedup.
- Anomalies detection in streaming, real-time applications generally better to handle using

20. REFERENCES

[1] Shilin He, Jieming Zhu, Pinjia He, and Michael R. Lyu, Experience Report: System Log Analysis for Anomaly Detection, 2016 IEEE 27th International Symposium on Software Reliability Engineering

[2] Subutai Ahmad, Alexander Lavin, Scott Purdy, Zuha Agha, Unsupervised real-time anomaly detection for streaming data, Neurocomputing, Volume 262, 1 November 2017, Pages 134-147

[3] T. Akidau, R. Bradshaw, C. Chambers, S. Chernyak, R. J. FernÃ¡ndez- Moctezuma, R. Lax, S. McVeety, D. Mills, F. Perry, E. Schmidt, and S. Whittle. The dataflow model: a practical approach to balancing correctness, latency, and cost in massive-scale, unbounded, out-oforder data processing. In PVLDB'15: Proc. of the VLDB Endowment, volume 8, pages 1792–1803. VLDB Endowment, 2015.

[4] M. Chen, A. X. Zheng, J. Lloyd, M. I. Jordan, and E. Brewer. Failure diagnosis using decision trees. In ICAC'04: Proc. of the 1st International Conference on Autonomic Computing, pages 36–43. IEEE, 2004.

[5] Q. Fu, J. Lou, Y. Wang, and J. Li. Execution anomaly detection in distributed systems through unstructured log analysis. In ICDM'09: Proc. of International Conference on Data Mining, 2009.

[6] J. Han, M. Kamber, and J. Pei. Data mining: concepts and techniques. Elsevier, 2011.

[7] P. He, J. Zhu, S. He, J. Li, and R. Lyu. An evaluation study on log parsing and its use in log mining. In DSN'16: Proc. of the 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2016.

[8] Y. Liang, Y. Zhang, H. Xiong, and R. Sahoo. Failure prediction in ibm bluegene/l event logs. In ICDM'07: Proc. of the 7th International Conference on Data Mining, 2007.

[9] Q. Lin, H. Zhang, J.G. Lou, Y. Zhang, and X. Chen. Log clustering based problem identification for online service systems. In ICSE'16: Proc. of the 38th International Conference on Software Engineering, 2016.

[10] J. Lou, Q. Fu, S. Yang, Y Xu, and J. Li. Mining invariants from console logs for system problem detection. In ATC'10: Proc. of the USENIX Annual Technical Conference, 2010.

[11] A. Makanju, A. Zincir-Heywood, and E. Milios. Clustering event logs using iterative partitioning. In KDD'09: Proc. of International Conference on Knowledge Discovery and Data Mining, 2009.

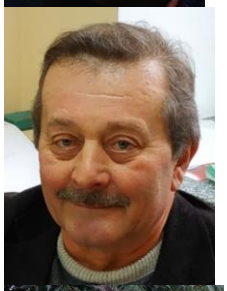
[12] A. Oliner and J. Stearley. What supercomputers say: A study of five system logs. In DSN'07:Proc. of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2007.

[13] F. Pedregosa, G. Varoquaux, A. Gramfort, et al. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

[14] G. Salton and C Buckley. Term weighting approaches in automatic text retrival. Technical report, Cornell, 1987.
[15] L. Tang, T. Li, and C. Perng. LogSig: generating system events from raw textual logs. In CIKM'11: Proc. of ACM International Conference on Information and Knowledge Management, pages 785–794, 2011.
[16] R. Vaarandi. A data clustering algorithm for mining patterns from event logs. In IPOM'03: Proc. of the 3rd Workshop on IP Operations and Management, 2003.
[17] W. Xu, L. Huang, A. Fox, D. Patterson, and M.I. Jordon. Detecting large-scale system problems by mining console logs. In SOSP'09: Proc. of the ACM Symposium on Operating Systems Principles, 2009.
[18] Yuwei Cui, Subutai Ahmad, Jeff Hawkins The HTM Spatial Pooler—A Neocortical Algorithm for Online Sparse Distributed Coding, Front. Comput. Neurosci., 29 November 2017, https://doi.org/10.3389/fncom.2017.00111
[19] D.L. Simon, A.W. Rinehart A model-based anomaly detection approach for analyzing streaming aircraft engine measurement data Proceedings of Turbo Expo 2014: Turbine Technical Conference and Exposition, ASME (2014), pp. 665-672, 10.1115/GT2014-27172
[20] Lee E.K., H. Viswanathan, D. Pompili Model-based thermal anomaly detection in cloud datacenters Proceedings of the IEEE International Conference on Distributed Computing in Sensor Systems (2013), pp. 191-198, 10.1109/DCOSS.2013.8
[21] T. Klerx, M. Anderka, H.K. Buning, S. Priesterjahn Model-based anomaly detection for discrete event systems Proceedings of the 2014 IEEE 26th International Conference on Tools with Artificial Intelligence, IEEE (2014), pp. 665-672, 10.1109/ICTAI.2014.105
[22] F. Knorn, D.J. Leith Adaptive Kalman filtering for anomaly detection in software appliances Proceedings of the IEEE INFOCOM (2008), 10.1109/INFOCOM.2008.4544581
[23] A. Soule, K. Salamatian, N. Taft Combining filtering and statistical methods for anomaly detection Proceedings of the 5th ACM SIGCOMM conference on Internet measurement, 4 (2005), p. 1, 10.1145/1330107.1330147
[24] Lee H., S.J. Roberts On-line novelty detection using the Kalman filter and extreme value theory Proceedings of the 19th International Conference on Pattern Recognition, (2008), pp. 1-4, 10.1109/ICPR.2008.4761918
[25] A. Morgan, Lytics Anomalyzer Blog, (2015). https://www.getlytics.com/blog/post/check_out_anomalyzer.
[26] Lee Y.J., Y.R. Yeh, Wang Y.C.F. Anomaly detection via online oversampling principal component analysis IEEE Trans. Knowl. Data Eng, 25 (2013), pp. 1460-1470, 10.1109/TKDE.2012.99
[27] A. Lakhina, M. Crovella, C. Diot Diagnosing network-wide traffic anomalies ACM SIGCOMM Comput. Commun. Rev, 34 (2004), p. 219, 10.1145/1030194.1015492
[28] N. Görnitz, M. Kloft, K. Rieck, U. Brefeld Toward supervised anomaly detection J. Artif. Intell. Res, 46 (2013), pp. 235-262, 10.1613/jair.3623
[29] The Numenta Anomaly Benchmark, https://github.com/numenta/NAB (accessed 2020-02-09)

Ліщитович Андрій Леонідович *– Medical Brain – темничній директор. Сфера наукових інтересів – інформаційні технології.*

Павленко Володимир Іванович *- професор кафедри Комп'ютерної інженерії. Сфера наукових інтересів – інформаційні технології.*



Шматок Олександр Станіславович *– директор інституту Комп'ютерних технологій університету «Україна». Сфера наукових інтересів – інформаційні технології.*

Фіненко Юрій Іванович *– інженер-дослідник лабораторії кіберфізичних систем університету «Україна». Сфера наукових інтересів – машинне навчання, захист інформації, програмування ПЛІС.*

DOI
УДК 004.896

ОПТИМІЗАЦІЯ СИСТЕМИ ПОЖЕЖНОГО МОНІТОРИНГУ З ВИКОРИСТАННЯМ МЕТОДУ ДЕФОРМОВАНИХ ЗІРОК

Антоневич Марина [1]
orcid.org/0000-0001-9797-6015
Дідик Анна [2]
orcid.org/0000-0003-2000-7835
Снитюк Віталій [3]
orcid.org/0000-0001-5154-5324

1) Україна, м. Київ, Київський національний університет імені Тараса Шевченка, marina.antonevich@gmail.com
2) Україна, м. Київ, Київський національний університет імені Тараса Шевченка, anna.didyk1999@gmail.com
3) Україна, м. Київ, Київський національний університет імені Тараса Шевченка, snytyuk@gmail.com

Історія статті:
Надійшла до редакції 17.02.2020
Прийнято 29.02.2020

Ключові слова:
система пожежної сигналізації;
метод деформованих зірок;
оптимізація.

Анотація: У статті розглядаються аспекти розв'язання задачі оптимізації функції двох змінних, яка, у загальному випадку є поліекстремальною та недиференційованою. Класичні методи неперервної оптимізації у цьому випадку є незастосовними. Одним із найбільш часто використовуваних методів розв'язання такої задачі є еволюційні алгоритми, які можна розділити на два класи. До першого класу належать алгоритми, де потенційний розв'язок-нащадок генерується двома батьківськими розв'язками, у другому випадку нащадок генерується одним батьківським розв'язком. Запропоновано метод деформованих зірок, де популяцію батьківських розв'язків становлять 3-х, 4-х та 5-ти точкові групи. Показано застосування пропонованого методу до розв'язання задачі оптимізації системи пожежного моніторингу будівель та споруд, що приводить до мінімізації часу її спрацювання. Розглянуті будівлі, де пожежне навантаження може мати як постійний, так і змінний характер. До таких будівель відносяться концертні зали, нічні клуби, супермаркети логістичні споруди тощо. Пожежі на таких об'єктах призводять до людських жертв та значних матеріальних збитків. Вчасне спрацювання системи пожежної сигналізації набуває великого значення. Цільова функція задачі визначена з урахуванням відстані від горизонтальних проекцій сповіщувачів до джерел виникнення пожежі та ймовірності спрацювання сповіщувачів. Розв'язком задачі є оптимізоване розміщення пожежних сповіщувачів із урахуванням їх кількості та пожежного навантаження приміщення. Показані переваги розробленого методу перед генетичними алгоритмами, еволюційними стратегіями та методом диференціальної еволюції як найбільш типовими еволюційними алгоритмами. Проведені чисельні експерименти, які засвідчили підвищену точність розрахунків та збільшену швидкість збіжності методу.

Abstract: In this paper are being considered the aspects of two variables function optimization problem solving, which, in general, is poly-extremal and undifferentiated. The classic methods of continuous optimization are not applicable in this case. One of the most commonly used methods of solving this problem is evolutionary algorithms, which can be divided into two classes. The first class includes algorithms where a potential offspring-solution is generated by two parent-solutions solutions, in the second case, the offspring-solution is generated by one parent-solution. There is deformed star method proposed where the population of parental solutions is 3, 4, and 5 point groups. The application of proposed method is shown to solve the optimization problem of fire monitoring system for buildings, which minimizes the time of its operation. The buildings where fire load can be both permanent and variable are considered. Such buildings include concert halls, nightclubs, supermarkets, logistics facilities and more. Fires at such buildings result in human sacrifice and serious material loss. Timely activation of the fire alarm system have great importance. The objective function of the problem is determined by the distance from the horizontal projections of the detectors to the sources of fire and the probability of triggering the detectors. The solution is optimizing location of fire detectors, taking into account their number and the fire load of the room. The advantages of the developed method over genetic algorithms, evolutionary strategies and differential evolution as the most typical evolutionary algorithms are shown. Numerical experiments were carried out, which showed the increased accuracy of calculations and the increased speed of method convergence.

1. ВСТУП

Більшість практичних задач пов'язана із необхідністю розв'язання задач оптимізації. У випадку, коли цільова функція є неперервною, гладкою, моноекстремальною застосовуються класичні методи неперервної оптимізації [1], в інших випадках, як правило, застосовуються методи стохастичної оптимізації [2]. Як перші, так і другі методи мають чисельні недоліки та особливості застосування. При мінімальних застереженнях до виду цільової функції та початкових даних раціонально використовувати еволюційні методи оптимізації [3-6].

Однією із таких задач є оптимізація системи пожежного моніторингу будівель та споруд і підзадача розміщення в приміщенні пожежних сповіщувачів [7]. Особливо актуальним оптимізація такого розміщення є у великих будівлях, в яких знаходиться велика кількість людей або значні обсяги матеріальних цінностей.

Зауважимо, що вказана задача є задачею оптимізації функції багатьох змінних. Одними із відомих методів розв'язання даної задачі можна назвати генетичні алгоритми [4], еволюційні стратегії [5], диференціальну еволюцію [6]. Вказані технології є класичними, найбільш часто застосовуваними еволюційними методами. Водночас, розв'язання задачі оптимізації розміщення пожежних сповіщувачів вимагає значної ресурсної бази. Ще один шлях до її ефективного розв'язання – розробка нових методів.

Таким методом є метод деформованих зірок (МДЗ) [8,9]. Він також належить до еволюційних алгоритмів. Далі буде показано переваги методу, де популяція потенційних розв'язків задачі складається із 3-, 4- та 5-точкових груп, у порівнянні з іншими методами. Також буде досліджена ефективність розв'язання задачі оптимізації розміщення пожежних сповіщувачів методом деформованих зірок із використанням різних варіантів даної техніки.

2. ЗАДАЧА ОПТИМІЗАЦІЇ ДВОМІРНОЇ ФУНКЦІЇ ТА МЕТОД ЇЇ РОЗВ'ЯЗАННЯ

Нехай, задана дана цільова функція:

$$f(x_1, x_2) \to \min,$$

$$x_1 \in [p_1, p_2],\ x_2 \in [q_1, q_2],\ p_1, p_2, q_1, q_2 \in R$$

Про її властивості нічого не відомо.

Розглянемо основні кроки МДЗ для цього випадку. Зауважимо, що оптимізація функції в одновимірному випадку наведена в [8].

На першому кроці задаємо номер поточної ітерації $t = 0$. Генеруємо початкову популяцію потенційних розв'язків

$$P_t = \{(x_1^1, x_2^1), (x_1^2, x_2^2), ..., (x_1^n, x_2^n)\},$$

рівномірно розміщених в прямокутнику $[p_1, p_2] \times [q_1, q_2]$, $|P_t| = n$. Далі, для всіх $(x_1^j, x_2^j) \in P_t$ знаходимо $f_i = f(x_1^j, x_2^j)$, які формують популяцію $F_t = \{f_j\}_{j=1}^n$. Генеруємо елементи для нової популяції P_z. Вибираємо значення $i, j, k = random\{1, 2, ..., n\}$. Точки $(x_1^i, x_2^i), (x_1^j, x_2^j)$ та (x_1^k, x_2^k) формують трикутник (не лежать на одній прямій).

Будемо робити стиснення і перенесення на відстань a в напрямку кращої точки (x_1^i, x_2^i) або (x_1^j, x_2^j), або (x_1^k, x_2^k), у залежності від значення f_i (далі будемо вважати (x_1^i, x_2^i) кращою вершиною). Для цього знаходимо центр мас трикутника (x_1^0, x_2^0).

Обчислюємо відстань R від центроїда до кращої точки. Задаємо параметр r – відстань зсуву точок: $r = coef \cdot R$, де $coef$ – деякий коефіцієнт на інтервалі $[0,1]$. Тоді відношення відрізків матиме вигляд: $a = r / R$. Здійснюємо зсув вершин у напрямку кращої вершини на відстань r (рис. 1).

Графічно, це можна зобразити наступним чином, де ABC - деякий трикутник, $m.B$ - найкраща. Новий трикутник, що буде утворено - $A_1B_1C_1$ (рис.1).

