by Anhalt University of Applied Sciences

PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON APPLIED INNOVATIONS IN IT

VOLUME 12

Issue 2

EDITION

Hochshule Anhalt

Proceedings of the 12th International Conference on Applied Innovations in IT

Volume 12 | Issue 2

Koethen, Germany
30 November 2024

This journal focuses on applied research in the fields of communication and information technologies, data processing, automation, power engineering, and information technology in economics. The research results are of interest for researchers and development engineers who deal with theoretical base and the application of the knowledge in the respective areas. The journal aims to contribute to scientific progress by sharing innovative research findings and fostering knowledge exchange in these critical areas.

ISBN: 978-3-96057-179-7 (Online)
ISSN: 2199-8876

All papers published in this jornal are full-open access:
immediately freely available to read and download.

Publisher: Anhalt University of Applied Sciences
Bernburg / Koethen / Dessau
http://www.hs-anhalt.de

Postfach 1458
D-06354 Koethen, Germany
Phone: +49 3496 67 2327
Email: eduard.siemens@hs-anhalt.de
Web: http://icaiit.org

Content

Section 1 Communication and Data Transport Technologies

Roman Tsarov, Serhii Siden, Lesya Nikityuk, Kateryna Shulakova and Liliia Bodnar
Assessment of the Impact of Intra-System Interference on the Throughput of a 5G-Based Telemedicine Network……. 1

Hassan Hadi Saleh, Israa Mishkal and Abd Ali Hussein
Enabling Smart Mobility with Connected and Intelligent Vehicles: The E-VANET Framework…..…… 9

Liliia Bodnar, Mykola Bodnar, Kateryna Shulakova, Oksana Vasylenko, Eduard Siemens, Roman Tsarev, Olha Yavorska and Olena Tyurikova
Advanced Techniques for IaC: Enhancing Automation and Optimization in Cloud-Based Infrastructure Management……..…….……..……. 19

Dmytro Sapozhnyk
Quantum Approximate Optimization Algorithm for the Max-Cut Problem: JavaScript Programming Language Implementation…… 27

Victor Tikhonov, Eduard Siemens, Yevhen Vasiliu, Valery Sitnikov, Abdullah Taher, Olena Tykhonova, Kateryna Shulakova and Serhii Tikhonov
Context-Defined Model of Open Systems Interaction for IoT Cybersecurity Issues Study…….….. 35

Tetiana Hromko, Oksana Vasylenko, Nataliya Chepelyuk and Yevheniia Hromko
Digital Transformation of Linguistic Pedagogy…….. 45

Viatcheslav Shuvalov and Irina Kvitkova
Comparative Analysis of the Unavailability Factors for Two Types of Optical Cable Section Repair under Conditions of Gradual and Sudden Failures…….……. 57

Nadezhda Abrosimova, Anna Belezekova, Evgeny Isenev and Maxim Shushnov
On the Need to Clarify Methods for Assessing the Quality of Sound Paths Through the Use of Psychoacoustic Criteria…… 65

P*aul Arhin and George Aggrey*
Enhancement of Elgamal Cryptosystem; A Review…… 71

Section 2 Data Analysis and Processing

Olha Kovalchuk, Nataliia Chudyk, Tetiana Drakokhrust, Valerii Kaniuka, Roman Ivanytskyy and Larysa Khokhlova
Classification Tree Model for Determining Society Unsafety Factors Convicted…… 77

Marianna Kovtoniuk, Olena Soia, Olena Kosovets and Ivanna Leonova
Development of a Textbook in the Modern Educational and Scientific Mathematical and Informatics Environment of Higher Education Students…… 83

Hour Alburaimi and In-Ju Kim
Optimizing UAE Food Supply Chain Management: Leveraging Fuzzy AHP for Strategic Selection of Optimal Blockchain Platforms…… 91

Igor Melnyk, Mykhailo Skrypka, Alina Pochynok and Olga Demyanchenko
Novel Computer Software for Interpolation and Approximation of Ravine and Stiff Digital Dependencies Using Root-Polynomoal and Root-Fractional-Rational Functions……. 99

Section 3 Control and Automation

Aleksy Kwilinski and Nataliia Trushkina
Impact of Cyber Risks and Threats on the Critical Infrastructure Development: Visualization of Scientific Research……..……. 107

Katerina Mitkovska-Trendova, Robert Minovski, Verica Bakeva, Simeon Trendov and Dimitar Bogatinov
Towards Automated Quality Control in Industrial Systems: Developing Markov Decision Process Model for Optimized Decision-Making…….……..…… 121

Section 4 Power Engineering

Boysori Yuldoshov, Mamasobir Tursunov, Gayrat Raimov, Shakhvoz Karshiev and Sirojiddin Toshpulatov
Mathematical Modeling of Heat Parameters of Photothermal Device……..…….. 131

Muhammadjon Tursunov, Khabibullo Sabirov, Ramazon Alikulov, Kholov Uygun, Shohimardonov Jamoliddin and Eshmatov Mansur
Testing the Energy Efficiency of a New Type of Photothermal Device in Dry Climate Conditions…….. 139

Muhammadjon Tursunov, Khabibullo Sabirov, Tohir Axtamov, Umirbek Abdiyev, Azamat Narbayev, Sirojiddin Toshpulatov and Tulkin Buzrukov
Annual Analysis of On-Grid Station with an Installed Capacity of 200.75kW: The Case of the City of Termez……... 145

Tulkin Gayibov and Boburjon Murodov
Optimal Planning of Short-Term Modes of Power Systems Containing Energy Storage Devices……. 151

Oleg Ivanov
Study of the Energy Efficiency of a Thermal Electrical Generator with a Hydraulic Heat Supply System…… 159

Boysori Yuldoshov, Maxamadi Chariyev, Azamat Narbayev, Jasur Khaliyarov, Sirojiddin Toshpulatov, Sojida Raupova and Tulkin Buzrukov
Annual Monitoring Analysis of Four On-Grid Stations in Termez…… 167

David Akpuluma, Wolf-Gerrit Früh, Neda Firoz, James Abam, Mohammed Umar Bello, Comfort Williams and Ambrose Onne Okpu
Comparative Analysis of LSTM-Based PV Power Forecasting Models with Climate-Adaptive Feature Selection in Abuja, Nigeria……. 173

Halina Falfushynska, Cornelia Scott and Markus Holz
Environmental Sustainability Assessment of Green Hydrogen Production from Seawater Using Life Cycle Assessment…….. 185

Section 5 Information Technology in Economics

Olha Kovalchuk, Kateryna Berezka, Oksana Tulai, Bohdan Ferens, Ludmila Babala and Oleksandr Zapolskyi
Modeling the Determinants of Global Economic Uncertainty…… 191

Iryna Polishchuk, Yuriy Dovhan, Iryna Kramar, Larysa Dovhan and Oksana Yakushevska
Marketing Logistics and its Digitalization……... 199

Olha Polous, Inna Mykhalchenko, Hanna Radchenko, Olena Lysenko and Kateryna Pysmenna
Formation of Product Promotion Strategies of Companies Producing Computer Games…….. 207

Olena Ivanova, Oleg Ivanov, Valentyna Ivanova and Alona Kashpruk
Development of an Information Management System for it Enterprise's Intellectual Potential…… 215

Oleksandra Tsyra, Yuliia Tochylina and Kateryna Tanashchuk
Study of Digital Skills Using ICT in the Development of the Digital Economy…… 223

Viktoriia Khaustova, Mykola Kyzym, Olena Reshetnyak, Mykyta Khaustov and Nataliia Trushkina
Methodical Approach to the Startup Ecosystem Formation: Foreign Experience and Opportunities for Ukraine…… 231

Iryna Stankevych, Inna Yatskevych, Hanna Sakun, Nataliya Shyriaieva, Oksana Vasylenko and Eduard Siemens
Assessment of the Impact of Social Responsibility of Integrated Business Forms on Regional Development……... 245

SECTION 1

Communication and Data Transport Technologies

Proceedings of the 12th International Conference on Applied Innovations in IT (ICAIIT)

NOVEMBER 2024

Assessment of the Impact of Intra-System Interference on the Throughput of a 5G-Based Telemedicine Network

Roman Tsarov[1], Serhii Siden[1], Lesya Nikityuk[1], Kateryna Shulakova[1,2] and Liliia Bodnar[3]
[1]*State University of Intelligent Technologies and Telecommunications, Kuznechna Str. 1, 65023 Odesa, Ukraine*
[2]*Anhalt University of Applied Sciences, Bernburger Str. 57, 06366 Köthen, Germany*
[2]*South Ukrainian National Pedagogical University, Staroportofrankyvska Str. 26, 65020 Odesa, Ukraine*
rcarev@gmail.com, ssiden@suitt.edu.ua, lesyanikityuk579@gmail.com, katejojo29@gmail.com, bodnarl79@pdpu.edu.ua

Keywords: Telemedicine, Telemedicine Network, 5G; Propagation models, SINR.