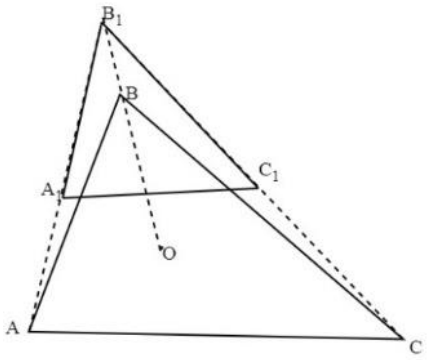


Рис. 1 - Перетворення трикутника

Якщо хоча б одна точка лежить в середині прямокутника можливих розв'язків, тоді зараховуємо цю точку до популяції P_z. Якщо хоча б одна точка лежить поза прямокутником, тоді нам необхідно застосувати додаткові перетворення і після цього зарахувати модифіковану точку до популяції P_z. Для всіх точок $(x_1^{i'}, x_2^{i'}), (x_1^{j'}, x_2^{j'}), (x_1^{k'}, x_2^{k'})$ знаходимо $f(x_1^{i'}, x_2^{i'})$, $f(x_1^{j'}, x_2^{j'})$ та $f(x_1^{k'}, x_2^{k'})$, які формують популяцію $F_Z = \{f_k^z\}_{k=1}^{3 \cdot k}$.

На наступному кроці генеруємо елементи нової популяції P_s. Вихідними є точки (x_1^i, x_2^i), (x_1^j, x_2^j) та (x_1^k, x_2^k), що формують трикутник.

Проведемо поворот трикутника на кут β навколо кращої вершини (рис. 2). У такому випадку вершина (x_1^i, x_2^i) залишається без змін, а інші будуть змінені. Точки $(x_1^{j'}, x_2^{j'})$ та $(x_1^{k'}, x_2^{k'})$ записуємо в популяцію P_s, якщо вони належать прямокутнику розв'язків. Інакше модифікуємо та записуємо в популяцію P_s. Для всіх точок $(x_1^{j'}, x_2^{j'})$ та $(x_1^{k'}, x_2^{k'})$ знаходимо $f(x_1^{j'}, x_2^{j'})$, $f(x_1^{k'}, x_2^{k'})$, які формують популяцію $F_s = \{f_k^s\}_{k=1}^{2\cdot k/3}$.

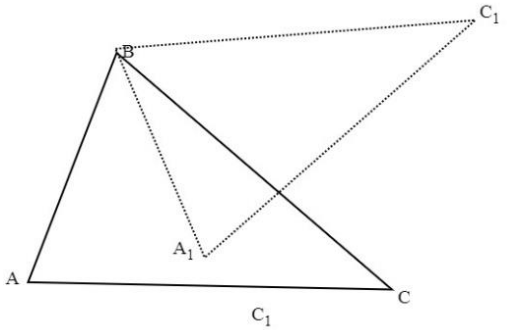


Рис. 2 - Поворот навколо найкращої точки

Далі генеруємо елементи для нової популяції P_w. Маємо точки $(x_1^i, x_2^i), (x_1^j, x_2^j)$ та (x_1^k, x_2^k), що формують трикутник. Здійснюємо поворот вершин трикутника на кут α навколо центру мас. Точки $(x_1^{i'}, x_2^{i'}), (x_1^{j'}, x_2^{j'})$ та $(x_1^{k'}, x_2^{k'})$ аналогічно записуємо в популяцію P_w Для всіх точок $(x_1^{i'}, x_2^{i'}), (x_1^{j'}, x_2^{j'}), (x_1^{k'}, x_2^{k'})$ знаходимо $f(x_1^{i'}, x_2^{i'})$, $f(x_1^{j'}, x_2^{j'})$ та $f(x_1^{k'}, x_2^{k'})$, які формують популяцію $F_w = \{f_k^w\}_{k=1}^{3\cdot k}$.

Об'єднуємо елементи кожної популяції $P_t \cup P_z \cup P_s \cup P_w = P$ та $F_t \cup F_z \cup F_s \cup F_w = F$. Елементи популяції F сортуємо за зростанням. Збільшуємо номер поточної ітерації $t = t + 1$. Створюємо нову популяцію P_t з елементів популяції P, що включатиме перші n елементів популяції F. Якщо критерій зупинки не виконано, то ітераційний процес продовжується. Якщо умову зупинки виконано, тоді значення потенційного розв'язку, в якому значення функції f є мінімальним, є шуканим розв'язком.

Зауважмо, що вище наведено основні елементи МДЗ, де зірки мають форму трикутника. Чотиривимірний та п'ятивимірний випадки є аналогічними.

3. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Для визначення ефективності МДЗ виконано порівняльний аналіз результатів його застосування із результатами класичних еволюційних методів: генетичного алгоритму (ГА), еволюційної стратегії (ЕС), диференціальної еволюції (ДЕ). Експерименти проводились при розв'язанні задачі оптимізації однієї із недиференційованих функцій з використанням різних критеріїв закінчення алгоритму: за кількістю ітерацій, за близькістю різниць середніх та максимальних значень фітнес-функцій для сусідніх популяцій.

Досліджено середній час виконання кожного із алгоритмів при 100 ітераціях (усереднено за 100 запусками програм), причому виконано підрахунок кількості результуючих значень, що досягли заданої точності та тих, що не досягли.

На рис. 3-4 представлено результати експериментів.

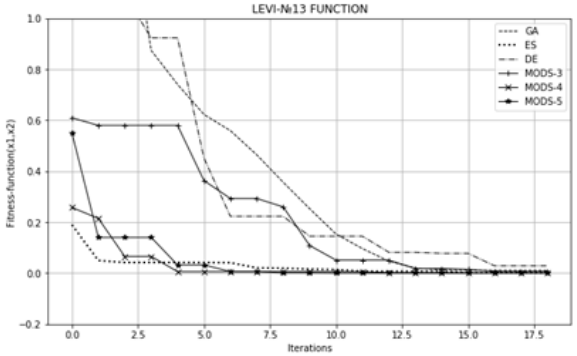


Рис. 3 – Умова зупинки: 100 ітерацій

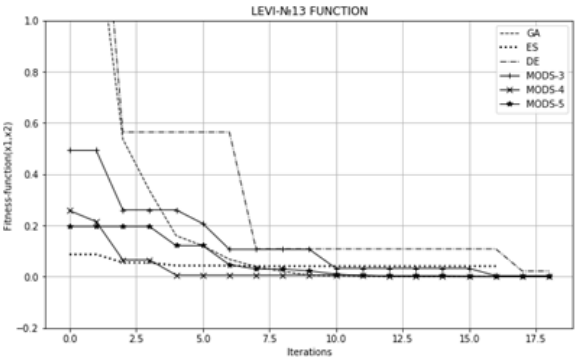


Рис. 4 – Умова зупинки: за середнім значенням фітнес-функцій

Результати, представлені на рис. 3-4, переконливо доводять переваги застосування МДЗ.

4.ПРАКТИЧНЕ ЗАСТОСУВАННЯ МЕТОДУ ДЕФОРМОВАНИХ ЗІРОК

У роботі [7] було розв'язано задачу оптимізації розміщення пожежних сповіщувачів у складському приміщенні методом деформованих зірок для мінімізація часу спрацювання системи пожежної сигналізації, наслідком чого є зменшення людських жертв та мінімізація матеріальних втрат. Розглянемо застосування МДЗ для розв'язання цієї задачі та його можливі варіації, а також різні штрафні функції та модифікації методу, де основними елементами є не точки на двовимірній площині, а вектори, що позначають множини пожежних сповіщувачів.

У цільовій функції задачі мають бути врахованими відстані між сповіщувачами та ймовірними джерелами виникнення пожежі, а також надійність системи пожежної сигналізації.

Припустимо, що всі сповіщувачі мають однакову ймовірність правильного спрацьовування p_d. Ефективність функціонування системи пожежного моніторингу, у першу чергу, визначається відстанню від i-го джерела небезпеки або точки виникнення пожежі до горизонтальної проекції сповіщувача або сповіщувачів.

Швидкість розвитку пожежі та час спрацювання сповіщувачів залежить від пожежного навантаження приміщення, тобто щільності розміщення та горючості матеріалів і об'єктів, які впливають на динаміку пожежі.

Розглянемо випадок постійного нерівномірного пожежного навантаження приміщення. До приміщень із постійним нерівномірним навантаженням належать склади із визначеними місцями для зберігання певних товарів, бібліотеки, житлові приміщення тощо.

Вважаємо відомою кількість пожежних сповіщувачів N, які мають бути встановленими у приміщенні, виходячи із відношення можливих збитків до витрат на створення системи пожежної сигналізації, де вартість сповіщувачів є основним фактором [7]. Така кількість може бути як більшою (як правило), так і меншою від визначеної нормативними документами. Нехай Ξ – горизонтальна проекція приміщення, де досліджується можливість оптимізації СПС, $\Xi = \{(x,y) \mid x \in [0,a], y \in [0,b]\}$.

Припустимо, що усі точки приміщення належать зоні відповідальності хоча б одного сповіщувача, яка є колом. Таким чином, горизонтальні проекції зон відповідальності перетинаються, утворюючи області Ξ_i з різною кратністю відповідальності сповіщувачів, $i = \overline{1, k_\Xi}$, де k_Ξ – кількість таких областей. Знаючи координати розміщення кожного сповіщувача (x_j, y_j), $j = \overline{1,N}$ і радіус зони його відповідальності r, можна одержати відповідну схему.

Припустимо, що $\Theta = \Theta_1 \cup \Theta_2 \cup ... \cup \Theta_n$, де Θ_i – ділянка приміщення, що характеризується рівнем пожежного навантаження η^i, $i = \overline{1,n}$, $\bigcap_{i=1}^{n} \Theta_i = \varnothing$. Очевидно, що ділянки Θ_i легко зобразити графічно, здійснити координатну прив'язку та вважати, що $\Theta_i = \{(x,y) \mid x \in [a_i^1, a_i^2], y \in [b_i^1, b_i^2]\}$. Границі ділянок $\Theta_i, i = \overline{1,n}$ є елементами матриці G.

У роботі [10] запропоновано розв'язання задачі визначення рівня пожежного навантаження. Для цього використано такі параметри, як Z_1 – тип матеріалу, який визначає пожежне навантаження, Z_2 – його об'єм на 1 см^2, Z_3 – нормальна швидкість горіння, Z_4 – швидкість вигорання, Z_5 – токсичність продуктів горіння, ξ – невраховані параметри. Очевидно, що тоді $\eta^i = h(Z_1^i, Z_2^i, Z_3^i, Z_4^i, Z_5^i, \xi^i)$, $i = \overline{1,n}$.

Зауважимо, що потужність та елементний склад кожного ξ^i встановлюється окремо у конкретному випадку. Для подальших досліджень потрібно було б мати ідентифіковану наведену вище залежність. Але, враховуючи унікальність кожного приміщення, її одержання в аналітичному вигляді є надто трудомістким процесом.

Вважаючи, що рівень пожежного навантаження η^i ділянки приміщення Θ_i, $i = \overline{1,n}$ відомий, переходимо до формування цільової функції, надаючи рівню пожежного навантаження η^i змісту вагового коефіцієнта, що визначає важливість надійності системи пожежного моніторингу та часу її спрацювання

Цільову функцію подамо таким чином:

$$F(W) = \sum_{l=1}^{n} \eta_l \cdot \sum_{i=1}^{M} \left(\sum_{j=1}^{N} \chi(d_{ij} < r) > 0 \,\&\, (T_i \in \Xi_l) \right) \times$$

$$\times \frac{1}{1-(1-p)^{\sum_{j=1}^{N} \chi((d_{ij} < r) \& (T_i \in \Xi_l))}} \cdot \min_j d_{ij} + P(W),$$

де W – структура системи сповіщувачів,

$X = (x_1, x_2, ..., x_N)$, $Y = (y_1, y_2, ..., y_N)$, (x_i, y_i) – координати сповіщувачів,

p_c^i – ймовірність спрацьовування сповіщувача або сповіщувачів з урахуванням резервування,

$\chi(*)$ – функція-індикатор,

d_{ij} – відстані від точок виникнення пожежі до пожежних сповіщувачів,

$T_i = (x_t^i, y_t^i)$ – точка виникнення пожежі,

$P(W)$ – штрафна функція.

Побудуємо два типи штрафних функцій: статичну та динамічну. Статична штрафна функція є такою:

$$S(W) = \sum_{l=1}^{n} \eta_l \cdot \sum_{i=1}^{M} \left(\sum_{j=1}^{N} \chi((d_{ij} < r) \& (T_i \in \Xi_l)) = 0 \right) \times$$

$$\times \min_j d_{ij}.$$

Динамічна штрафна функція має вигляд:

$$D(W) = \lambda(t) \cdot S(W)^2,$$

$$\lambda(t) = \begin{cases} C \cdot e^{\frac{t}{v}}, & при\ \lambda < \lambda_{\max} \\ \lambda_{\max}, & при\ \lambda \geq \lambda_{\max}. \end{cases},$$

де C – початкове значення множника штрафної функції,

v – швидкість зростання,

$\lambda_{\max}$ – максимальне значення множника,

t – номер ітерації.

Параметри $C, v, \lambda_{\max}$ визначаються в результаті експериментів.

Оскільки класичний МДЗ працює із розв'язком-точкою, а у даній задачі розв'язок є сукупністю точок (вектором), алгоритм було модифіковано. Потрібно знайти розв'язок-вектор, що має найкраще значення цільової функції.

Розглянемо приклад, що показаний на рис. 5. Дано три вектори - A, B і C, що складаються із семи точок $(x_i; y_i), i = \overline{1...7}$.

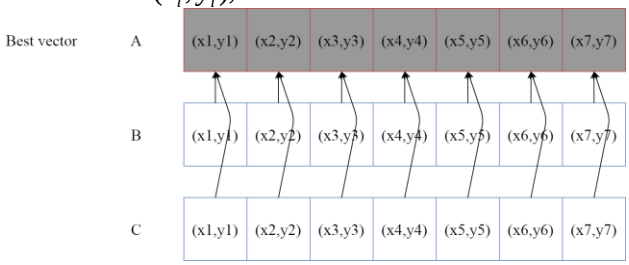


Рис. 5 – Приклад розв'язків

Припустимо, що найкращим є вектор A. Тоді нам необхідно покращити значення цільової функції векторів B і C. Для всіх точок векторів паралельно робимо однакові перетворення із однаковими вхідними параметрами. Точки $(x_1; y_1)_B$ та $(x_1; y_1)_C$ зсуваємо до точки $(x_1; y_1)_A$.

У залежності від того, який випадок використовується (тривимірний, чотиривимірний чи п'ятивимірний), використовують відповідні обчислення.

Такі дії виконуємо над усіма компонентами векторів. Тобто, далі виконуємо перетворення над $(x_2; y_2)_B$ та $(x_2; y_2)_C$, зсуваючи їх до $(x_2; y_2)_A$, і т.д. Одержимо нові вектори B^* і C^*, як показано на рис. 6.

Такі паралельні перетворення здійснюємо для стиснення, перенесення, повороту тощо.