Abstract: The article is devoted to the use of 5G networks in the field of telemedicine, namely for the deployment of telemonitoring clusters with a high density of telemedicine sensors and detectors. As part of the study, the possibility of fifth-generation technology for organizing a telemonitoring system was assessed, provided that a guaranteed data transfer rate is ensured, considering the simultaneous connection of many devices and sensors. The assessment was carried out considering that any device within the cluster is a potential source of additional electromagnetic interference and affects other devices. The telemonitoring system based on fifth-generation networks was modeled for different initial conditions (distance to the base station, interference power, etc.) under the influence of intra-system interference. The distribution of intra-system interference sources and their distance to the object of study is realized as a random process based on the Monte Carlo method. The results obtained show that the actual throughput of a telemedicine network based on 5G technology largely depends on the number of sources of intra-system interference, which in turn affects the capacity of the telemonitoring cluster (number of endpoints) and the actual density of endpoints is several times less than stated by the technology developers.

1 INTRODUCTION

Today, thanks to information technology and network development, many different areas are undergoing a paradigm shift, and the medical field is no exception. The integration of information technology into the healthcare industry has contributed to the development of telemedicine. Telemedicine is a healthcare area that involves information technology to increase the availability of medical services and facilitate communication between patients and doctors in cases where distance is a critical factor, i.e., personal communication is not possible at the current time [1, 2]. The development of telemedicine has contributed to creating a wide range of services and applications to facilitate the exchange of data in various forms and volumes within the provision of medical services. In accordance with the Sustainable Development Goals [3] and national healthcare development programs, the priority areas of healthcare development are both increasing the availability of medical services and timely monitoring of the health of citizens. Within the framework of telemedicine, this direction is realized through such a service as telemonitoring.

Modern telemonitoring is developing through the convergence of information technology and the concept of the Internet of Things (IoT)/Internet of Medical Things (IoMT). The importance of telemedicine in general, and telemonitoring in particular, became apparent during the outbreak of the COVID-19 pandemic. The pandemic has demonstrated the critical importance of telemonitoring systems, which made it possible to monitor patients' vital signs in real-time remotely and, at the same time, ensure doctors' health. Today, the demand for telemonitoring systems continues to grow. This is because the likelihood of a recurrence of the pandemic remains very high, as well as the growing number of military conflicts and wars. This situation requires creating large-scale, functional, and efficient telemedicine monitoring systems to monitor

critical vital signs and inform doctors about them in real-time. All of this suggests, however, that telemonitoring systems will become more complex in their architecture and increase in scale, and while today's telemonitoring systems contain hundreds of sensors, shortly telemonitoring systems will include thousands of such sensors and sensors. To create such systems, it is necessary to use new information and communication technologies capable of providing the required density of devices within the telemonitoring system, the required bandwidth, etc. One of these technologies is the fifth-generation technology and network 5G.

This paper aims to analyze the possibilities of implementing 5G networks in developing and improving telemonitoring systems, taking into account possible intra-system interference.

2 TELEMONITORING SYSTEMS

Telemonitoring is one of the essential services of telemedicine [2], which makes it possible to control vital signs and reduce the risks of developing and complicating chronic diseases, such as heart disease and diabetes, by continuously monitoring patients using data analytics and decision support systems. A modern telemonitoring system (Fig. 1) is an information and communication system for collecting and transmitting indicators.

From an architectural point of view, the monitoring system consists of three components [4,5]:

1) The level of actuators. It is represented by sensors and transducers that collect various physiological data of the patient in real time.
2) Data transmission level. It is represented by gateways and an information network that transmits the collected data from sensors to storage and analysis systems.
3) Analytical processing level. This level is represented by various local and/or cloud storage, analytics, and decision support systems that provide storage, processing, and analysis of data received from the system's sensors and detectors.

The system is based on actuators - sensors and detectors. A sensor is a device that directly measures a particular physiological parameter (e.g., temperature, pressure, etc.) and converts it into a signal suitable for transmission over a network. A sensor cannot display and analyze data [6]. A detector is a more complex device containing several sensors and additional components for displaying and analyzing data. Sensors and detectors are divided into electrical, optical, mechanical, and chemical according to the principle of operation. According to how they are used, they are divided into wearable and implantable, and according to how the collected data is transmitted, sensors and detectors are divided into wired and wireless. It is obvious that, in most cases, flexible and scalable telemonitoring systems can be created based on wearable wireless sensors and detectors. Such devices are connected to a data transmission network using appropriate radio technology [6].

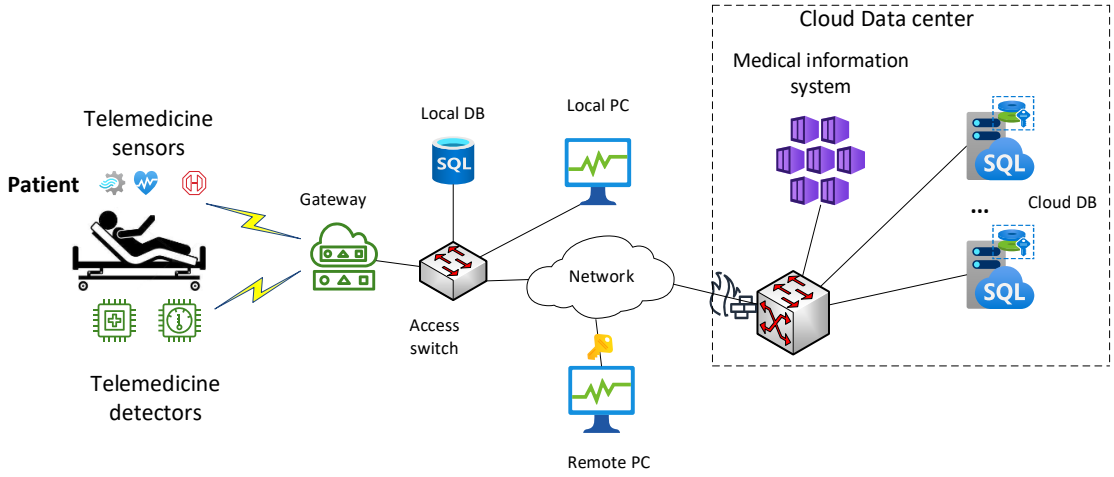


Figure 1: Basic architecture of the telemonitoring system.

To create a reliable telemonitoring system, the data transmission layer must be based on radio technologies that provide high-speed data transmission, low latency, and high density of devices per unit area. Existing technologies, such as 4G, are not able to fully meet the current and future needs of the telemonitoring system (telemonitoring systems require a high density of devices per unit area), so it is advisable to use 5G technologies, which have several advantages over 4G technology, to create a data transmission layer [7]. An alternative to 5G technology is to create a system based on WI-FI 7 technology or on ZigBee with a gateway to a wired Ethernet network. However, such solutions have certain limitations. 5G technology provides greater mobility, better bandwidth, lower latency, and greater scalability. Scalability and mobility play an important role in cases where it is necessary to serve thousands of devices in a small area and provide connection to the system for mobile medical teams [11]. WI-FI 7 or ZigBee should be considered as a solution for creating a local telemonitoring system only inside the building, while a solution based on 5G technology is universal and allows you to create large-scale telemonitoring systems.

The primary data source within the telemonitoring service is sensors and sensors that collect medical data and transmit it for further processing over the network. Channels with a bandwidth of 0.1 – 1 Mbps are sufficient for telemonitoring data transmission [8]. Providing such bandwidth is not a problem for existing technologies, but a high density of sensors and sensors per unit area characterizes modern telemonitoring systems.

It is known that the maximum network capacity for a MIMO channel is determined by the following (1):

$$C = M \cdot \left(\frac{BW}{n} \right) \log_2 \left(1 + \frac{P_s}{N} \right), \quad (1)$$

where M is the number of MIMO levels, BW is the channel bandwidth, n is the number of subscribers, P_s and N is the power of the useful signal and interference, respectively.

The signal strength at the receiver input depends on many factors: operating frequency, terrain, weather conditions, etc. Various radio wave propagation models are used to estimate the signal strength at the receiver input. The telemonitoring cluster can operate in different areas, and it is recommended to use radio wave propagation models for various types of terrain to model the loss level (Rural, Urban Macro, Urban Micro) [9, 10] using (2), (3) and (4). These models are part of the 3GPP recommendation and are widely used in real-world applications as basic models.

$$PL_{Rural} = 20\lg(40\pi d_{3d}\ f/3) + \min(0.03h^{1.72}, 10)\lg(d_{3d}) - \\ -\min(0.044h^{1.72}, 14.77) + 0.002\lg(h)d_{3d}; \quad (2)$$

$$PL_{Urban\,Macro} = 28 + 22\log(d_{3D}) + 20\log(f), \quad (3)$$

$$PL_{Urban\,Micro} = 32,4 + 21\log(d_{3D}) + 20\log(f). \quad (4)$$

The following (5) determines the minimum noise level of the receiver:

$$N = -174 + 10\log(BW) + N_f, \quad (5)$$

N_f is receiver noise level equal to 0.3 dB.

It is possible to determine the maximum range of the telemonitoring service as a function of the 5G channel bandwidth and signal strength under different radio wave propagation conditions and to assess how this indicator is affected by such a factor as the number of connected devices (sensors) using (1)-(5).

Table 2 shows how the radius of telemonitoring service provision, the density of connected devices (sensors), and the technical parameters of the 5G network are interrelated [8].

The table shows that in the base case, the service radius, the number of devices, and the MIMO configuration (M) are directly proportional to all three components. 5G technology allows scalable telemonitoring systems with different configurations, but the service coverage radius is relatively small. For example, you can create a system for 1000 devices, but the maximum range is only 28 meters, with the maximum MIMO configuration (M = 8).

Table 2: The Dependence of the distance of the remote monitoring service on the number of connected devices and 5G network parameters.

Environment	Number of devices					
	100		500		1000	
	$M = 2$	$M = 8$	$M = 2$	$M = 8$	$M = 2$	$M = 8$
Rural	60	115	27	54	17,5	38
Urban Macro	53	110	22	47	13,5	32
Urban Micro	45	88	19	40	12	28

3 ASSESSMENT OF THE IMPACT OF INTRA-SYSTEM INTERFERENCE ON TELEMEDICINE NETWORK THROUGHPUT

The studies in [8] were carried out for a system with one base station without considering intra-system interference in the cluster. In real conditions, the telemonitoring cluster, like any other cluster of the IoT concept, contains many different radio devices (base stations, various sensors and sensors, gateways, smartphones), which are a source of intra-system interference. In conditions of low density of end devices within a cluster, intra-system interference can be ignored. Clusters created based on 5G technology have hundreds of times higher density of end devices. Accordingly, intra-system interference will significantly impact network capacity, which in turn may lead to a change in the radius of service provision and a reduction in the number of devices in the cluster. In such circumstances, the model described by (1)-(5) needs to be adapted to function in the face of intra-system interference.

The communication channel capacity is determined by the Shannon-Hartley theorem:

$$\mathrm{C} = \Delta f \ log_2 \ (1 + SINR), \qquad (6)$$

where Δf is is the bandwidth, SINR is the signal-to-noise ratio.

In order to determine the SINR value, it is necessary to determine the signal power at the receiver input P_s, the noise spectral density P_n and the level of total interference from other devices $P_{\Sigma i}$. To find the useful signal power, use the formula:

$$P_s(d) = P_t + G_t + G_r - PL(d), \qquad (7)$$

where P_t is the power of the transmitter (base station), G_t and G_r are the gain of the transmitting and receiving antennas, respectively, $PL(d)$ is the path loss, which depends on the operating frequency, distance, and propagation medium.

The power level of other devices was estimated using a similar (8):

$$P_i = P_d + G_d + G_r - PL(d_i), \qquad (8)$$

where P_d is the power of the interfering transmitter, G_d and G_r are the gain of the interfering antenna and the device under study, $PL(d_i)$ is the path loss, d_i is the vector distance between the devices.

After determining the power of each individual interference, the total interference level is calculated:

$$P_{\Sigma i} = \sum_{j=1}^{N} P_i, \qquad (9)$$

The minimum noise level of the receiver P_n is determined as follows:

$$P_n = -174 + 10 \log(\Delta f) + N_f, \qquad (10)$$

N_f is the receiver noise level of 0.3 dB.

The Urban Macro model was chosen as the radio wave propagation medium. According to the ETSI recommendation, the interference level can be determined by the following (11) [9].

$$PL(d) = 28 + 22 \log(d) + 20 \log(f), \qquad (11)$$

where f is the working frequency, GHz.

The analysis of the formalized model allows us to conclude that the number of devices and the cluster radius depend on two parameters:

- the number of additional sources of interference acting on the object of study;
- the distance between the interference source and the object of study.

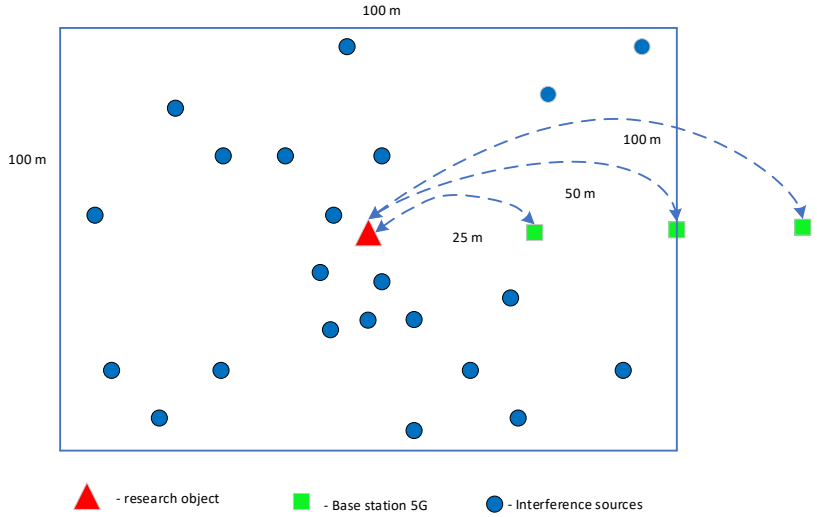
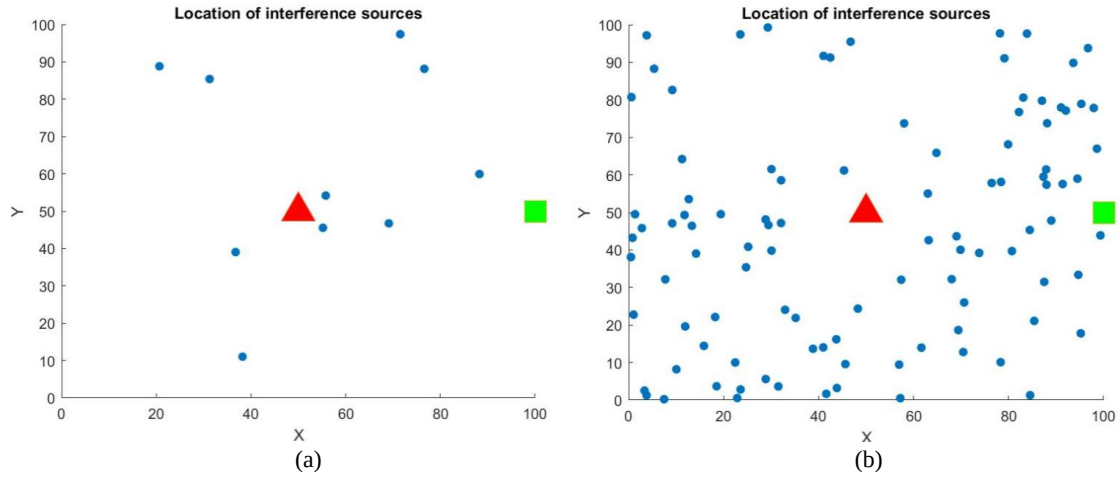