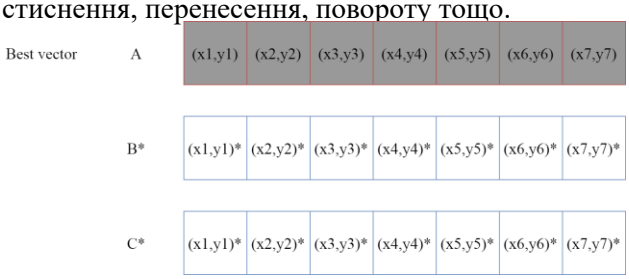


Рис. 6 – Результат перетворень

5. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Моделювання проведено за таких початкових умов:

– розмір приміщення 10x10;
– кількість пожежних сповіщувачів = 30;
– радіус реакції сповіщувача дорівнює 2;
– ймовірність спрацювання сповіщувача дорівнює 0.9;
– 3 зони пожежного навантаження;
– $C = 0.1, v = 1, \lambda_{\max} = 1$.

Критерії завершення алгоритму (далі КЗА1, КЗА2, КЗА3):

1. За кількістю ітерацій (200 прогонів);
2. За збіжністю різниць середніх значень фітнес-функцій для двох популяцій (точність 0.000001);
3. За збіжністю різниць максимальних значень фітнес-функцій для двох популяцій (точність 0.000001).

Введемо наступні позначення для відображення результатів на рисунках:

– прямокутні зони – зони різного пожежного навантаження;
– точки – координати виникнення пожежі;
– кола – це пожежні сповіщувачі, які мають певний радіус дії.

При використанні статичної штрафної функції даним методом для п'ятикутника отримуємо результати, зображені на рис. 7-8.

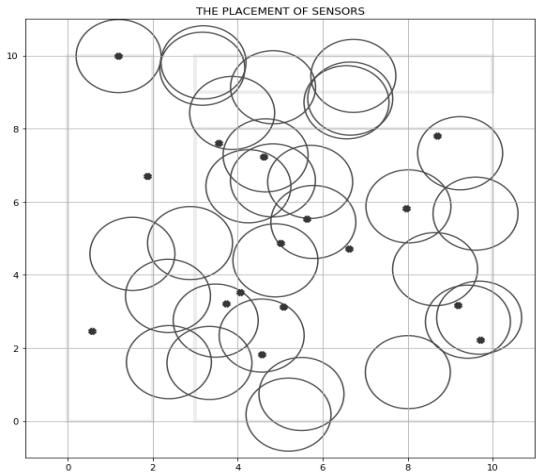


Рис. 7 – Розміщення сповіщувачів: 100 ітерацій

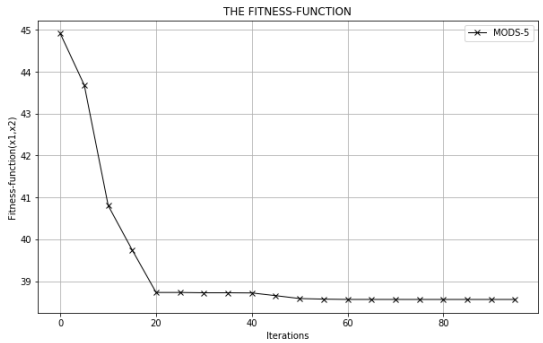


Рис. 8 – Динаміка зміни цільової функції: 100 ітерацій

Результати при використанні динамічної штрафної функції зображені на рис. 9-10.

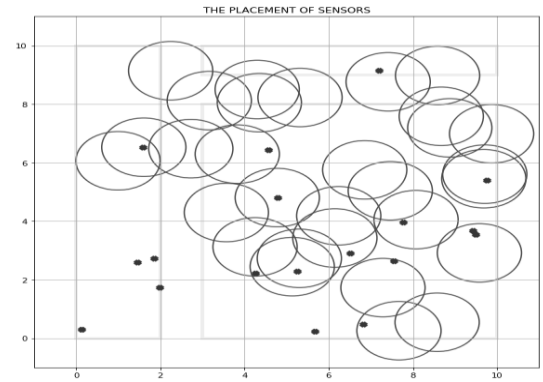


Рис. 9 – Розміщення сповіщувачів: 100 ітерацій

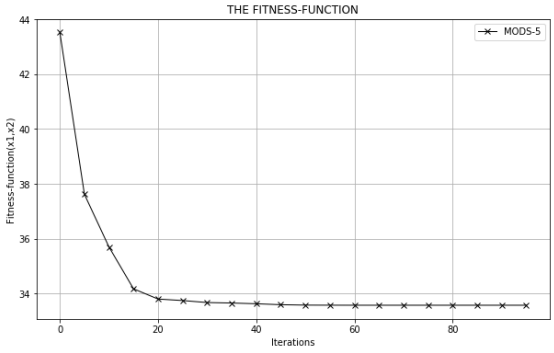


Рис. 10 – Динаміка зміни цільової функції: 100 ітерацій

Нижче в табл. 1 та табл. 2 наведені узагальнені результати для 10 запусків програми.

Таблиця 1. Узагальнений результат при використанні статичної штрафної функції

Метод	Критерій 1	Критерій 2	Критерій 3
МДЗ-3	27.979	30.54	38.604
МДЗ-4	19.406	21.193	45.856
МДЗ-5	38.457	38.703	35.023

Таблиця 2. Узагальнений результат при використанні динамічної штрафної функції

Метод	Критерій 1	Критерій 2	Критерій 3
МДЗ-3	24.346	28.11	21.053
МДЗ-4	15.443	14.686	18.103
МДЗ-5	33.176	29.097	26.16

6. ВИСНОВОК

Метод деформованих зірок демонструє переконливі результати своєї ефективності. Головна його ідея – більш глибоке дослідження області зміни аргументів, врахування взаємовпливу потенційних розв'язків, спрямованість пошуку та легкість реалізації. Метод є параметричним і допускає значну кількість налаштування, що може значно підвищувати як швидкість збіжності, так і точність розв'язання оптимізаційних задач.

При розв'язанні актуальної проблеми оптимізації розміщення пожежних сповіщувачів за допомогою методу деформованих зірок найкращі результати показав чотиривимірний випадок даного методу. Мінімальне значення цільової функції склало, в середньому, 22.4. Іншими еволюційними методами одержали значення цільової функції 28.4 та 33.4,що дозволяє стверджувати про ефективність методу деформованих зірок при розв'язанні практичних задач.

7. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] Du, D. Z.; Pardalos, P. M.; Wu, W. (2008). "History of Optimization". In Floudas, C.; Pardalos, P. (eds.). Encyclopedia of Optimization. Boston: Springer. pp. 1538–1542.

[2] *Spall, J. C. (2003).* Introduction to Stochastic Search and Optimization. *Wiley.* ISBN 978-0-471-33052-3.

[3] Vikhar, P. A. (2016). "Evolutionary algorithms: A critical review and its future prospects". Proceedings of the 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC). Jalgaon. pp. 261–265.

[4] J.A. Vasconcelos, J.A. Ramirez, R.H.C. Takahashi, R.R. "Saldanha Improvements in genetic algorithms", IEEE Transactions on Magnetics,Vol. 37, Issue 5, Sept. 2001.

[5] S. Van Rijn, H. Wang, M. van Leeuwen, T. Bäck, "Evolving the structure of Evolution Strategies", in 2016 IEEE Symposium Series on Computational Intelligence (SSCI), 6-9 Dec. 2016.

[6] A. Ukhov, "Differential Evolution: A tool for global optimization", Cornell Hospitality Report, 16(24), 3-44, 2016.

[7] Землянский А.Н., Каверина Н.П., Снитюк В.Е. Проектирование систем пожарного мониторинга в условиях неопределенности / Искусственный интеллект. – 2010. – № 3. – С. 483-488.

[8] Antonevych M. Optimization of functions of two variables by deformed stars method / M. Antonevych, A. Didyk, V. Snytyuk. // 2019 IEEE International Conference on Advanced Trends in Information Theory. – 2019. – С. 475–480.

[9] Snytyuk V., Method of Deformed Stars for Multi-extremal Optimization. One- and Two-Dimensional Cases // In Proc. Int. Conf. "Mathematical Modeling and Simulation of

Systems. MODS 2019", Advances in Intelligent Systems and Computing, vol 1019, Springer, Cham.
[10] Быченко А.А., Джулай А.Н., Снитюк В.Е. Эволюционные технологии принятия решений в пожаротушении. – Черкассы: Маклаут, 2008. –264 с.

***Снитюк Віталій Євгенович**, Вищу освіту здобув у Київському державному університеті імені Т.Г. Шевченка у 1991 році. У 2009 році захистив дисертацію на здобуття наукового ступеня доктор технічних наук за спеціальністю «Інформаційні технології». Посада: декан факультету інформаційних технологій Київського національного університету імені Тараса Шевченка*
Наукові інтереси: Прийняття рішень в умовах невизначеності, обчислювальний інтелект.

***Антоневич Марина Миколаївна,** студентка Київського національного університету імені Тараса Шевченка факультету інформаційних технологій за спеціальністю 122 «Комп'ютерні науки».*
Наукові інтереси: еволюційні обчислення, штучні нейронні мережі, інтелектуальний аналіз даних

***Дідик Анна Андріївна,** студентка Київського національного університету імені Тараса Шевченка факультету інформаційних технологій за спеціальністю 122 «Комп'ютерні науки».*
Наукові інтереси: еволюційні обчислення, веб-розробка, штучні нейронні мережі

DOI
УДК 004.51

УДОСКОНАЛЕНА МЕТОДИКА ЗАСТОСУВАННЯ ОБ'ЄКТНОЇ МОДЕЛІ ВЗАЄМОДІЇ DSP-SSP СИСТЕМ ЧЕРЕЗ AD EXCHANGE

Коротін Денис [1]
orcid.org/0000-0001-7239-2101

Поперешняк Світлана [2]
orcid.org/0000-0002-0531-9809

Коротін Сергій [3]
orcid.org/0000-0003-2123-6103

1) Київський національний університет ім. Тараса Шевченка, 01601, Київ, вул. Володимирська, 64/13, korotinmain@gmail.com
2) Київський національний університет ім. Тараса Шевченка, 01601, Київ, вул. Володимирська, 64/13, spopereshnyak@gmail.com
3) Національний університет оборони України ім. Івана Черняховського, 03049, м. Київ, Повітрофлотський проспект, 28, kororin2008@gmail.com

Історія статті:
Надійшла до редакції 28.11.2019
Прийнято

Ключові слова:
рекламні ресурси;
Інтернет контент;
об'єктна модель;
алгоритм роботи;
коефіцієнти виграшу.

Анотація: У статті проведено аналіз взаємодії DSP-SSP систем через Ad Exchange. За результатами проведеного аналізу сформовано мету наукового дослідження, яке полягає в тому, що на основі удосконаленої методики застосування об'єктної моделі взаємодії систем DSP-SSP провести моделювання та обґрунтувати доцільність використання реалізованих платформ для автоматизації купівлі та продажу медіа контенту на світовому ринку. Для досягнення поставленої мети було з'ясовано, що Ad Exchange є посередником між DSP та SSP платформами і реалізує модуль один до багатьох. Визначено, що така система дає можливість працювати на цифровому ринку, де видавці та рекламодавці збираються разом для торгівлі цифровими рекламними ресурсами. Зроблено висновок, що Ad Exchange – автономна платформа, яка полегшує і спрощує програмну покупку оголошень. Авторами представлено модель роботи системи Ad Exchange, вказані її переваги та недоліки. Зроблено висновок, що найважливішою складовою цієї моделі є вперше запропонований авторами блок Validator, який виставляє рекламу лише для встановленого певного контингенту людей. Реалізований в моделі Validator є основною перевагою представленої в статті системи над іншими. Запропоновано удосконалену методику використання об'єктної моделі взаємодії DSP-SSP систем через Ad Exchange. Визначено, що важливою складовою для реалізації даної методики є Аналітика, яка була реалізована на DSP платформі, де клієнт може слідкувати за важливими коефіцієнтами виграшу, кліків тощо. За рахунок включення блоку Validator удосконалено відомий алгоритм роботи DSP, SSP. На основі запропонованої моделі здійснено моделювання, де реалізовано блок Validator. Зроблено висновок, що Ad Exchange є інструментом для автоматизації процесу купівлі та продажу Інтернет контенту з можливістю отримувати більший заробіток та економити час на здійснення операцій. Розроблені авторами рекомендації відстежують неякісні пропозиції і борються з ризиками.

Abstract: In the article analyzed the interaction of DSP-SSP systems through Ad Exchange. According to the results of the analysis, the purpose of scientific research is formed, which is that for automate process of the buying and selling of media content, to carry out simulations and to prove the feasibility of using the implemented platforms on the basis of the advanced method of application of the object model of interaction of DSP-SSP systems. The authors found that Ad Exchange is an intermediary between DSP and SSP platforms and implements the module one to many. Has been identified that this system enables the digital marketplace where publishers and advertisers come together to trade digital inventory. It is concluded that Ad Exchange is a standalone platform that facilitates and simplifies programmatic ads buying. The authors presented the model of the Ad Exchange system, presents advantages and disadvantages are indicated. It is concluded that the Validator block, which advertises only for a specific contingent of people, was first proposed by the authors as the most important component of this model. The Validator block is the main advantage of the system presented in the article over others. An advanced technique for

the Validator block, the well-known DSP and SSP algorithm is improved. Based on the proposed model, a simulation was performed where the Validator block was implemented. It is concluded that Ad Exchange is a tool for automating the process of buying and selling Internet content. This allows you to earn more money and save time on transactions. The recommendations developed by the authors track poor quality proposals and deal with risks.

1. ВСТУП

На сьогоднішній день поряд з DSP та SSP платформами на сучасному ринку реалізовано Ad Exchange систему. Ad Exchange є посередником між DSP та SSP платформами і реалізує модуль один до багатьох.

Ad Exchange – це цифрові ринки, на яких видавці та рекламодавці збираються разом для торгівлі цифровими рекламними ресурсами, такими як дисплей, рідні, відео, мобільні пристрої та програми. Купівля та продаж відбуваються в аукціонах реального часу, уповноваженими технологією RTB (торги в реальному часі). Ad Exchange керує аукціонним механізмом для посередництва, який не обслуговує ні покупця, ні продавця; це автономна платформа, яка полегшує і спрощує програмну покупку оголошень [1].

Отже, між DSP та SSP існує зв'язок "один до одного", але існує система, яка автоматизує та з'єднує ці два процеси і робить зв'язок "один до багатьох", така система називається Ad Exchange [1].

Власники веб-сайтів, онлайнових журналів або блогів, відомих як "видавці", використовують SSP (платформа на стороні постачання) Ad Exchange і роблять свій цифровий рекламний простір доступним для покупців.

Незалежні маркетологи, рекламні агентства та рекламні мережі називаються "рекламодавцями", які використовують DSP (платформу на стороні попиту), щоб підключитися до Ad Exchange і придбати рекламний простір. Рекламні мережі також купують рекламний простір з обмінів об'яв, позначають його та продають для отримання прибутку. ATDs (агентство Trading Desk) використовують Ad Exchange, щоб купувати великі рекламні запаси і продавати оптом окремим рекламодавцям [2].

Процес програмування купівлі оголошень в провідних країнах світу здійснюється за допомогою програмного забезпечення та алгоритмів, що дозволяє видавцям отримувати найкращу ціну для свого рекламного простору та рекламодавців, щоб вони могли звернутися до цільової аудиторії в потрібний час і в контексті.

В той же час, існує нагальна потреба на українському ринку у розвитку сучасних інтернет-платформ.

Мета дослідження – на основі удосконаленої методики застосування об'єктної моделі взаємодії систем DSP-SSP провести моделювання та обґрунтувати доцільність використання реалізованих платформ для автоматизації купівлі та продажу медіа контенту на світовому ринку.