Figure 2: Geometric model of a telemedicine cluster.



(a) (b)

Figure 3: Location of devices in the selected area with n = 10 (a) and 100 (b).

To assess the impact of such interference sources on such important indicators as the density of devices and the radius of the telemedicine cluster, we will conduct a simulation in Matlab based on a formalized model (described by formulas 6-11).

A computer model for estimating the maximum throughput was created to determine the maximum number of simultaneously connected devices to the 5G telemetry network, which ensures the required data transmission rate of 1 Mbps. The modeling was performed for the following initial data: the area under consideration S is 100 m2, the number of devices N is from 2 to 1000, and the distance to the base station is 25, 50, and 100 m. The general modeling scheme is shown in Figure 2.

Since the territorial distribution of subscribers is random, the Monte Carlo method was used to place telemetry devices [10]. The network space is modeled as a two-dimensional plane in which interference sources (other devices) are randomly placed. For each source, the location coordinates and power of the generated interference are determined, taking into account the vector distance between the devices, the propagation medium, and equipment parameters.

This allows us to calculate the total interference level as the sum of the power from all interference sources and to estimate the throughput for each scenario using the (1).

To obtain more accurate estimates, the described experiments were repeated 10 times. After that, the throughput is averaged over all experiments, and its dependence on the number of active interferers is estimated.

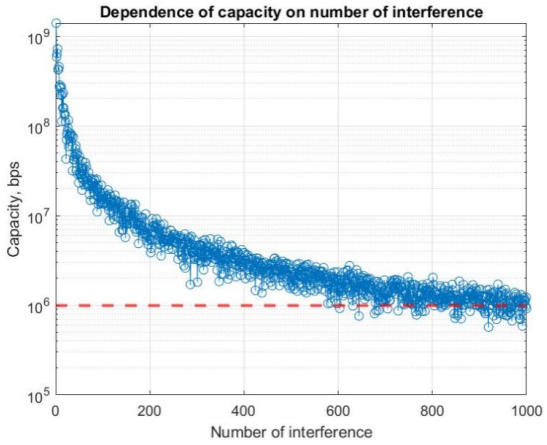


Figure 4: Dependence of bandwidth on the number of devices at 25 meters to the BS.

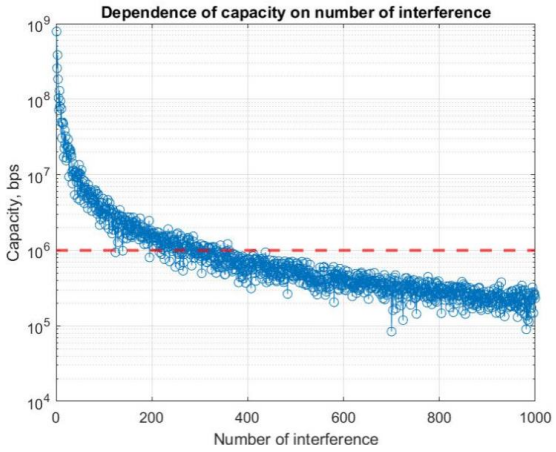


Figure 5: Dependence of bandwidth on the number of devices at 50 meters to the BS.

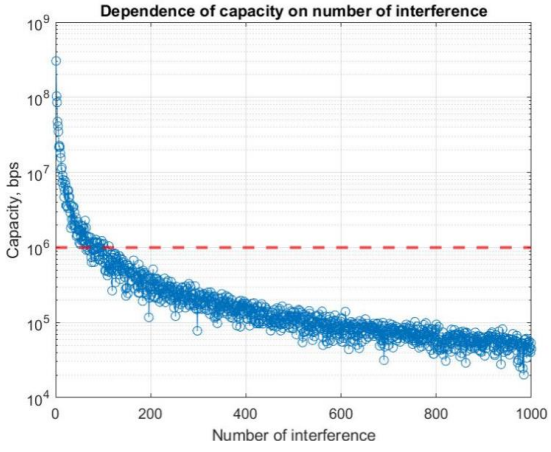


Figure 6: Dependence of bandwidth on the number of devices at 100 meters to the BS.

The described model for the example of 10 and 100 devices is shown in Figs. 3a and 3b.

Figures 4-6 show the dependence of the maximum bandwidth of the device on the number of other devices at a fixed distance to the base station.

The presented dependencies allow you to estimate the maximum number of simultaneously functioning devices at different distances to the base station. For example, at a distance of 25 meters, you can provide a guaranteed data rate for 640 devices, while at a distance of 100 meters to the BS - for only 80 devices.

4 CONCLUSIONS

The high density of devices per unit area declared in fifth-generation networks creates a potential problem - it forms a set of sources of intra-system interference, the power of which is almost equal to the network capacity. This poses a significant problem for systems and networks with a high density of devices deployed on the basis of 5G, such as telemedicine telemonitoring networks. To meet the requirements for the density of connected devices specified in the requirements for telemonitoring networks, it is necessary to take into account the mutual influence of all telemonitoring devices operating simultaneously, since the mutual influence leads to a significant reduction in network quality indicators, such as network speed and reliability.

The proposed method of assessing mutual influence based on the Monte Carlo method allows us to estimate the level of influence of intra-system interference on network efficiency and predict optimal conditions for telemedicine systems.

Reducing the mutual influence of devices is possible due to parametric optimization of system characteristics, such as frequency range, transmitter power, configuration of MIMO and beamforming technologies.

The results of this study can be the basis for further design and optimization of telemedicine networks based on 5G. In particular, the proposed method can be implemented to form a strategy for effective management of radio frequency resources in conditions of high density of devices and scenarios requiring high reliability of data transmission. This will ensure a high level of quality of service (QoS) and minimize the risk of loss of communication during data transmission, which is critical for the

successful functioning of telemedicine services based on 5G technology.

ACKNOWLEDGMENTS

We acknowledge support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Open Access Publishing Fund of Anhalt University of Applied Sciences.

REFERENCES

[1] R. Tsarov, I. Tymchenko, V. Kumysh, K. Shulakova, and L. Bodnar, "Extended Classification Model of Telemedicine Station," in Proc. Int. Conf. Appl. Innov. IT, vol. 11, no. 1, pp. 37-42, 2023. doi: 10.25673/101908.

[2] "Recommendations on the application of modern technical solutions in the design of e-health systems, including telemedicine networks," [Online]. Available: https://www.itu.int/en/ITU-D/Regional-Presence/CIS/Documents/RI-WTDC17/ONAT_RI2%20Recommendations_Rev2.pdf.

[3] "Sustainable Development Goals," [Online]. Available: https://sdgs.un.org/goals.

[4] X. Hu, L. Wang, P. Gao, X. Dong, Z. Ji, F. You, et al., "Study on Disease Screening and Monitoring System Based on Wireless Communication and IOT," in Proc. 2012 Spring Congr. Eng. Technol., Xi'an, China, 2012, pp. 1-6. doi: 10.1109/SCET.2012.6341961.

[5] J. Mohammed, C. H. Lung, A. Ocneanu, A. Thakral, C. Jones, and A. Adler, "Internet of Things: Remote Patient Monitoring Using Web Services and Cloud Computing," in Proc. 2014 IEEE Int. Conf. Internet Things (iThings), Taipei, Taiwan, 2014, pp. 256-263. doi: 10.1109/iThings.2014.45.

[6] A. Sabban, "Innovation and Review in Health Monitoring Systems and Wearable Medical Sensors," Ann. Clin. Med. Case Rep., vol. 11, no. 12, pp. 1-15, 2023.

[7] Y. Hao, "Investigation and Technological Comparison of 4G and 5G Networks," J. Comput. Commun., vol. 9, pp. 36-43, 2021. doi: 10.4236/jcc.2021.91004.

[8] M. Saad, S. Serhii, R. Tsaryov, and L. Nikityuk, "Assessment of the Possibility of Using 5G to Build Telemedicine Networks in Various Environments," in Proc. IDAACS 2023, pp. 1125-1129. doi: 10.1109/IDAACS58523.2023. 10348929.

[9] ETSI TR 138 901, "Study on channel model for frequencies from 0.5 to 100 GHz," 3GPP TR 38.901 version 14.0.0 Release 14, 2017.

[10] W. Krauth, Introduction to Monte Carlo algorithms, 2006, 41 p. [Online]. Available: https://cel.archives-ouvertes.fr/cel-00092936/document.

[11] Alenoghena, C. O., Ohize, H. O., Adejo, A. O., Onumanyi, A. J., Ohihoin, E. E., Balarabe, A. I., Okoh, S. A., Kolo, E., & Alenoghena, B. (2023). Telemedicine: A Survey of Telecommunication Technologies, Developments, and Challenges. Journal of Sensor and Actuator Networks, 12(2), 20, [Online]. Available:https://doi.org/10.3390/jsan12020020.

Enabling Smart Mobility with Connected and Intelligent Vehicles: The E-VANET Framework

Hassan Hadi Saleh [1], Israa Mishkal [2] and Abd Ali Hussein[3]

[1] *Department of Computer Science, College of Education for Pure Science, University of Diyala, MJJ2+R9G, 32001 Baqubah, Iraq*
[2] *Department of Computer Science, Sciences College, University of Diyala, MJJ2+R9G, 32001 Baqubah, Iraq*
[3] *Computer Department, College of Engineering, University of Diyala, MJJ2+R9G, 32001 Baqubah, Iraq*
hassan.hadi@uodiyala.edu.iq, israaadnan@uodiyala.com, abdalihussein@uodiyala.edu.iq

Keywords: Vehicular Ad Hoc Networks (VANETs), Intelligent Transportation Systems (ITSs), Intelligent Transportation Systems (ITSs), Vehicle-to-Vehicle (V2V), Vehicle -to-Infrastructure (V2I), Smart Mobility.

Abstract: Vehicular Ad Hoc Networks (VANETs) have become a trending technology for enabling intelligent and connected vehicles in the smart mobility era. This paper presents the ideas, architectures, applications, and challenges of VANETs, as well as their influence on Intelligent Transportation Systems (ITSs), in this comprehensive study. It explores the characteristics and architecture of VANETs, including communication protocols, mobility models, and network architectures. Additionally, it analyses various communication protocols and standards, including IEEE 802.11p/WAVE and Dedicated Short-Range Communications (DSRC). Using the fundamental traffic flow equation ($Q=k\cdot v$), urban deployment scenarios demonstrated a 20% improvement in traffic efficiency, reducing average travel times by approximately 10 minutes during peak hours. In highway scenarios, VANET-enabled collision avoidance systems, employing the Time-to-Collision (TTC) metric ($TTC= d\ Vr$), reduced potential crash scenarios by 30%, preventing up to 15 accidents per 1,000 vehicles. This paper discusses the role of VANETs in wise traffic management, eco-friendly transportation systems, advanced driver assistance structures (ADAS), and autonomous vehicles. Finally, it emphasizes VANETs' potential in enabling sustainable, efficient, and secure mobility, providing valuable insights for researchers, practitioners, and policymakers to shape the future of transportation systems.

1 INTRODUCTION

VANETs are a key to making smart travel and smart roads work. These networks enable vehicles communicate with each other and with objects on the roadside, fostering a connected world for safety and environmentally friendly travel [1]. VANETs matter because they let vehicles and roads talk to each other in real time. By sharing info about traffic, dangers on the road, and the weather, VANETs help drivers and traffic systems make smart choices. This leads to smoother drives, less jamming up, and quicker trips [2]. Through vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) chats, they facilitate the effective operation of safety tools such as group crash dodge, lane-changing assistance, and crossroad vehicle systems. By giving vehicles a full look at what's around them and warning drivers of possible risks, VANETs aid in stopping crashes and make roads safer overall [3].

VANETs help make green travel systems. By sharing live traffic news, VANETs let smart traffic systems improve how vehicles' movement, reduce jams, and use less fuel. This cuts the polluted air and supports kinder travel methods [4]. VANETs are a key to making smart travel and clever roads work. They help talk in real-time, making traffic move better, making roads safer, assisting self-driving vehicles to grow, and keeping our travels green. They utilize simple communication to establish a connection, enabling both vehicles and roads to immediately understand and respond to change. This makes getting around smoother, keeps us safe while we move, helps vehicles drive on their own, and is kind to our planet [5]. The story of VANETs starts in the late '90s when experts began studying how vehicles can talk to each other to make roads safer and traffic flow better. Since then, the VANET study has experienced significant growth and milestones that have shaped its

trajectory. One early big step in the VANET study was making the DSRC standard, also known as IEEE 802.11p/WAVE, in 2003. This standard sets up a way for vehicles to communicate in a way that fits the road setting. It provides a base for VANET research and building [6]. In the following years, numerous larger-scale tests and studies were connected on talking vehicles to assess their effectiveness. Key efforts were the VII project in the US, the C2C-CC group in Europe, and the ASV plan in Japan. These studies investigated how vehicles communicating with each other and roads could enhance driving safety, and improve traffic flow through cooperative efforts [7]. As the researchers on VANETs progressed, they developed various test tools and methods to mimic vehicle movement, assessing the effectiveness and functionality of vehicle networks. These presented approaches, such as the Random Waypoint, alter the provided text and adhere to certain guidelines: utilize concise and straightforward language, replace complex words with simple ones, and vary the length of sentences. Use the most-used English words when you can Keep the number of words the same [8]: Manhattan Grid and SUMO (Simulation of Urban Mobility) [9] helped experts grasp how vehicles move, create smart ways to talk, and guide paths for VANETs.In the study, we discuss VANET communication, movement, network setup, and security. We also present the potential applications of VANETs, such as in smart traffic city control. Additionally, this study investigates the big hurdles with VANETs, such as modelling movements, finding the best paths, managing the airwaves, and making sure the service is good [10]. Also it introduces the world of VANETs, focusing on their role in making smart movement and ITS possible. It digs into the history and major moments of VANET research and investigates what's new in this technology.

2 CHARACTERISTICS AND ARCHITECTURE OF VANETS

VANETs are a kind of on-the-go network that lets vehicles and road setups talk to each other without a central point. Make the text shorter, use simple words, and mix up how long the sentences are. Use the simplest and most familiar words you can, but keep the number of words the same (see figure 1) [11].
This supports a range of needs, from keeping traffic moving well to helping the environment and making driving safer. When choosing these methods, it’s crucial to consider the network’s construction, its primary objectives, and its intended use [19]. VANETs are special because they're different from other network types.