Для досягнення поставленої мети у процесі дослідження вирішені наступні завдання:

проведено аналіз взаємодії DSP-SSP систем через Ad Exchange;

представлено модель роботи системи Ad Exchange, її переваги та недоліки;

запропоновано удосконалену методику використання об'єктної моделі взаємодії DSP-SSP систем через Ad Exchange;

на основі запропонованої моделі здійснено моделювання;

обґрунтовано доцільність використання реалізованих платформ для автоматизації купівлі та продажу медіа контенту на українському ринку.

Результати дослідження впроваджено в DSP, SSP та Exchange систем компанії EPOM.

2. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Аналіз ринку щодо автоматизації процесу купівлі та продажу реклами дає можливість встановити, що процес пошуку і купівлі рекламних місць, а також купівля інвентарю на будь-яких біржах була витратним і неефективним процесом: не було стандартизованого інтерфейсу налаштування кампанії, для кожної площадки її доводилося налаштовувати "з нуля"; реклама демонструвалася одному і тому ж цільовому користувачу кілька разів через різні канали, а плата стягувалася багаторазово; такий стан справ ускладнювало ведення загальної аналітики [3, 4, 5, 6].

У [7] визначено, що ще у 2018 році обсяг

дисплейної і відеореклами, проданої за допомогою програматики, склав $ 14,2 млрд (+ 49% у порівнянні з 2017). Це більше 30% загальних витрат на банерну і відеорекламу [8]. Тільки на провідному ринку, в США, у найближчий період очікувані витрати на рекламу за допомогою програматики-технологій складуть $ 7,7 млрд, 79% з яких – витрати в RTB. До кінця 2019 року глобальні витрати на програматики в світі складуть близько $ 37 млрд [8]. Рівень проникнення програматики в країнах Західної Європи становить, в середньому, 10-30%.

Незважаючи на певні труднощі в оцінці українського ринку реклами в RTB, згодом RTB на українському ринку займе своє місце за аналогією з іншими країнами. Є підстави припустити, що вже за підсумками поточного року, частка RTB в Україні у найближчому майбутньому може скласти більше 10% ринку медійної реклами [9].

Враховуючи вищевикладене, а також з метою опанування українського медіа ринку, авторами статті було прийнято рішення розпочати реалізовувати даний проект за рахунок удосконаленої методики застосування об’єктної моделі взаємодії систем DSP-SSP.

3. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Ad Exchange надає різноманітні механізми контролю видавцю, такі як:

1) встановлення мінімальних CPM для одиниць запасів; фільтрування та блокування оголошень із конфіденційним вмістом або оголошеннями конкурентів;

2) визначення типів оголошень на веб-сторінці, з обранням формату об'яв і стилю оголошення;

3) створення індивідуальних комбінацій шрифтів, кольорів і кутового стилю та їх застосування для кількох графічних оголошень одночасно;

4) вибір місця розташування оголошень на веб-сторінці.

Рекламодавці отримують переваги від Ad Exchange з можливістю:

1) встановити параметри бюджетного ходу; обрати варіанти орієнтування, можливості встановлення ставок, поведінкове профілювання, налаштування цін;

2) визначення чорного списку певних веб-сайтів і аудиторій;

3) визначення обмежень на кількість разів, коли однакові оголошення відображаються одному користувачеві;

4) перенацілення на кількох обмінах об’явами.

3.1. ПОРЯДОК РОБОТИ AD EXCHANGE.

Видавець надає свої рекламні ресурси на Ad Exchange за допомогою SSP. Він надає повну інформацію про рекламні ресурси, такі як розташування сторінки, URL-адреса, аудиторія, теми тощо.

Коли користувач вводить веб-сайт або мобільний додаток видавця, показ оголошень автоматично відображається на аукціоні. Дані про користувача збираються, надсилаються на сервер видавця, а потім передаються на Ad Exchange. Потім Ad Exchange надсилає запит на надходження ставок до DSP та рекламних мереж. Кожен DSP перевіряє запит ставки та всю інформацію, пов'язану з нею (демографічні дані, ідентифікатор користувача, геопозиціювання, обмеження частоти показів, розділ на день та інші параметри націлювання тощо). Після перевірки платформа вирішить, чи буде враження цікавим для рекламодавця. Якщо це так, DSP надсилатиме відповіді на Ad Exchange з максимальною сумою ставки та місцем розташування рекламної копії, яка буде розміщена на доступному рекламному просторі [10]. Ad Exchange перевіряє рекламодавців, які надають ставки на показ та виключає рекламодавців, які не відповідають вимогам видавця.

Наприклад, рекламне місце доступне на китайському веб-сайті, а видавець встановлює обмеження для показу лише на китайській мові. Таким чином, Ad Exchange скасував би всі не-китайські оголошення з аукціону [11]. Зібравши всі дані, Ad Exchange аналізує ставки та продає лот найвищою ціною.

Виграшна рекламна копія з'являється на сайті видавця перед користувачем. Користувач Інтернету не знає про процес RTB, оскільки аукціони тривають не більше кількох мілісекунд, приблизно на час завантаження веб-сторінки. Процес не заважає користувачеві і не зменшує швидкість завантаження сторінки.

3.2. ТИПИ ОБМІНУ ОГОЛОШЕННЯМИ.

Open Ad Exchange / Public Marketplace / Відкритий аукціон – це відкритий цифровий ринок з великим інвентарем від декількох видавців, доступних для всіх покупців [12].

Open Ad Exchange пропонує широкий перелік видавців. Проте покупці не мають детальної інформації про видавця, як у випадку з приватним ринком. Покупці, які шукають більш широку публічність, обирають відкриту Ad Exchange.

Private Ad Exchange / Private Marketplace

(PMP) – це закрита платформа, яка дозволяє видавцю керувати цією програмою.

При обміні об'явами тисячі видавців можуть зробити рекламні місця, (які нереалізовані), доступними для рекламодавців, щоб вони могли пропонувати ставку автоматично – на відміну від традиційних переговорів про продаж один-на-один. Купівля та продаж відбувається програмно, що в основному є вигадливим терміном для автоматичної (перевірки нашого програмного забезпечення), між програмними платформами [13]. Рекламодавці та агентства, як правило, підключаються до обміну оголошеннями, використовуючи так звану платформу на стороні DSP – MediaMath, Turn, DataXu та AppNexus є деякими з великих DSP. Рекламні мережі також часто купують покази обміну оголошеннями, а потім помічають їх і перепродають покази покупцям у мережах. Зазвичай видавці надають доступ до показів обміну через платформу постачання (SSP).

На біржах об'яв рекламодавці можуть орієнтуватися на аудиторії, тобто групи користувачів, які демонструють певну поведінку веб-переглядача або відповідають певним інтересам своїх клієнтів – у великій кількості видавців. У момент, коли користувач переходить на сторінку на веб-сайті або в додаток, показ об'яв на сторінці може з'явитися на аукціоні на біржі. Платформа торгів покупця може аналізувати дані з декількох джерел – cookies користувача або мобільний ідентифікатор, дані від видавця, демографічне та купівельне поведінка від сторонніх постачальників даних або власні дані покупця – щоб з'ясувати, чи пропонувати ставки на показ і як багато [14].

3.3. ВІДКРИТІ БІРЖІ ТА ПРИВАТНІ РИНКИ.

Сьогодні досить часто використовується термін “відкритий обмін”. Це просто означає, що практично будь-яка система покупки може отримати доступ до пропозицій, наявних на біржі. Вона відкрита для всіх покупців. Відкриті біржі відрізняються від рекламних мереж таким чином. А процес обміну оголошеннями, як правило, є більш прозорим, ніж рекламні мережі, оскільки покупці можуть бачити ринкову ціну показу [15]. В той час, як відкриті біржі забезпечують більшу прозорість для середовища торгів, ніж запропоновані рекламні мережі, а також контроль над тим, хто є покупцями та продавцями. Приватні ринки є свого роду гібридом прозорості обох пропозицій,

Рекламодавці та агенції, які “нервують” через шахрайство у відкритих біржах (з поважною причиною), часто хочуть бути більш прихильними до того, де з'являться їхні оголошення. Крім того, видавці часто хочуть мати більший контроль над тим, які покупці отримують доступ до своїх запасів. Таким чином, більшість відкритих бірж тепер пропонує те, що називають приватними ринками або приватними біржами [16]. Приватні ринки все ще є автоматизованими умовами призначення ставок, але видавці можуть запрошувати бажаних рекламодавців або агентство до приватного пулу показів. Видавці також можуть встановлювати ціни для інвентаризації, що продаються на приватних ринках.

Як правило, інвентаризація, яка доступна на приватних ринках, також вважається більш якісною, ніж те, що видавці надаватимуть на відкритих біржах [17].

Видавці використовують платформи на стороні постачальника (SSP), зокрема, для забезпечення доступності запасів для програмної купівлі, і кілька видавців можуть підключатися до SSP.

Лінії між SSP розмиваються, оскільки технологія просунулася, і SSP залучили критичну масу видавців. На додаток, SSP призначені для підключення до кількох обмінів об'явами, рекламних мереж або будь-якої комбінації відкритих бірж і приватних ринків, а також пропонують інструменти, розроблені спеціально для видавців для оптимізації своїх продажів і покращення ставок (головне завдання – продати більше оголошень) на найвищих CPM [18].

Сьогодні все більше компаній керуються тим, що є, по суті, обмінними гібридами SSP. І надалі, зростає кількість компаній, які пропонують “повний стек” рекламних технологій з продуктами, включаючи продавців, які здійснюють аукціони. Рубікон Проект, DoubleClick і ONE від AOL є одним з прикладів компаній, які пропонують продукти для обох сторін [19].

3.4. ПЕРСПЕКТИВИ РОЗВИТКУ ІНТЕРНЕТ-ПЛАТФОРМ НА УКРАЇНСЬКОМУ РИНКУ.

Google DoubleClick Ad Exchange (AdX) вважаються найвідомішими [20]. Verizon, що належить AOL, керує гібридами SSP для відображення відео, які називаються ONE AOL Display та ONE AOL Video. Rubicon Project, AppNexus, PubMatic і OpenX є одними з найбільших незалежних бірж. Також однією з найширших та найперспективніших систем українського ринку є DSP, SSP та Ad Exchange

платформи компанії Epom, на прикладі яких розглянемо системний аналіз закупівлі та продажу реклами у вигляді об'єктної моделі.

Як вже було визначено вище – Ad Exchange є посередником між SSP та DSP платформами.

Звідси витікає висновок, що Ad Exchange є інструментом для автоматизації процесу купівлі та продажу Інтернет контенту в автоматичному режимі з можливістю отримувати більший заробіток і економити час на здійснення операцій. Розглянемо таку систему на прикладі продукту Ad Exchange компанії Epom. На рис. 1 представлена модель роботи цього рішення. Найбільш важливі функції цієї моделі вважаються такими:

– на Ad Opportunities розповсюджуються запити від SSP між усіма підключеними DSP;

– перехресно між двома групами вище;

– на Ad Opportunities розповсюджуються запити від усіх клієнтів Ad Server між усіма клієнтами Ad Server з підключеною DSP;

– Optimizer регулює розподіл трафіку між DSP, на основі статистичних даних для досягнення максимального win rate при колишніх QPS;

– Bid Probe намагається знайти нижній поріг і вищий поріг bid при прийнятному win rate;

– Filters робить фільтрацію запитів для оптимального розподілу між DSP;

– Analytics збирає всю інформацію щодо подій і повні запити для оптимізації торгів, шляхом визначення найкращих взаємозв'язків.

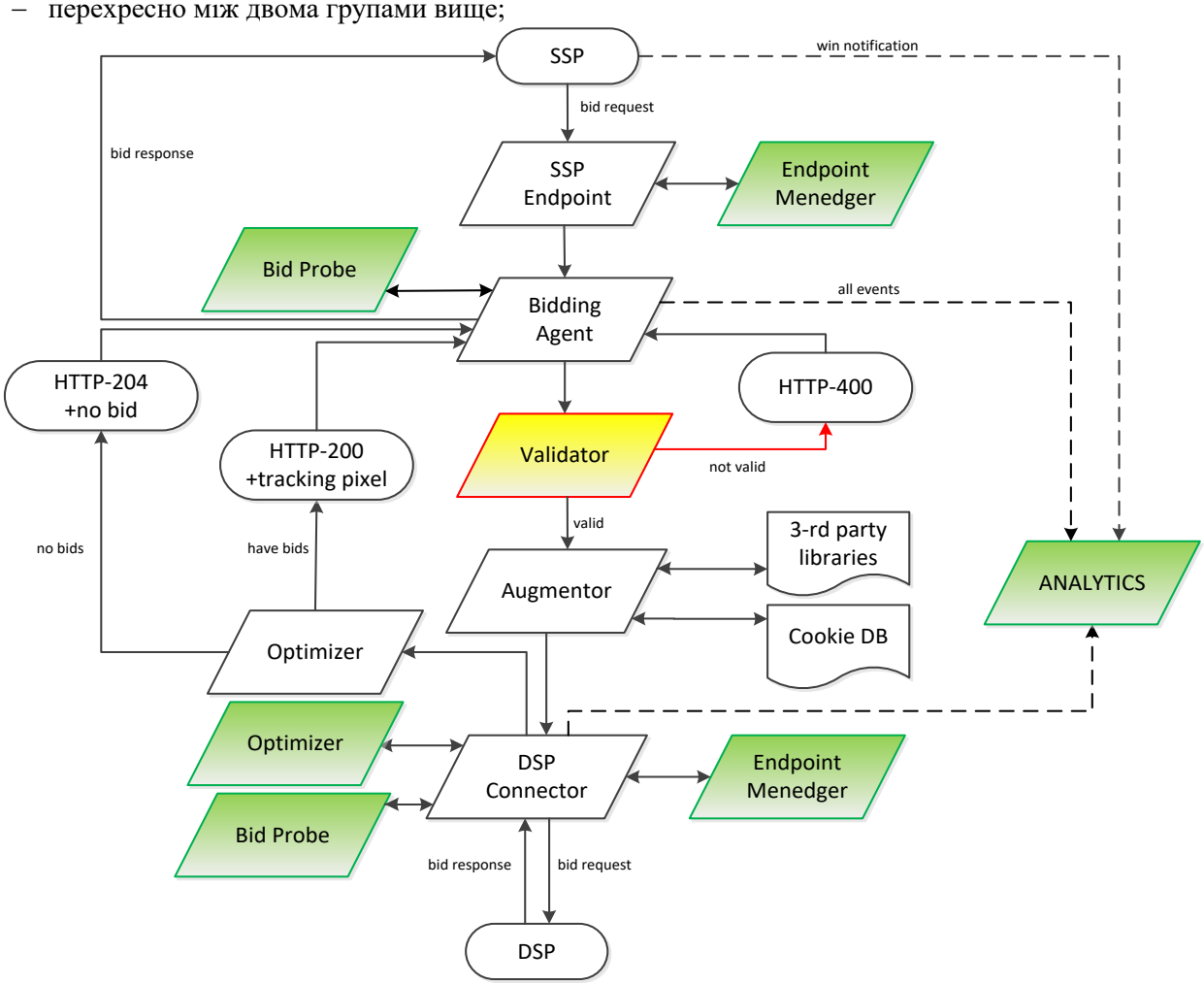


Рис. 1 – Модель роботи системи Ad Exchange

А найважливішою складовою цієї моделі є вперше запропонований авторами Validator (в моделі позначено жовтим кольором), який виставляє рекламу лише для встановленого певного контингенту людей. Реалізований в моделі Validator є основною перевагою представленої системи над іншими.

Найціннішою й найважчою складовою для реалізації нашого продукту є Аналітика, яка була авторами реалізована на DSP платформі (рис. 5). Це єдиний ресурс у всьому додатку, де клієнт може слідкувати за важливими коефіцієнтами виграшу, кліків тощо. Ad Exchange надає можливість контролювати видавцю:

1) встановлення мінімальних CPM для одиниць запасів; фільтрування та блокування оголошень із конфіденційним вмістом або оголошеннями конкурентів;

2) визначення типів оголошень на веб-сторінці, з обранням формату об'яв і стилю оголошення;

3) створення індивідуальних комбінацій шрифтів, кольорів і кутового стилю та їх застосування для кількох графічних оголошень одночасно;

4) вибір місця розташування оголошень на веб-сторінці.

Ad Exchange є посередником між SSP та DSP платформами. За рахунок включення блоку Validator удосконалено відомий алгоритм роботи DSP, SSP.

Звідси витікає висновок, що Ad Exchange є інструментом для автоматизації процесу купівлі та продажу Інтернет контенту з можливістю отримувати більший заробіток та економити час на здійснення операцій.

Розглянемо DSP-SSP системи компанії EPOM, як інструмент розв'язання задачі визначення обсягів купівлі та продажу медіа-контенту.

Для вирішення такої задачі в статті проведено моделювання на основі DSP платформи.

3.5. МОДЕЛЮВАННЯ НА ОСНОВІ DSP ПЛАТФОРМИ.

Вхід до системи починається з контрольної панелі, де рекламодавець може контролювати інтенсивність натисків по його рекламі; яка конверсія та скільки грошей було знято з його рахунку тощо (рис 2).

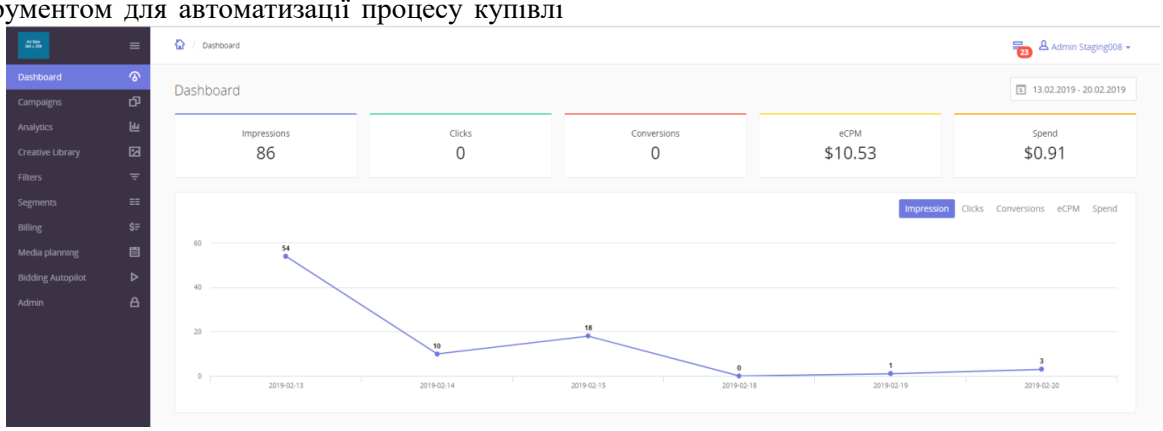


Рис. 2 – Контрольна панель DSP системи

Далі йде сторінка кампаній, де користувач може додавати креативи (типи реклами: банери, відео тощо), рис. 3-4.

З рис. 4 видно, що у однієї кампанії може бути багато креативів, різних різновидів та розмірів.

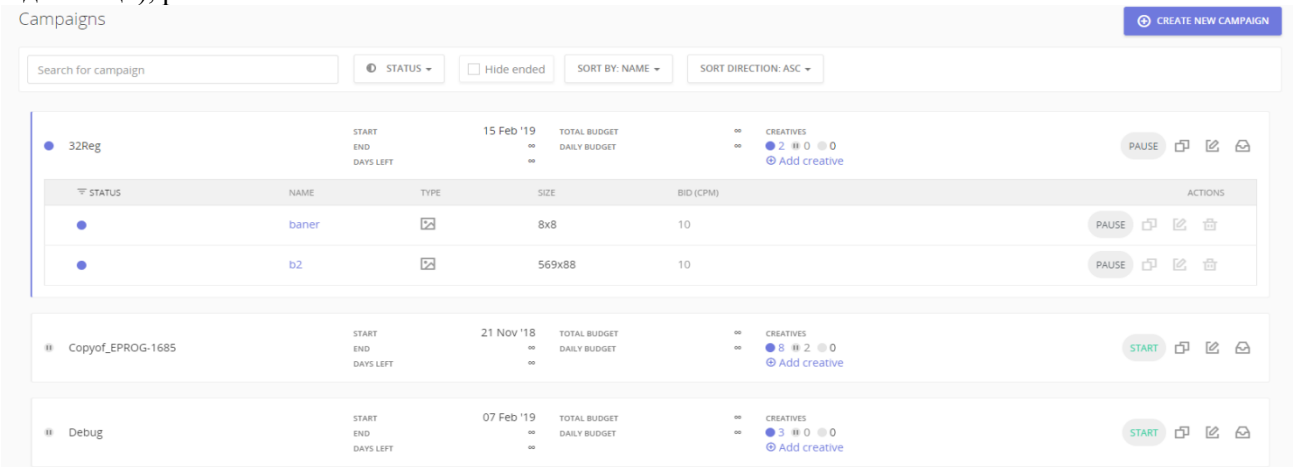


Рис. 3 – Список кампаній з креативами

Вслід за сторінкою управління кампаніями та креативами йде одна з найголовніших та найскладніших областей DSP системи – це аналітика, рис. 5.

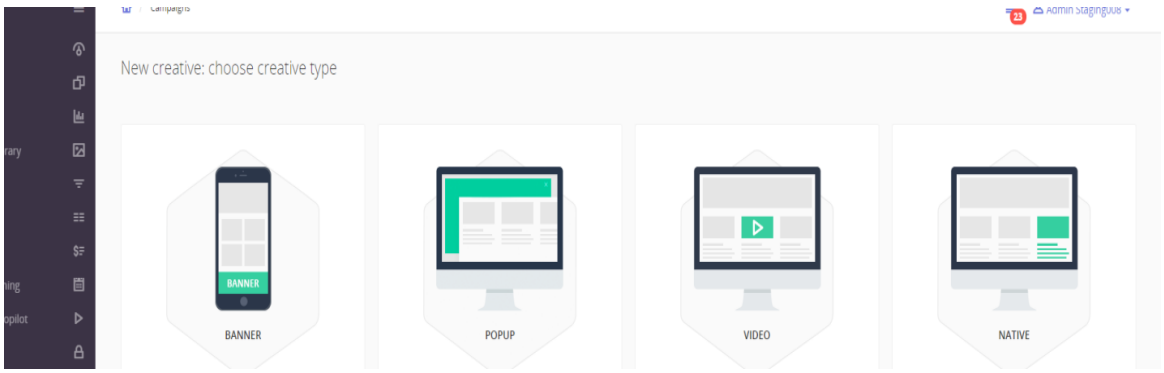
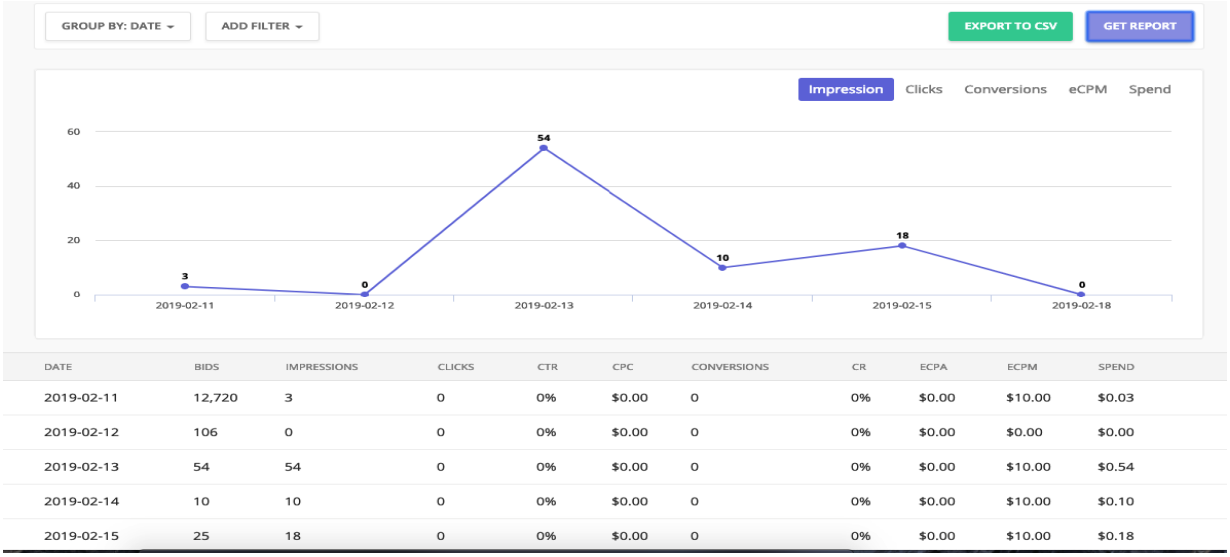


Рис. 4 – Типи креативу

Вслід за сторінкою управління кампаніями та креативами йде одна з найголовніших та найскладніших областей DSP системи – це аналітика, рис. 5.

На рис. 6 зображено таблицю аналітики продукту DSP компанії EPOM.



DATE	BIDS	IMPRESSIONS	CLICKS	CTR	CPC	CONVERSIONS	CR	ECPA	ECPM	SPEND
2019-02-11	12,720	3	0	0%	$0.00	0	0%	$0.00	$10.00	$0.03
2019-02-12	106	0	0	0%	$0.00	0	0%	$0.00	$0.00	$0.00
2019-02-13	54	54	0	0%	$0.00	0	0%	$0.00	$10.00	$0.54
2019-02-14	10	10	0	0%	$0.00	0	0%	$0.00	$10.00	$0.10
2019-02-15	25	18	0	0%	$0.00	0	0%	$0.00	$10.00	$0.18

Рис. 5 – Аналітика DSP платформи

1	2	3	4	5	6	7	8	9	10	11
DATE	BIDS	IMPRESSIONS	CLICKS	CTR	CPC	CONVERSIONS	CR	ECPA	ECPM	SPEND
2018-05-28	3,908,055	1,454,309	0	0%	0.00	0	0%	0.00	$0.08	$119.71
2018-05-29	3,866,707	1,507,460	0	0%	0.00	0	0%	0.00	$0.08	$124.74
2018-05-30	6,136,218	2,061,448	0	0%	0.00	0	0%	0.00	$0.08	$165.92
2018-05-31	7,090,107	2,550,861	0	0%	0.00	0	0%	0.00	$0.07	$178.69
2018-06-01	7,064,578	2,682,524	0	0%	0.00	0	0%	0.00	$0.06	$168.47

Рис. 6 – Результати аналізу роботи DSP платформи

На рис. 6 стовбці за № 1-11 позначено наступне:

1) Group By value (дату, кампанію, ОС тощо).
2) Bids – кількість об'єктів показів надсилається у відповідь на ставку (bid_imps_count).
3) Impressions – кількість отриманих показів, що розраховується за формулою:

$$imp\_count + imp\_nurl\_count; \quad (1)$$

4) Clicks – кількість отриманих кліків за допомогою пікселя показів (click_count).
5) CTR – %, кліків / показів, що розраховується за формулою:

$$\frac{click\_count}{\left(\frac{imp\_count}{imp\_nurl\_count}\right)} \cdot 100\%; \quad (2)$$

6) CPC – витрати / кліки:

$$\frac{dsp\_adv\_price\_sum}{dsp\_adv\_price\_sum} \cdot 1000; \quad (3)$$

7) Conversions – кількість отриманих дій (action_count);
8) CR – %, конверсії / кліки:

$$\left(\frac{action\_count}{click\_count}\right) \cdot 100\%; \quad (4)$$

9) eCPA – витрати / переходи:

$$\frac{dsp\_adv\_price\_sum}{1000 \cdot action\_count}; \quad (5)$$

10) eCPM:

$$\frac{dsp\_adv\_price\_sum}{win\_count}; \quad (6)$$

11) Spend – сума цін пропозиції реклами

$$(dsp\_adv\_price\_sum). \quad (7)$$

У цій системі обов'язково ведеться історія усіх транзакцій, прикріплених за цим аккаунтом, яку було розроблено мною (рис. 7).

Transaction history

ID	DATE AND TIME	TRANSACTION TYPE	DESCRIPTION	SUM	STATUS	PAYMENT TYPE	INVOICE
66a70a132208	06.08.2018 12:00:30	Debit	dvdfg	$3.00	COMPLETED	PayPal	
0b167edf28d4	06.08.2018 11:59:59	Debit	jgj	$55.00	COMPLETED		
e2077dc38bd1	06.08.2018 11:59:22	Debit	345345	$345.00	COMPLETED		
9bfeda2c6a9a	20.07.2018 13:12:43	Debit	dasdasda	$2,342.00	COMPLETED		

4 items

Рис. 7 – Історія транзакцій

3.6. ОБҐРУНТУВАННЯ ДОЦІЛЬНОСТІ ВИКОРИСТАННЯ РЕАЛІЗОВАНИХ ПЛАТФОРМ ДЛЯ АВТОМАТИЗАЦІЇ КУПІВЛІ ТА ПРОДАЖУ МЕДІА КОНТЕНТУ.

Виходячи з отриманих результатів, доцільність використання розглянутих платформ полягає у наступному:

1. Значна економія бюджету. У той час, як більшість рекламних мереж продовжують стягувати плату за тисячі показів невідомій аудиторії – платформи DSP дозволяють купити покази на платформах з зацікавленою аудиторією, тому покупці не переплачують за порожній трафік.
2. Є доступ до джерел трафіку по всьому світу. Програмні платформи продають інвентар, який надходить безпосередньо від постачальників. Останні, в свою чергу, працюють з рекламними майданчиками, сайтами і мобільними додатками. В результаті, після реєстрації рекламодавець отримує доступ до якісних джерел трафіку по всьому світу.
3. Можливість застосування точного таргетингу. DSP покращує свої функціональні можливості за допомогою даних користувачів, щоб допомогти покупцям медійної мережі орієнтуватися на певні аудиторії і залучати потенційних клієнтів за допомогою персоналізованих повідомлень [21].
4. Можливість застосування правил бідингу.
5. Для зручності використання RTB платформ можна встановлювати правила, за якими певний креатив буде припиняти свою роботу, наприклад, якщо конверсія нижче 2 умовних одиниць (рис. 8).