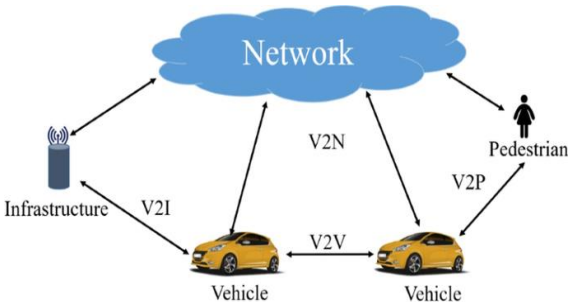


Figure 1: Vehicles communication types [11].

As shown in Table 1, it lists all these differences. These unique traits enable VANETs to effectively serve smart road applications such as enhanced traffic control, crash prevention, green travel, and driver assistance. By using the changing ways of vehicle networks and making communication easy, VANETs can make roads safer, cut down on jams, and make travel better overall [12]. The design of VANETs has three key parts: the onboard units (OBUs) in vehicles, the roadside units (RSUs) set up along roads, and the network that lets them talk to each other. The methods they use to share information ensure smooth data movement and support various applications [18].
Table 2 illustrates the primary communication methods utilized in VANETs. The ones you pick depend on things like the network setup, what you need it for, and how you plan to use it. They vary in style, but they all strive ensure seamless information sharing between vehicles and other roadside equipment. This supports a range of needs from keeping traffic moving well to helping the environment and making driving safer. When choosing these approaches, it’s crucial to consider the network’s construction, its primary objectives, and its intended use [19].

3 COMMUNICATION PROTOCOLS AND STANDARDS

DSRC lets vehicles talk to each other and share important information, like how fast they are going, where they are, how they are moving, and other things that can help them avoid crashes, control traffic, and work together. Different areas have changed and

Table 1: Distinctive types of communication networks.

Ref.	Type	Feature
[12]	Dynamic Network Topology	VANETs have a very fast-changing network layout because vehicles move quickly. This network shape shifts a lot as vehicles come in and go out of reach, making talking and path-finding hard.
[13]	V2V Communication	VANETs let vehicle s talk to each other, sharing info on how fast they're going, where they are, how quickly they're speeding up, and more. This vehicle -to-vehicle chat is the core of many team apps designed to make roads safer and traffic flow better.
[14]	V2I Communication	Besides vehicle -to-vehicle talk, VANETs also let vehicles chat with things on the side of the road, like traffic lights, signs, and watch systems. Vehicle -to-infrastructure talk lets vehicles and things swap up-to-the-minute road news, warnings, and commands.
[15]	Wireless Communication Technologies	VANETs use short-range radio, like IEEE 802.11p/WAVE DSRC, LTE-V2X, and 5G NR-V2X. These techs give the speed and quick chat needed for live use in vehicles.
[16]	Time-Sensitive Communication	VANETs often need quick chat because late messages can cause big problems. Apps like crash alert systems and crossroad control need fast and sure info sharing for good choices.
[17]	Security and Privacy Concerns	VANETs have their own tough security and privacy issues. Making sure messages are safe, guarding against bad attacks, and keeping private data secret are key parts in setting up and running VANETs.

adjusted DSRC to meet their needs [30]. Table 3 shows some significant differences. Different methods of DSRC aim to facilitate vehicle communication and enhance the efficiency of smart transportation. The DSRC has developed new approaches for sending and receiving messages using cells, enabling vehicles to communicate and exchange crucial information such as their speed, location, and movement patterns, which can aid in preventing crashes, managing traffic, and collaborating. The DSRC has undergone modifications and adjustments to meet the requirements of various regions [30].

Table 3 illustrates a few key differences. Different methods of DSRC aim to facilitate vehicles communication and enhance the efficiency of smart transportation. Some developed new cell-based message-sending and receiving methods to complement the traditional radio-wave methods. Two ways to use cells to make vehicles talk to each other are LTE-V and 5G NR-V2X [29].

They are shown in Table 4. The designs of inter-vehicle Communication (IVC) protocols enable direct communication between motors in vehicular VANETs, eliminating the need for infrastructure elements. These protocols enable the green exchange of records, in addition to facilitating protection warnings, site visitor updates, and cooperative manoeuvres among neighbouring vehicles. There have been several proposed IVC protocols, each with unique characteristics and relevance to VANETs [32]. Safety-critical packages regularly rely on dedicated IVC protocols with low latency and reliable conversation, while different applications may leverage more flexible protocols to guide diverse. The selection of a suitable IVC protocol relies upon the unique wishes of the VANET utility and the available verbal exchange technologies within the community [33].

Here are some exceptional IVC protocols for consideration. The applicability of IVC protocols in VANETs depends on various factors, which include the intended use case, communication variety requirements, community scalability, and interoperability. VANETs specifically inspired the design of IEEE 802.11p/WAVE, a communication standard. IEEE 802.11p/WAVE operates at the 5.9 GHz frequency and is a protocol that forms the basis of wireless in-vehicle and vehicle-to-infrastructure communications. This enables Remote Direct V2V and V2I communication [26]. This standard aims to facilitate the transmission of safety-related information, traffic updates between vehicles, and other relevant data. The design provides high-speed, low-latency communication channels essential for real-time applications in VANETs. In particular, the standard describes the physical signal transmission and medium access protocols grounded on it. The standard provides definitions for PHY, medium access control, and modulation, among others things. IEEE 802.11p/WAVE provides the basis for the functioning of VANETs. It supports cooperative system possibilities, ensures an increase in road safety, and contributes to ETS [27]. DSRC is a way for cars and other vehicles to talk to each other using radio waves. Vehicles that operate in groups, such as buses and trains, can utilize it. It uses a certain type of radio waves and has its own special way of talking to other cars and things on the road.

Table 2: Types of Communication protocols.

Ref.	Protocol	Function
[20]	IEEE 802.11p/WAVE (DSRC)	This system, called DSRC, is made for VANETs. It works in the 5.9 GHz band and lets vehicle s and stuff talk fast and with no delay. IEEE 802.11p/WAVE sets the base for vehicle -to-vehicle and vehicle -to-thing chats in VANETs.
[21]	LTE-V2X and 5G NR-V2X	These setups use cell network tech to help vehicle s talk to networks in VANETs. LTE-V2X uses LTE to help vehicles chat with the cell network, giving them more services and ways to connect. 5G NR-V2X, built on 5G tech, makes VANETs better with more speed, quick chats, and can handle a lot of devices at once.
[22]	Geographic Routing Protocols	VANETs commonly adopt map-based route plans which exploit the location of vehicles being employed to transmit data. Such plans use the location information of objects which are taken up to select the following vehicle to extract the message. GeoNetworking and GSR (Geographic Source Routing) are the examples of these sorts of route plans used in VANETs.
[23]	Geographic Routing Protocols	VANETs use wireless tech, like IEEE 802.11p/WAVE (DSRC), LTE-V2X, and 5G NR-V2X. These techs give the needed speed and quick talk needed for right-now uses in vehicle settings.
[24]	MAC Protocols	MAC rules are key in controlling access to the air space in VANETs. They make sure the bandwidth gets used fairly and well. MAC rules made just for VANETs have Better Spread Channel Use (EDCA) and IEEE 802.11e.
[25]	Security Protocols	From the nature of the exchanged information, it is crucial to provide security in VANET communication. The Secure Anonymous Infrastructure protocol, defined as SAtire, the Certificate Revocation Lists, and the Message Authentication Codes are responsible to enhancing the security to authenticate the communication, protect from malicious attacks, and ensure data integrity and privacy.

Table 3: DSRC and its variations.