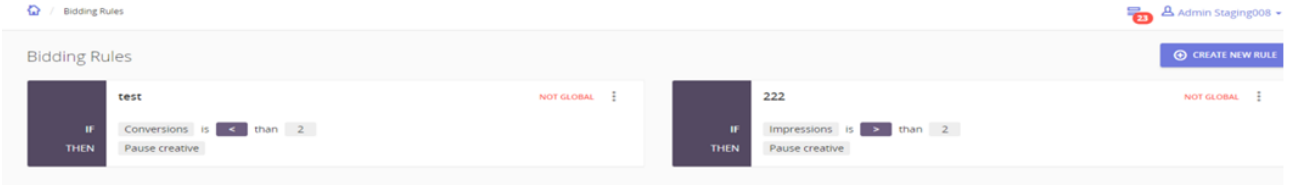


Рис. 8 – Біддинг правила для креативу

1) Можливість здійснення якісного аналізу і статистики.

Реклама DSP піддається аналізу прямо на платформах. Це дозволяє поліпшувати результати і не витрачати бюджет даремно.

2) Прозорість здійснення покупок.

Розроблені авторами правила (рис. 8) відстежують неякісні пропозиції і борються з ризиками.

4. ВИСНОВКИ

Авторами обґрунтовано доцільність використання RTB платформ для автоматизації купівлі та продажу медіа контенту, показано практичне застосування DSP системи, розглянуто аналітику системи, варіації реклами, панель, яка показує клієнтові всю аналітичну інформацію по його рекламі, а також доведено, чим ця система більш досконала від прямих продаж.

З'ясовано, що представлений в статті підхід дає змогу повністю замінити людську роботу і розміщувати рекламу лише для вузько заданого направлення контингенту.

5. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] *Martech Landscape: What Is An Ad Exchange?* [Online]. Available: https://martechtoday. com/ martech -landscape- what- is- an-ad- exchange-161947.

[2] *Trading Desk.* [Online]. Available: https://www.investopedia. com/ terms/ t/ trading desk.asp.

[3] *Зачем используют программатик - баинг?* [Online]. Available: https:// netpeak. net/ru/blog /kak- rabotaet-programmatic/.

[4] *Что такое Programmatic Direct?* [Online]. Available: https://www. cossa.ru/ trends/ 118519/.

[5] *Что такое программатик реклама?* [Online]. Available: https://www.ve.com/ru/blog/ pro-grammatik - reklama.

[6] *Прямая реклама*. [Online]. Available: https:// adbutton. net/ru/ webmaster/profit/advertising-direct.

[7] *Что такое RTB: новые технологии интернет-рекламы*. [Online]. Available: https://habr. com/ru /post/169267/.

[8] *RTB в Украине и мире: что есть и чего ждать*. [Online]. Available: https://sostav. ua/ pub- lication /rtb-v-ukraine-i-mire-chto-est-i-chego-zhdat-68585.html.

[9] *RTB: шаг навстречу новой технологии*. [Online]. Available: https://habr. com/ru/ company/ realweb/blog/295120/.

[10] Е.Л. Головлёва *Основы рекламы*. Академический проект, М, 2018, 191 с.

[11] Г. Грачёв, И. Мельник, Манипулирование личностью: организация, способы и технология информационно-психологического воздействия, Астра-Пресс, М, 2016, 397 с.

[12] *What is Real Time Bidding? How Does it work and Why you Need it in your PPC Campaigns?* [Online] Available: https://www. Omni coreagency. com/what-is-real-time-bidding.

[13] Х. Кафтанджиев, Гармония в рекламных коммуникациях, Эксмо, М, 2015, 368 с.

[14] *Ad Exchange: краткое руководство и что нужно знать издателям*. [Online] Available: http://programmatic.com.ua /2017/09/ad-ex change-kratkoe-rukovodstvo-chto-nyzhno-znat-publisher/.

[15] *Deep Reinforcement Learning for Sponsored Search Real-time Bidding.* [Online] Available: https://arxiv. org/pdf/1803.00259. pdf.

[16] Т.В. Науменко Психологические методы воздействия на массовую аудиторию, *Вопросы психологии*, Вып. 6, С. 63−71, 2014.

[17] *DSP, SSP и Ad exchange – что это такое и как они все дружат?* [Online]. Available: https://maddata.agency/mad-blog/ dsp-ssp-i-ad-exchange-chto-eto-takoe.

[18] М.В. Петрушко. Реклама: культурный контекст, РИП-Холдинг, М., 2014, 186 с.

[19] *Сервер Ad Exchange – не как у других.* [Online]. Available: https://habr.com /ru/ company/ maxilect/blog/421471/.

[20] К. Юнг Человек и его символы, Б.С.К., М., 2018, 304 с.

[21] А.В. Юрасов, Основы электронной коммерции, Дело, М., 2015, 480 с.

Денис Коротін *студент 4-го курсу кафедри програмних систем і технологій факультету інформаційних технологій Київського національного університету ім. Тараса Шевченка.*

Наукові інтереси: інформаційні технології, розроблення web-додатків, розроблення програмного забезпечення.

Світлана Поперешняк *доцент кафедри програмних систем і технологій факультету інформаційних технологій Київського національного університету ім. Тараса Шевченка.*
Наукові інтереси: програмна інженерія, автоматизація процесів виробництва, інформаційні технології, захист інформації, використання багатовимірних статистик для тестування бітової послідовності на випадковість.

Сергій Коротін*, к.т.н., доцент, заступник начальника кафедри авіації інституту авіації та протиповітряної оборони Національного університету оборони України імені Івана Черняховського.*
Наукові інтереси: напрямки підвищення ефективності функціонування інформаційних технологій, теорія автоматичного управління, сучасні методи чисельного інтегрування

DOI
УДК 621.396.946

ІННОВАЦІЙНЕ РІШЕННЯ ДЛЯ LEO-СИСТЕМИ З АРХІТЕКТУРОЮ «РОЗПОДІЛЕНОГО СУПУТНИКА»

Сайко Володимир[1]
orcid.org/0000-0002-3059-6787

Наритник Теодор[2]
orcid.org/0000-0002-4118-0226

Гладких Валерій[3]
orcid.org/0000-0002- 5868-9504

Сивкова Наталія[4]
orcid.org/0000-0002-4934-4109

[1] 02224, м. Київ, Україна. вул. Бальзака, 4, кв. 283, E-mail: vgsaiko@gmail.com
[2] 03170, м. Київ, Україна, вул. Зодчих, 62-Б, кв.93, E-mail: director@mitris.com
[3] 65029, м. Одеса, Україна. вул. Кузнечна, 1, E-mail: informatics.onas@gmail.com
[4]02166, м. Київ, Україна, вул. Мілютенка 23, кв. 192, E-mail: onazkafedratk@gmail.com

Історіястатті:
Надійшла до редакції 19.02.2020
Прийнято 01.03.2020

Ключові слова:
лінії терагерцового діапазону;
розподілений супутник;
низькоорбітальні супутникові системи;
Інтернет речей;
хмарні технології;
тумані обчислення;
граничні обчислення

Анотація: Запропоновано інноваційне рішення для практичної реалізації у LEO- системі з архітектурою «розподіленого супутника», яке може бути використано для забезпечення зв'язку низькоорбітальних космічних апаратів з наземними станціями та користувачами супутникових послуг 5G/IoT. Суть запропонованої розробки в системі низькоорбітального супутникового зв'язку із FC-архітектурою, полягає в тому, що для зменшення затримки при передачі сигналів споживачам та ймовірності перевантаження мережі в перспективну систему низькоорбітального супутникового зв'язку, яка містить штучні супутники Землі, кожен з яких функціонує на навколоземній орбіті і оснащений бортовими ретрансляторами, міжсупутниковий зв'язок, мережу наземних станцій зв'язку і управління штучними супутниками Землі, угруповання низькоорбітальних космічних апаратів (LEO-система), яке включає угрупування кореневих (ведучих) супутників та супутників-ретрансляторів (ведених), навколо кожного кореневого супутника формується мікро-угрупування супутників-ретрансляторів, а функції кореневого супутника в обраній фазовій точці орбітальної площині робочої орбіти виконують міні- або мікро-супутники, які пов'язані в кільцеву мережу лініями зв'язку між супутниками, а - функції супутників-ретрансляторів – кубсати, новим є введення багаторівневої системи граничних хмар, яка представляє собою гетерогенну розподілену обчислювальну хмарну структуру. При цьому граничні хмари багаторівневої системи з'єднані за допомогою надвисокошвидкісних бездротових ліній радіозв'язку терагерцового діапазону та бездротових оптичних систем зв'язку. Наведена методика оцінки часу доступу в запропонованій структурі «туманних обчисленнях» на основі моделі доступу в «туманних обчисленнях» з дозволом колізій джерел даних, що реалізують режим опитування.

Abstract: An innovative solution for practical implementation in a LEO system with a "distributed satellite" architecture that can be used to provide low-orbital spacecraft communications with ground stations and users of 5G / IoT satellite services is proposed. The essence of the proposed development in the system of low-orbital satellite communication with FC-architecture is that to reduce the delay in signaling to consumers and the probability of overloading the network into a prospective system of low-orbital satellite communication, which contains artificial Earth satellites, each of which functions in Earth orbit and equipped with onboard repeaters, inter-satellite communications, a network of ground-based communication and control systems for artificial satellites of the Earth, a grouping of low-orbiting space their devices (LEO-system), which includes the grouping of root (leading) satellites and satellites-repeaters (slave), around each root satellite is formed micro-grouping of satellites-repeaters, and functions of the root satellite in the selected orbital phase of the orbital -or micro-satellites that are connected to the annular network by communication lines between satellites, and - functions of satellites-repeaters - kubsat, new is the introduction of a multilevel boundary cloud system, which is a heterogeneity

well distributed computing cloud structure. At the same time, the boundary clouds of the multilevel system are connected by ultra-high-speed wireless terahertz radio lines and wireless optical communication systems. The technique of estimation of access time in the proposed structure of "fog computing" on the basis of the model of access in "fog computing" with the resolution of collisions of data sources implementing the survey mode is presented.

.

1. ВСТУП

Бурхливий розвиток і стандартизація наземної частини мереж 5G/IMT-2020, а також обмеження для глобального покриття бездротовими стільниковими мережами при використанні міліметрового та субтерагерцового діапазонів хвиль змушують розробників мереж супутникових телекомунікацій звертати увагу і на цей сегмент ринку мобільного супутникового зв'язку [1-2]. Концепція застосування супутникового сегмента 5G, що розглядається сьогодні, заснована на таких передумовах [3]:

• супутниковий сегмент буде інтегруватися з іншими мережами мобільного та фіксованого зв'язку, а не будуть автономною мережею і інтеграція супутникового і наземного сегмента 5G є ядром цього;

• системи космічного зв'язку є фундаментальними компонентами для надійного надання послуг 5G не тільки на території всієї Європи, а й у всіх регіонах світу, весь час і за доступною ціною;

• супутниковий сегмент буде сприяти характеристикам глобальності, збільшення можливостей послуг 5G і вирішення проблем, пов'язаних з підтримкою зростання мультимедійного трафіку, повсюдного покриття, між машинного зв'язку і критично важливих телекомунікаційних місій при оптимізації вартості послуг для кінцевих користувачів;

• космічний сегмент може стати частиною гібридної конфігурації мережі, що складається з поєднання широкомовної і широкосмугової інфраструктур, керованих таким чином, щоб вони забезпечували безперебійну і негайну конвергенцію послуг 5G для всіх кінцевих користувачів.

Вимоги до супутникового сегменту мережі п'ятого покоління будуть визначатися насамперед сукупністю послуг, що надаються мережами 5G, які об'єднані трьома основними бізнес-моделями: розширений мобільний широкосмуговий доступ (Enhanced mobile broadband - eMBB), масове з'єднання пристроїв машинного типу (Massive Machine - Type Communications - mMTC) і вкрай надійний зв'язок з низьким рівнем затримки (uRLLC - ultra-Reliable Low Latency Communications) [4-5]. Супутникові системи зв'язку відомі своєю надійністю і можливістю забезпечувати вимоги щодо затримок сигналів в мережі. Затримка сигналів при використанні геостаціонарних космічних апаратів (КА) буде прийнятною для багатьох додатків мереж 5G. Більш чутливі до затримок програми можуть підтримуватися за допомогою нових низькоорбітальних супутникових мереж, які будуть розгорнуті в майбутньому.

Аналіз пропозицій і технологічних проектів провідних виробників по використанню супутникових мереж для розширення можливостей мереж 5G показує [5-6], що багато телекомунікаційних компаній вже зробили спроби презентації своїх проектів, придатних для розгортання супутникового сегмента. Але недоліком відомих технічних рішень є:

- складність реалізації ресурсу різних супутникових систем;
- суттєве підвищення вартості за рахунок використання значної кількості супутників;
- постійної потреби у визначенні взаємної дислокації супутників низькоорбітальних космічних апаратів.

При цьому розширення території обслуговування із прийнятною імовірністю необхідної якості передачі суттєво зменшується оскільки в каналі передачі використовують супутники, які постійно рухаються відносно геостаціонарного супутника із значною швидкістю і на значних відрізках території. Таким чином, на даний час існує багато проектів, ідей, планів по реалізації і впровадженню нових систем, але зараз вони «малоефективні» тому не існує точної впевненості в практичній реалізації і довговічності роботи систем. Наприклад, у LEO-HTS вже зараз існують серйозні побоювання щодо спільного використання спектра з існуючими мережами і внаслідок виникнення частотних завад.

2. ПОСТАНОВКА ПРОБЛЕМИ

З даної точки зору для реалізації нових послуг становить інтерес розроблене авторами технічне рішенні [7], в якому система низькоорбітального супутникового зв’язку, представляє угруповання низькоорбітальних космічних апаратів (LEO-система) з архітектурою "розподіленого супутника", та яка включає угрупування кореневих (ведучих) супутників та супутників-ретрансляторів (ведених). Навколо кожного кореневого супутника формується мікро-угрупування супутників-ретрасляторів, яке називається «розподілений супутник». Функції кореневого супутника в обраній фазовій точці

орбітальної площині робочої орбіти виконують міні-або мікро-супутники, а функції супутників-ретрансляторів – кубсати. Кореневі супутники пов'язані між собою в кільцеву мережу високошвидкісними лініями зв'язку між супутниками. Геометричний розмір «розподіленого супутника» - область навколо кореневого супутника радіусом приблизно 1 км. Це означає, що кубсати здійснюють груповий політ на відстані не більше 1 км від кореневого супутника. Космічний сегмент LEO-системи складається з декількох орбітальних площин, що мають однакову кількість розподілених супутників, однаковий спосіб, і відрізняються довготою висхідного вузла. У кожній орбітальній площині розподілені супутники рівномірно розміщені з однаковою відносною істинної аномалією при цьому кожен розподілений супутник пов'язаний з двома сусідніми розподіленими супутниками в своїй орбітальній площині і з двома найближчими розподіленими супутниками в двох сусідніх орбітальних площинах - по одному в кожній.

Архітектура «розподіленого супутника» дозволяє проводити вимірювання параметрів орбітального руху кубсата на базі вимірювань їх взаємного розташування. В результаті можливо підвищити ефективність використання кубсатів, відключивши частину обладнання, необхідного для вимірювання параметрів. До складу корисного навантаження супутників-ретрансляторів включені SDR-модулі (програмно-конфігуровані модулі), призначені для функціонування і обробки транспортних потоків для передачі і прийому інформації. Для забезпечення зв'язку всередині «розподіленого супутника» використовується бездротова мережа.