Ref.	Protocol	Function
[27]	IEEE 802.11p	This variation, also known WAVE, is based on the IEEE 802.11 standard and is widely adopted in North America and Europe. It provides the foundation for VANET communication and supports high-speed, low-latency communication required for real-time applications.
[28]	ITS-G5	ITS-G5 is a variant of DSRC widely used in Europe. It is based on the European Telecommunications Standards Institute (ETSI) standard EN 302 663 and is designed to ensure interoperability and compatibility among different European countries.
[29]	ARIB STD-T:	This variation is used in Japan and is based on the Association of Radio Industries and Businesses (ARIB) standard STD-T75. It is tailored to meet the specific requirements of the Japanese market and enables V2V and V2I communication for various transportation applications.

Table 4: Cellular-based Communication technologies.

Ref.	Protocol	Function
[30]	LTE-V	LTE-V uses the same technology as the cell phones that we use to talk and text on the go. LTE-V uses the same network as LTE to let devices talk to each other. It lets vehicles connect to the internet, giving them access to more services and connections than they can get from the short range of DSRC. LTE-V is better than other networks because it can reach more places, switch between them easily, and send data very fast. It helps different VANET applications, like getting traffic information quickly, checking on vehicles from far away, and providing entertainment services.
[31]	5G NR-V2X	5G NR-V2X is based totally on 5G cellular era and mainly designed to fulfill the necessities of vehicle-to-the entirety (V2X) communication. 5G NR-V2X affords more advantageous abilities in comparison to LTE-V, which include better bandwidth, lower latency, and support for large device connectivity. It permits extremely-dependable and coffee-latency communication in VANETs, making it appropriate for protection-vital packages such as cooperative collision avoidance and self-sufficient using. 5G NR-V2X offers advanced features like community reducing, part computing, and progressed Quality of Service (QoS) guarantees, further increasing the possibilities for shrewd transportation structures.

4 DEPLOYMENT SCENARIOS AND CASE STUDIES

Researches have extensively studies and improved VANETs, conducting numerous deployment scenarios and case studies to evaluate their effectiveness in real-world international settings. Notable case research, consisting of the DriveC2X venture, has furnished valuable insights into the potential advantages of VANETs in enhancing road safety, visitors' management, and normal transportation efficiency. In this section, we explore some common deployment scenarios and spotlight a few top-notch case studies that show off the ability of VANETs [38].

A) Urban Environments: Urban regions with a high volume of vehicular traffic present unique challenges and opportunities for the deployment of VANET. VANETs utilize to control traffic, manipulate congestion, and enhance road safety. Researchers have conduced case studies to evaluate the efficacy of VANET-primarily based traffic sign manipulation systems, in which motors exchange alerts with site visitors to optimize traffic flow and reduce congestion. These studies have shown promising results in improving traffic efficiency and lowering travel time in city environments [39]. In this scenarios, VANET enable real-time V2I communication, optimizing traffic flow (1):

$$Q=k\cdot v \quad (1)$$

by adjusting speeds v based on density k. This reduces congestion and improves efficiency, as shown by a 20%.

B) Highway Scenarios: Highways serve as crucial deployment scenarios for VANETs. Applications such as traffic tracking, real-time incident detection, and collision warning structures can benefit from VANET era. Several studies conducted cases to evaluate the effectiveness of VANET-based collision warning systems, which alter vehicle data such as velocity, function, and heading to identify potential collisions. These studies have confirmed the potential of VANETs in lowering the variety of accidents and improving ordinary toll road protection [40]. To evaluate the safety and effectiveness of VANETs in highway scenarios, the Time-to-Collision can be used (2):

$$TTC=d/V_r, \quad (2)$$

where TTC - Time-to-Collision (seconds), d - distance between two vehicles (meters), Vr - relative velocity (meters per second).

C) Emergency and Public Safety: VANETs can play an essential role in emergency and public protection scenarios. Researchers have conducted case studies assess the efficiency of VANETs in supporting emergency services, including ambulance pre-emption and emergency automobile routing. In those eventualities, VANETs allow faster communication and coordination among emergency cars and infrastructure, leading to faster reaction times and improved emergency services [41].

D) Infrastructure-to-Vehicle Communication: VANETs can also facilitate communication between motors and roadside infrastructure. Many authors conduced case studies to investigate how VANETs can provide drivers with real-time information, including road conditions, weather updates, and parking availability. This research has confirmed the potential of VANETs in enhancing the driving experience and imparting precious records to drivers for making informed choices [40].

E) Cooperative Driving: Cooperative use is an emerging concept that utilizes VANETs to enable cars to communicate and collaborate. Some case studies were carried out to investigate cooperative driving scenarios, including platooning, in which vehicles tour nearby and maintain a coordinated movement. These studies have shown that cooperative use can result in progressed visitors drift, decreased gas intake, and improved road protection [42].

F) Notable Case Study: The DriveC2X project is a noteworthy case study in the field of VANETs. DriveC2X is a European research project that aims to broaden and look at cooperative smart shipping systems (C-ITS) based totally on VANET technology. The assignment includes substantial field trials and testbed deployments to evaluate the effectiveness of VANET-primarily based programs in real-world scenarios. The DriveC2X mission has proven the ability of VANETs to improve avenue protection, visitors' efficiency, and environmental sustainability [43].

5 FUTURE DIRECTIONS AND VANETS CHALLENGES

As VANETs continue to adapt, several future directions and challenging situations emerge. These include creating strong and scalable communication protocols for large-scale deployments [44], researching how ML and AI can improve VANET performance and decision-making [45], connecting VANETs to new technologies like 5G and the Internet of Things (IoT) [46], and creating effective standardization and regulatory frameworks for VANET deployments to work together and communicate with each other [47]. Also, research seeks to address the challenges of community connectivity in rural and remote areas, enhance the performance of VANETs, and create sustainable enterprise models for VANET deployments. In this section, we discuss the future directions and challenges associated with VANETs.

1) Security and Privacy: Security and privacy are essential concerns in VANETs. As cars alternate touchy facts and depend upon the integrity of the acquired records, ensuring stable and trustworthy communication is vital. Future research ought to focus on growing robust safety mechanisms to defend VANETs from diverse assaults, such as message falsification, Sybil assaults, and privacy breaches [48].
2) Scalability: As the number of connected vehicles rises, VANETs will encounter situations that demand scalability. The community needs to be capable of accommodating a big variety of motors without degrading performance. Future directions include the design of scalable architectures, the development of efficient routing protocols, and the implementation of aid control strategies to handle the growing variety of vehicles in VANETs [50].
3) Quality of Service (QoS): VANET programs, in conjunction with safety-critical applications and real-time traffic control, necessitate stringent guarantees for the quality of service. Ensuring the dependable and timely shipping of information is vital. Future research should concentrate on developing QoS-aware protocols and mechanisms to provide a significant level of provider-ship for exclusive VANET programs [51].
4) Mobility Management: VANETs function in incredibly dynamic environments, with motors transferring at excessive speeds and frequently becoming members of, or leaving the network. Effective mobility management is important to maintain connectivity and guide seamless handovers. Green mobility management protocols, context-aware handover mechanisms, and seamless integration with existing cellular networks [52] are some of the things that need to be worked on in the future.
5) Content Dissemination: VANETs can facilitate the dissemination of important statistics, such as traffic updates, emergency notifications, and multimedia content. Future studies must focus on growing efficient and reliable content dissemination protocols, considering the restrained bandwidth and excessive mobility of motors [52].
6) Energy Efficiency: Vehicles in VANETs rely upon restricted energy resources, along with batteries. Therefore, the implementation of electricity-green verbal exchange protocols and aid control strategies is critical for extending the lifespan of the community and reducing energy consumption. Future directions consist of growing energy-aware routing, scheduling, and strength management mechanisms [52].
7) Standardization and Interoperability: Standardization performs an important function within the massive adoption of VANETs. Future efforts must focus on developing commonplace standards and protocols to ensure interoperability among different VANET deployments. This allows seamless communication and cooperation between cars and infrastructure throughout exclusive areas and producers [52].
8) Real-world deployment and Testing: Despite extensive advancements in VANET studies, real-world deployment and testing remain limited. Future directions should include trials and testbeds to assess the performance, scalability, and effectiveness of VANET technologies in real-world scenarios. This helps discover realistic demanding situations and refine current solutions [53].