Недоліком відомого рішення є те, що у ньому не визначена архітектура міжсупутникового каналу зв’язку, що не дозволяє оцінити його технічні характеристики та відповідно технічні можливості системи низькоорбітального супутникового радіозв’язку для надання якісних послуг в інтегрованих мережах зв’язку 5G та IoT. Крім того, для реалізації в космічному сегменті LEO- системи FC-архітектури в склад «розподіленого супутника» включений окремий супутник-обчислювач, завданням якого є виконання необхідних обчислень для забезпечення функціонування обслуговування пристроїв IoT в межах зони обслуговування "розподіленого супутника". Таке рішення не забезпечує гнучкість та високу надійність системи при перевантаженнях мережі та при виході з ладу окремого супутника-обчислювача. Задачею запропонованого нового рішення є удосконалення системи низькоорбітального супутникового зв’язку шляхом:

1. Реалізації міжсупутникового зв’язку на основі радіоканалу, створеного в терагерцовому 140 ГГц діапазоні.

2. Зменшення затримки при передачі сигналів споживачам.

3. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

3.1 ТЕХНІЧНІ АСПЕКТИ ІННОВАЦІЙНОГО РІШЕННЯ

Поставлена задача вирішується тим, що в систему низькоорбітального супутникового зв’язку додатково введено багаторівневу систему граничних хмар, що представляє собою гетерогенну розподілену обчислювальну хмарну структуру. При цьому, автори виходили з того, що таку запропоновану систему можна розглядати як тришарову систему, засновану на структурі мережі 5G/IMT2020, запропонованої NGMN [8]. У цій структурі розділені апаратне і програмне забезпечення, а для сумісності послуг на рівні додатків використовуються інтерфейси прикладного програмування (API).

Ідея нового підходу до побудови багаторівневої хмарної системи для низькоорбітальних супутникових мереж зв'язку полягає в тому, що зони обслуговування окремих супутників-ретрансляторів з'єднуються з піко хмарами з досить невеликими обчислювальними можливостями для виконання граничних обчислень. Ці піко хмари з'єднаються з макро хмарами, які в свою чергу, поєднуються з окремими супутниками-ретрансляторів і мають великі обчислювальні можливості. Крім того, кожен супутник-ретранслятор з'єднується за допомогою радіоліній терагерцового діапазону з мікро хмарою з можливості яких обмежені. Ядро «розподіленого супутника» (кореневий супутник) забезпечує взаємодію мікро хмар в системі в цілому. Запропонована система зменшує затримку при передачі сигналів споживачам і ймовірність перевантаження мережі, одночасно збільшуючи гнучкість її побудови і доступність.

Сутність запропонованого рішення пояснюється блок схемою, де зображена схема реалізації запропонованої системи низькоорбітальної системи супутникового зв’язку мережі 5G/IoT (рис.1).

На ньому позначені:

1 – Кореневий супутник.

2,3,4,5 – Супутники-ретранслятори.

6 – Фідерні лінії.

7 – Зона обслуговування системи супутникового зв’язку.

8 – Піко хмара хмарної структури.

9 – Мікро хмара хмарної структури.

10 – Макро хмара хмарної структури.

11 – Головна хмара хмарної структури.

12 – Міжсупутникова лінія передачі.

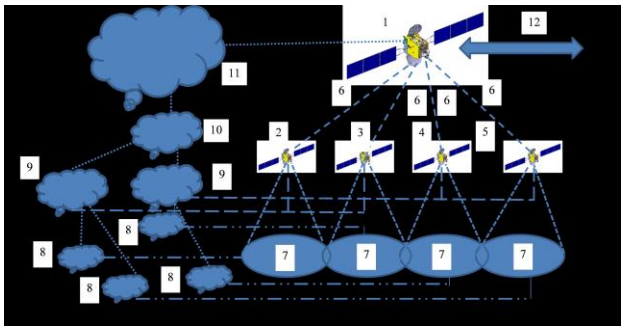


Рис. 1 - Блок-схема реалізації запропонованої системи низькоорбітальної системи супутникового зв'язку мережі 5G/IoT

Схема функціонування «розподіленого супутника» в складі низькоорбітальної системи супутникового зв'язку мережі 5G/IoT працює наступним чином.

Інформаційним та інтелектуальним ядром розподіленого супутника є кореневий супутник 1 (див. блок-схему). Супутники-ретранслятори 2,3,4,5 розподіленого супутника формують промінь / промені користувачів з обмеженою зоною обслуговування. Сукупність променів, які формуються супутниками-ретрансляторами, складають зону обслуговування 7 LEO-системи. Вимоги щодо інтегральної зони обслуговування LEO-системи (географічна зона обслуговування) визначають вимоги до кількості розподілених супутників в системі в цілому.

Фідерні лінії 6 забезпечують з'єднання кореневого супутника з супутниками-ретрансляторами та призначені для передачі транспортного цифрового потоку відповідного формату до кінцевих користувачів. Фідерна лінія 6 між кореневим супутником і супутником - ретранслятором є внутрішньою лінію зв'язку між супутниками в складі розподіленого супутника. Ця лінія - комбінована радіолінія, яка забезпечує дуплексну передачу інформації, вимір похилої дальності і взаємного кутового положення між кореневим супутником 1 і супутниками-ретрансляторами 2,3,4,5.

Розподілені супутники в LEO-системі пов'язані між собою міжсупутниковими лініями 12 терагерцового зв'язку, які формують магістральну мережу LEO - системи. Кожен розподілений супутник пов'язаний з двома сусідніми розподіленими супутниками в своїй орбітальній площині і з двома найближчими розподіленими супутниками в двох сусідніх орбітальних площинах - по одному в кожній з них. У складі розподіленого супутника функції підтримки лінії зв'язку між супутниками покладені на кореневий супутник 1 який оснащений модулями фідерних ліній 6 та відповідними високочастотними приймально-передавальними супутниками-ретрансляторами.

Запропонована багаторівнева система для низькоорбітальної супутникової системи радіозв'язку мережі 5G/IoT являє собою розвиток систем хмарних обчислень для мереж зв'язку від централізованої системи до гетерогенної розподіленої системі.

Особливістю запропонованої системи, яка містить чотири рівні хмар 8-11, є те, що хмари з'єднані за допомогою надвисокошвидкісних бездротових ліній радіозв'язку терагерцового діапазону. Така побудова передбачає і використання бездротових оптичних систем зв'язку, які є однією з найважливіших особливостей майбутніх мереж 5G. При цьому відстань між призначеним для користувача обладнанням і найближчою хмарою в мережі становить всього один бездротовий перехід. Це дозволяє забезпечити зменшення затримка передачі даних та ймовірність перевантаження в мережі, збільшити доступність мережі та поліпшити безпеку, оскільки додатки надаються користувачам всередині мережі.

3.2 Методика оцінки часу доступу у запропонованій структурі в «туманних обчисленнях».

Вивантаження даних від користувача на рівень хмари у запропонованому рішенні включає в себе три складових затримки: затримку по висхідним і спадним лініях, затримку поширення і затримку обробки інформації.

Для оцінки часу доступу в «туманних обчисленнях» пропонується модель доступу в них з дозволом колізій джерел даних, що реалізують режим опитування. В системі з режимом опитування є центральний вузол (маршрутизатор, шлюз або сервер) і деяка кількість джерел даних, яка в силу динамічних змін заздалегідь «невідомо» вузлу. Джерела даних починають передачу даних тільки за запитом центрального вузла. Якщо у джерел даних немає підготовленого для передачі пакета даних, то спеціальний логічний механізм джерел даних формує ідентифікаційний пакет, що містить унікальний ідентифікаційний номер джерел даних. Завдання центрального вузла ідентифікувати всі джерела даних, які знаходяться в зоні його обслуговування і якщо у джерел даних є інформація для передачі, то прийняти їх.

Процес опитування - часовий інтервал, який розділений на блоки - вікна. Вікна - нефіксовані відрізки часу, розмір яких визначається кількістю слотів, на які він ділиться. Розмір слота - фіксована величина для кожного джерела даних. Так як слот - це теж часовий інтервал, його розмір визначається швидкістю передачі

даних від джерел даних до вузла, тобто його визначає обладнання, що використовується в системі.

Тоді загальна затримка на передачу по висхідним і низхідним лініях (Tv, Tn) може бути обчислена на основі наступних рівнянь [9,10]:

$$Tv = (1 + Vbo)(Db / Vb),$$

$$Tn = (1 + Vno)(Dn / Vn),$$

де Vbo, Vno - швидкості передачі для висхідної і низхідної ліній відповідно, при яких настає відмова в обслуговуванні, Db, Dn - загальне число переданих біт по висхідній і низхідній лініях відповідно і, нарешті, Vb, Vn - швидкості передачі даних для висхідної і низхідної ліній відповідно.

Затримка поширення Tr є функцією відстані і може бути обчислена в такий спосіб:

$$Tr = Rob / Vr,$$

де Rob - відстань між мобільним користувачем і хмарою, а Vr - швидкість поширення. Затримка через обробку даних ґрунтується на числі необхідних операцій і швидкості обробки інформації процесором в хмарі. Затримка через обробку даних $Tobr$ обчислюється таким чином:

$$Tobr = N / Vproc,$$

де N - загальне число необхідних операцій і $Vproc$ - швидкість процесора хмари. Резюмуючи все вищесказане, загальна затримка може бути обчислена як сума зазначених трьох затримок. Тоді загальна затримка вивантаження даних в хмару $Tzag$ обчислюється таким чином:

$$Tzag = Ty + Tn + Tr + Tobc,$$

Відповідно повний цикл взаємодії M джерел даних і центрального вузла становить $T = MTzag$.

Якщо кожне джерело даних в середньому λ пакетів в секунду, то при наявності в системі M джерел даних, загальна інтенсивність потоку даних складатиме λM пакетів в секунду, а середній інтервал між їх надходження $1/\lambda M$. Отже, щоб уникнути необмеженого зростання черг, час обслуговування $Tzag$ повинно відповідати умові:

$$Tzag \le 1/\lambda M,$$

Вирішуючи (6) при умові $Vb = Vn$ щодо необхідної швидкості передачі даних для висхідної і низхідної радіоліній $V = Vb + Vn,$ отримаємо умову гарантування відсутності черги для кожного джерела даних і відповідно можна сформулювати вимоги до пропускної спроможності радіоканалу.

Ґрунтуючись на останньому виразі, можна сказати, що загальна затримка, в основному, залежить від відстані. Таким чином, в разі введення макро хмар затримка менша, ніж затримка при використанні ресурсів ядра мережі.

Введення рівня макро хмари на супутниках-ретрансляторах і забезпечення взаємозв'язку макро хмари за допомогою бездротових високошвидкісних радіосистем передачі терагерцового діапазону дозволяє також використовувати новий спосіб вивантаження трафіку, що при спільному застосуванні з технологією взаємодії пристрій-пристрій D2D (Developer-to-Developer) значно покращує характеристики якості обслуговування мережі 5G/IoT.

Результати попереднього моделювання показали, що в порівнянні з традиційним рішенням з побудови ядра мережі Evolved Packet Core, заснованим на версії 15 3GPP, запропоноване рішення забезпечує зменшення кругової затримки в середньому на 30-40%.

Дійсним технічним рішенням для зменшення взаємних завад між угрупуванням кореневих (ведучих) супутників та супутників-ретрансляторів (ведених), які знаходяться в радіусі приблизно 1 км, та відповідно мінімізації спотворень інформаційного сигналу при реалізації міжсупутникового зв'язку дану лінію створено в терагерцовому 128-134 ГГц діапазоні [11-13].

Подальша перспектива використання даного рішення вбачається в наступному. Для створення інноваційної архітектури низькоорбітальної інфокомунікаційної системи зв'язку сьогодні активно проводяться науково-технічні розробки. Результати таких досліджень знайдуть застосування в реалізації концепції Інтернет-речей, в системах дослідження Землі з космосу для наукових і економічних завдань зондування поверхні Землі, для інформаційного забезпечення силових структур і т.д. У вітчизняних науково-дослідних роботах, на відміну від світових, розробляється і досліджується комплекс моделей і методів для низькоорбітальних супутникових мереж, побудованих на основі архітектури «розподіленого супутника». Запропонована авторська розробка дозволяє вирішити ряд науково-технічних задач для створення нового виду мереж - низькоорбітальних супутникових мереж, які, як правило, є частиною комплексу взаємодіючих мереж електрозв'язку систем зв'язку загального користування (СЗЗК) і призначені для надання нових послуг як

користувачам СЗЗК, так і пристроям цих мереж при міжмашинній взаємодії M2M (Machine-to-Machine) при реалізації концепції Інтернету речей.

Серед науково-практичних завдань, які необхідно вирішувати для цього, основними є наступні:

• аналіз концепції Інтернету речей й сучасних технологій телекомунікацій та обґрунтування необхідності створення нового виду мереж зв'язку загального користування - низькоорбітальних супутникових мереж, побудованих на основі архітектури «розподіленого супутника»;

• визначення теоретичних і практичних напрямків досліджень в області низькоорбітальних супутникових мереж, побудованих на основі архітектури «розподіленого супутника»;

• розробка моделей і способів підвищення завадостійкості каналів радіозв'язку низькоорбітальних супутникових мереж, побудованих на основі архітектури «розподіленого супутника»;

• розробка методів використання низькоорбітального системи супутникового зв'язку з архітектурою «розподіленого супутника» для бездротових сенсорних мереж з метою збільшення їх зв'язності.

4. ВИСНОВКИ

Запропоновано багаторівневу систему туманних обчислень для підвищення ефективності функціонування низькоорбітальних супутникових систем зв'язку при забезпечені вимог мереж і систем п'ятого покоління 5G, що представляє собою розвиток систем хмарних обчислень для мереж радіозв'язку від централізованої системи до гетерогенної розподіленої системі. Наведена методика оцінки часу доступу в запропонованій структурі на основі моделі доступу в «туманних обчисленнях» з дозволом колізій джерел даних.

5. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] В. Тихвинский, "Спутниковая связь в будущей инфраструктуре 5G", Connect, № 7-8, с. 104-107, 2018.

[2] "Рынок спутникового Интернет вещей в перспективе до 2024-2030 года", Технологии и средства связи. Специальный выпуск "Спутниковая связь и вещание -2019", № 4, с. 25-30, 2019.

[3] Jonas Eneberg - Inmarsat - Satellite Role in 5G. 2017. [Online]. Available: https://www.slideshare.net/TechUK/jonas-eneberg-inmarsat-satellite-role-in-5g. Accessed on: May 18, 2017.

[4] А.Минов, А.Бабин, "Спутниковая связь для Интернета вещей", Connect, № 5-6, с. 112-116, 2017.

[5] В. Тихвинский, М. Стрелец, "Перспективы создания спутникового сегмента 5G" , Первая миля, № 1, с. 104-107, 2018.

[6] "Подготовка к внедрению 5G: возможности и проблемы," Отчет МСЭ, 2018. [Электронный ресурс]. Режим доступа: https://www.itu.int/dms _pub/itu-d/opb/pref/D-PREF-BB.5G_01-2018-PDF-R.pdf. Дата обращения: Янв.19, 2020.

[7] Т.М. Наритник, В.Г. Сайко, Г.Л. Авдєєнко, В.Я. Казіміренко, С.В. Сарапулов, " Система низькоорбіталь-ного супутникового зв'язку", Н 04 В 7/185. патент 134409 Україна, травень 10, 2019.

[8] 5G white paper | NGMN. Next generation mobile networks, Alliance 2017.[Online]. Available: https://www.ngmn.org/work-programmer/5g-white-paper.html. Accessed on: May 19, 2017.

[9] [A. Mukherjee, and D. G. Roy "A Power and Latency Aware Cloudlet Selection Strategy for Multi-Cloudlet Environment," IEEE Transactions on Cloud Computing Mag., vol. 3, pp.44-48, July 2016.

[10] D. Szabo, A.Gulyas, F. Fitzek, and D. Lucani, "Towards the tactile internet: Decreasing communication lateen-cy with network coding and software defined networking," In European Wireless 2015. 21th European Wireless Conference, May 2015. pp. 1-6.

[11] V. Saiko, S. Toliupa, V. Nakonechnyi, and Dakov Serhii, "The method for reducing probability of incorrect data reception in radio channels of terahertz frequency range," 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Feb. 2018.

[12] В.Г.Сайко, та Т.М. Наритник, Безпроводові системи зв'язку терагерцо-вого діапазону: монографія. Deutsch: Видавництво "LAP LAMBERT Academic Publishing ", 2019.

[13] G. Avdeetiko, "Simulation of a terahertz band wireless telecommunication system based on the use of IR-UWB signals," Telecommunications and Radio Engineering, vol.78 (10), pp. 891 – 909, 2019.

Сайко Володимир Григорович

Доктор технічних наук, професор, професор кафедри прикладних інформаційних систем Київського національного університету імені Тараса Шевченка.

Інтереси: радиоінженерія, інформаційні технології.

Наука: автор більше 270 наукових праць, 4 монографій, 30 патентів на винахід.

Наритник Теодор Миколайович

Кандидат технічних наук, професор, директор Інституту електроніки та зв'язку Української академії наук.

Інтереси: телекомунікаційні системи терагерцового діапазону,.

Наука: автор більше 330 наукових праць, 40 патентів на винаходи.

Гладких Валерій Миколайович

Кандидат технічних наук, завідувач кафедри телекомунікацій Одеської національної академії імені О.С. Попова.

Інтереси: інформаційні технології, інженерія програмного забезпечення.

Наука: автор більше 105 наукових праць, 15 патентів на винахід.

Сивкова Наталія Максимівна

Аспірант кафедри телекомунікацій Одеської національної академії імені О.С. Попова.

Інтереси: інформаційні технології, інженерія програмного забезпечення.

Наука: автор більше 10 наукових праць.

DOI
УДК

1НСТРУКЦ1Я ПО П1ДГОТОВЦ1 СТАТТІ ДО ДРУКУ

Перший автор 1м'я 1), другий автор 1м'я 2)

1) Поштона адреса, адреса електронно! пошти, веб-адреса (URL)
2) Поштова адреса, адреса електронно! пошти, веб-адреса (URL)

Історія статті:
Надійпла до редакції ...
Прийнято ...

Ключові слова:
ключове слово;
ключове слово;
ключове слово;
...
ключове слово (Слова до 10).

Анотація: Цей документ являє собою необхідний макет документів, що подається для публікації в науковому журналі Інформаційні системи та технології безпеки. Тези повинні підсумувати зміст роботи і має складати не менше 1800 знаків.

Abstract: This document presents the required layout of papers to be submitted for publication at the scientific journal Information systems and technologies security. The abstract must summarize the paper content and should have a reasonable number of characters -1800.

1. ФОРМАТУВАННЯ

1.1 ЗАГОЛОВОК

Назва статті має бути інформативною в межах розумної кількості слів, але не може містити більше 90 символів. Будь ласка, уникайте скорочень і формул, де це можливо.

1.2 АВТОР ПУБЛІКАЦІЇ І АФІЛІАЦІЯ

Там, де прізвище може бути неоднозначним (наприклад, подвійне ім'я), прохання вказати це явно. Вкажіть афіліацію авторів (де була зроблена фактичної робота) нижче імен. Вкажіть повну поштову адресу кожного учасника, включаючи назву країни і адресу електронної пошти кожного автора. Будь ласка, відзначте відповідного автора.

1.3 АНОТАЦІЯ

Потрібне коротке і фактологічне резюме. У анотації коротко викладаються мета дослідження, основні результати і основні висновки. Анотація часто видається окремо від статті, тому вона має розкривати зміст статті. З цієї причини слід уникати посилань, але якщо це необхідно, то посилатися на автора(авторів) і рік(роки). Крім того, слід уникати нестандартних або незвичайних абревіатур, але в разі необхідності вони повинні бути визначені при їх першій згадці в самій анотації. Анотація повинна підсумувати зміст статті і мати 1800 знаків.

1.4 КЛЮЧОВІ СЛОВА

Відразу після анотації вкажіть максимум 10 ключових слів, використовуючи американську орфографію та уникаючи загальних і множинних термінів і декількох понять (уникайте, наприклад, "і", "з"). Будьте обережні з абревіатурами: тільки абревіатури, чітко встановлені в цій області, можуть бути використані. Ці ключові слова будуть використовуватися для цілей індексації.

1.5 ВСТУП

Необхідно мати адекватне резюме для опису нинішнього стану справ або результатів.

1.6 Матеріали і методи

Прохання представити досить детальну інформацію, з тим щоб можна було відтворити цю роботу. Методи, вже опубліковані, повинні бути позначені посиланням: повинні бути описані тільки відповідні зміни.

1.7 РЕЗУЛЬТАТИ

Результати повинні бути чіткими і короткими.

1.8 ВИСНОВКИ

Основні висновки дослідження можуть бути представлені в короткому розділі висновків, який може бути самостійним або являти собою підрозділ обговорення або розділу результатів і обговорення.

1.9 ПІДРОЗДІЛИ

Розділіть роботу на чітко визначені і пронумеровані розділи. Підрозділи повинні бути пронумеровані 1.1, 1.2, і т. д. (Анотація не включена в нумерацію розділів). Використовуйте цю нумерацію також для внутрішніх перехресних посилань. Будь-якому підрозділу може бути дано короткий заголовок. Кожен заголовок повинен відображатися в окремому рядку.

1.10 ФОРМУЛИ

Математичні вирази і грецькі або інші символи повинні бути написані ясно з достатньою відстанню.

1.11 ОДИНИЦІ

Дотримуйтесь міжнародно визнаних правил і умовностей: використовуючи міжнародну систему одиниць (СІ). Якщо згадуються інші одиниці, будь ласка, дайте їх еквівалент в СІ.

1.12 ВИНОСКА

Виноски слід використовувати економно. Кількість їх послідовно по всій статті нумерується

арабськими цифрами нижнього індексу. Багато wordprocessors будують виноски в тексті, і ця функція може бути використана. Якщо це не так, вкажіть положення виносок в тексті і представте самі виноски окремо в кінці статті. Не включайте виноски в список посилань.

1.13 ПОСИЛАННЯ

Посилання повинні відображатися у вигляді окремої бібліографії в кінці статті, пронумеровані цифрами в квадратних дужках.

Посилання в тексті наводяться в квадратних дужках. Назви журналів повинні бути скорочені. Прохання вказати Тома журналів, номери випусків (номерів) і номери сторінок.

Кількість посилань налічує більше 20. Відсоток власних публікацій повинен бути менше 20%.

1.14 ФІГУРИ

Електронна версія в EPS (Encapsulated Postscript) найбільш прийнятна. Ілюстрації повинні бути чіткими, без шуму і гарною контрастністю.

1.15 БУДЬ-ЛАСКА НЕ:

- Використовуйте файли, оптимізовані для використання на екрані (наприклад, GIF, BMP, PICT, WPG) розрішення занадто низьке;
- Використовуйте файли, які мають занадто низьке розрішенням;
- Використовуйте рисунки, які непропорційно великі для вмісту.

2. ФОРМАТУВАННЯ

Автори повинні представити готовий документ у форматі MS Word по електронній пошті. Використовуйте формат паперу А4 (210 х 297 мм). Залиште поля 20 мм вгорі, внизу, зліва і справа. Розмір готової до друку статті має складати не більше восьми сторінок без авторських фотографій і резюме. Авторам рекомендується залишити на останній сторінці місце для фотографій і коротких авторських резюме. Будь ласка, не використовуйте нумерацію сторінок у своїх роботах.

3. МАКЕТ

Текст повинен бути набраний через один інтервал. Використовуйте Times New Roman шрифт для звичайного тексту.

Центр заголовка (шрифт Arial, 14 пунктів, напівжирний), ім'я автора (авторів) (шрифт Arial, 12 пунктів, напівжирний) організація(ї), адресу(адреси), адресу електронної пошти та веб-адресу (Times New Roman, розмір 10 пунктів, звичайний). Залиште один порожній рядок під заголовком і приналежності.

Потім вставте коротку анотацію на папері курсивом (не більше 150 слів). Для слова анотація використовуйте напівжирний шрифт. Після цього вставте один порожній рядок.

Нижче вставте список (з одним і двома порожніми рядками) до 10 ключових слів. Ключові слова виділяти жирним.

Введіть текст статті (11 пунктів, звичайний) в дві колонки шириною 8,25 мм, розділені пропуском 5 мм. збалансуйте довжину стовпців, особливо на останній сторінці статті. Перший рядок кожного абзацу повинен мати відступ 5 мм. Заголовки повинні бути пронумеровані. Не використовуйте римські номери. Використовуйте тільки один рівень заголовків (Arial, 12 пунктів, напівжирний) і передуйте кожному з них одним порожнім рядком. Використовуйте інтервал 6 пунктів після абзацу.

Помістіть повний список літератури в кінці документа. Будь ласка, розмістіть посилання відповідно до порядку їх появи в тексті. Укладіть числа в квадратні дужки.

4. ФОРМУЛИ

Рівняння повинні бути поміщені в окремі рядки і пронумеровані. Використовуйте окремі клітинки таблиці для розміщення рівняння і його номери. Не використовуйте рівняння всередині абзацу. Цифри повинні бути укладені в квадратні дужки і вирівняні по правому краю. Будь ласка, для формул використовуйте Microsoft Equation з таким шрифтом: звичайний – 12 пт, великий індекс – 7 пт, дрібний індекс – 5 пт, великий символ – 18 пт, дрібний символ – 12 пт. Залиште один порожній рядок над і під кожним рівнянням. Наприклад, вираз

$$y = a * x + b, \qquad (1)$$

де a - коефіцієнт, x - аргумент, b - зміщення.

5. РИСУНКИ І ТАБЛИЦІ

Рисунки і таблиці повинні бути пронумеровані, мати самостійний заголовок. Підписи до рисунків повинні бути нижче рисунків; підписи до таблиць повинні бути вище таблиць. Крім того, уникайте розміщення рисунків і таблиць до їх першої згадки в тексті.

Текст підпису до рисунка повинен бути 10 пт і жирним шрифтом. Використовуйте інтервал 6пт до і після. Залиште один рядок після таблиць, якщо заголовок розділу не слід. Заголовки таблиць і малюнків повинні бути центровані. Великі цифри і таблиці можуть охоплювати обидва стовпці. Залиште один порожній рядок перед заголовком таблиці. Не відступайте від підписів. Не використовуйте символи менше 10 пт на рисунках.

Всі рисунки, графіки та фотографії повинні бути пронумеровані і вказані в основному тексті. Абсциси і ординати всіх графіків повинні бути позначені символами і одиницями, як на рис.1.

Використовуйте скорочення «Рис.» навіть на початку абзацу. Текст на малюнках має бути Times New Roman.

Не використовуйте скорочення в назвах, якщо вони не є неминучими.

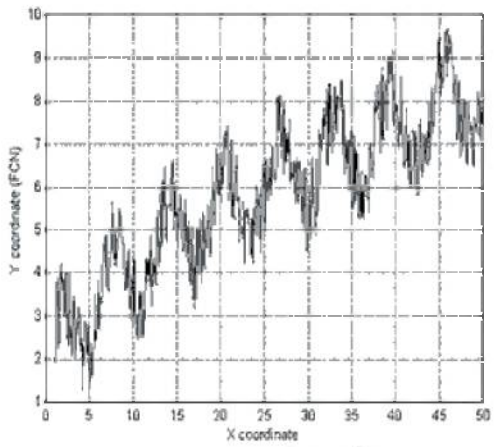


Рис. 1 - Приклад графа. Зверніть увагу, як центрований підпис

Всі рисунки, графіки і фотографії повинні бути чорно-білими (або відтінків сірого).

Розміри шрифтів, які будуть використовуватися в різних частинах документа наведені в таблиці 1.

Таблиця 1. Розміри шрифтів

Розділ документа	Розмір шрифту і стиль
Назва	14, Arial, напівжирний
Ім'я автора	12, Arial, напівжирний
Адреса	10, Times New Roman, Regular
Анотація	10, Times New Roman, курсив
Ключові слова	10, Times New Roman, курсив
Заголовки	12, Arial, усі великі літери, напівжирний
Підзаголовки	12, Arial, усі великі літери, Regular
текст тіла	11, Times New Roman, Regular
Підписи до таблиць	10, Times New Roman, міжрядковий інтервал 6 до і після, напівжирний
Підписи до рисунків	10, Times New Roman, міжрядковий інтервал 6 пунктів до і після, напівжирний

6. ВИСНОВОК

Ми заздалегідь дякуємо вам за надання ретельно підготовлених документів, які можуть бути відправлені для публікації без змін.

7. СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

Стиль посилання

Текст: вказати посилання за номером (номерами) у квадратних дужках відповідно до тексту. Можна посилатися на фактичних авторів, але завжди повинен бути вказаний номер посилання (и).

Зразок.":.... як показано в роботах [3,6]. Барнабі і Джонс [8] отримали інший результат ...."

Список: пронумерувати посилання (цифри в квадратних дужках) в списку в тому порядку, в якому вони вказані в тексті.

Приклад:

Посилання на публікацію журналу:

[1] J. van der Geer, J.A.J. Hanraads, R.A. Lupton, "The art of writing a scientific article," *Journal of Science Communication*, Vol. 16, Issue 3, pp. 51-59, 2000.

Посилання на книгу або звіт:

[2] W. Strunk Jr., E.B. White, *The Elements of Style*, third ed., Macmillan, New York, 1979, 350 p.

Посилання на главу в відредагованою книзі:

[3] G.R. Mettam, L.B. Adams, *How to prepare an electronic version of your article*, in: B.S. Jones, R.Z. Smith (Eds.), Introduction to the Electronic Age, E-Publishing Inc., New York, 1999, pp. 281-304.

Посилання на конференцію:

[4] Chu-Hsing Lin, Jung-Chun Liu and Chun-Wei Liao, "Energy analysis of multimedia video decoding on mobile handheld devices," in *Proceedings of the International Conference on Multimedia and Ubiquitous Engineering,* Seoul, Korea, April 26-28, 2007, pp. 120-125.

Інтернет видання:

[5] M. Shell, *The IEEEtran.cls Package*, 2007, [Online]. Available: http://www.ctan.org/tex-archive/macros/latex/contrib/IEEEtran/

Будь ласка, поставте тут відскановану фотографію автора 3 х 4 см.

***Повне ім'я автора**, Коротке резюме автора (10-20 рядків), в тому числі: освіта, місце роботи і область наукових інтересів, відскановане фото.*