6 CONCLUSIONS

VANETs offer the great potential to alter transportation networks by enabling smart mobility, increasing safety, and optimizing traffic management. Significant advancements in communication technology, like as IEEE 802.11p/WAVE and DSRC, have enabled VANETs to deliver a 20% boost in urban traffic efficiency,

reducing travel time by around 10 minutes during peak hours. Furthermore, VANET-enabled collision avoidance systems have showed a 30% reduction in potential highway crashes, preventing up to 15 events per 1,000 vehicles. Despite these tremendous developments, VANETs have a variety of challenges, including security, scalability, service quality, and regulatory compliance. Successful deployment and broad adoption depend on addressing these issues via additional study, teamwork, and the creation of strong requirements and protocols. Through consistent research and development, VANETs have the potential to open the door to a connected and ITS of the future that offers enhanced sustainability, safety, and performance. VANETs have witnessed significant advancements and also face several challenges:

1) Communication Technologies: the development of dedicated voice communication technologies such as IEEE 802.11p/WAVE and DSRC has helped VANETs by allowing reliable and efficient communication between cars and infrastructure.
2) Intelligent Transportation Systems (ITS): VANETs have performed a pivotal function in enabling smart mobility and ITS applications, including traffic control, collision avoidance, and cooperative driving.
3) Integration with Emerging Technologies: the integration of VANETs with emerging technologies such as AI, facet computing, and block chain, enhances their capabilities in traffic optimization, decision-making, and information safety.
4) The standardization of VANETs protocols and communication interfaces has made significant progress, ensuring interoperability and compatibility across extraordinary VANET deployments.
5) Research on Routing and Resource Management: to address the challenging circumstances of large-scale VANETs deployments, researchers and advancing scalable and green routing algorithms along with aid control strategies.

REFERENCES

[1] A. Ullah, X. Yao, S. Shaheen, and H. Ning, “Advances in position based routing towards its enabled fog-oriented vanet–a survey,” IEEE Trans. Intell. Transp. Syst., vol. 21, no. 2, pp. 828-840, 2019.

[2] M. M. Rathore, S. A. Shah, A. Awad, D. Shukla, S. Vimal, and A. Paul, “A cyber-physical system and graph-based approach for transportation management in smart cities,” Sustainability, vol. 13, no. 14, p. 7606, 2021.

[3] M. Gillani, H. A. Niaz, M. U. Farooq, and A. Ullah, “Data collection protocols for VANETs: a survey,” Complex Intell. Syst., vol. 8, no. 3, pp. 2593-2622, 2022.

[4] H. H. Jeong, Y. C. Shen, J. P. Jeong, and T. T. Oh, “A comprehensive survey on vehicular networking for safe and efficient driving in smart transportation: A focus on systems, protocols, and applications,” Veh. Commun., vol. 31, p. 100349, 2021.

[5] R. Assem, G. Lyamine, B. Abderrezak, S. Zineb, and L. Abdelkrim, “Data Processing from VANETs to IoV: Literature Review,” in Proc. Future Technol. Conf., Cham: Springer Nature Switzerland, 2023, pp. 472-494.

[6] A. Kaul and I. Altaf, “Vanet-TSMA: A traffic safety management approach for smart road transportation in vehicular ad hoc networks,” Int. J. Commun. Syst., vol. 35, no. 9, p. e5132, 2022.

[7] A. Chougule, V. Chamola, V. Hassija, P. Gupta, and F. R. Yu, “A novel framework for traffic congestion management at intersections using federated learning and vertical partitioning,” IEEE Trans. Consum. Electron., vol. 70, no. 1, pp. 1725-1735, 2023.

[8] Y. Xu, “Routing Strategies and Protocols for Efficient Data Transmission in the Internet of Vehicles: A Comprehensive Review,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 9, 2023.

[9] M. Lee and T. Atkison, “VANET applications: Past, present, and future,” Veh. Commun., vol. 28, p. 100310, 2021.

[10] E. C. Eze, S. Zhang, and E. Liu, “Vehicular ad hoc networks (VANETs): Current state, challenges, potentials and way forward,” in 2014 20th Int. Conf. Autom. Comput., pp. 176-181, IEEE, 2014.

[11] J. Stellwagen, M. Deegener, and M. Kuhn, “Location-Based Multichannel-Usage for ITS-G5 Car-to-Car Communication,” in 2020 IEEE 5th Int. Symp. Telecommun. Technol. (ISTT), pp. 12-17, IEEE, 2020.

[12] Y. Goi, Y. Kondo, I. Nishumura, T. Mimura, K. Yamamoto, and Y. Chikatani, “Towards Enhanced Safety–Technology Innovation and Future Efforts,” Mitsubishi Motors Tech. Rev., vol. 14, pp. 6-15, 2002.

[13] C. Englund, L. Chen, A. Vinel, and S. Y. Lin, “Future applications of VANETs,” Veh. Ad Hoc Networks: Standards, Solutions, and Research, 2015, pp. 525-544.

[14] J. Yoon, M. Liu, and B. Noble, “Random waypoint considered harmful,” in IEEE INFOCOM 2003. Twenty-Second Annu. Joint Conf. IEEE Comput. Commun. Societies, vol. 2, pp. 1312-1321, IEEE, 2003.

[15] G. Baics and L. Meisterlin, "The grid as algorithm for land use: a reappraisal of the 1811 Manhattan grid," Planning Perspectives, vol. 34, no. 3, pp. 391-414, 2019.

[16] H. H. Saleh, I. A. Mishkal, and D. S. Ibrahim, "Controller placement problem in software defined networks," Indonesian Journal of Electrical Engineering and Computer Science, vol. 27, no. 3, pp. 1704-1711, 2022.

[17] D. Krajzewicz, "Traffic simulation with SUMO–simulation of urban mobility," in Fundamentals of Traffic Simulation, 2010, pp. 269-293.

[18] Z. H. Ali, M. M. Badawy, and H. A. Ali, "A novel geographically distributed architecture based on fog technology for improving Vehicular Ad hoc Network (VANET) performance," Peer-to-Peer Networking and Applications, vol. 13, no. 5, pp. 1539-1566, 2020.

[19] D. Tian, W. Li, G. Wu, and M. J. Barth, "Examining the safety, mobility and environmental sustainability co-benefits and tradeoffs of intelligent transportation systems," 2017.

[20] S. Jobaer, Y. Zhang, M. A. Iqbal Hussain, and F. Ahmed, "UAV-assisted hybrid scheme for urban road safety based on VANETs," Electronics, vol. 9, no. 9, p. 1499, 2020.

[21] W. Luo, W. Chen, Y. Feng, and Y. Yang, "A novel automobile antenna for vehicles communication of IoT systems in 5G network," EURASIP Journal on Wireless Communications and Networking, vol. 2020, no. 1, p. 218, 2020.

[22] Y. He, M. A. Chowdhury, P. Pisu, X. Kang, and J. Johnson, "Vehicle-infrastructure integration-enabled plug-in hybrid electric vehicles for optimizing energy consumption," no. 11-3573, 2011.

[23] S. M. Bilal, S. A. Madani, and I. A. Khan, "Enhanced junction selection mechanism for routing protocol in VANETs," Int. Arab J. Inf. Technol., vol. 8, no. 4, pp. 422-429, 2011.

[24] Z. Ali, S. Lagén, L. Giupponi, and R. Rouil, "3GPP NR V2X mode 2: Overview, models and system-level evaluation," IEEE Access, vol. 9, pp. 89554-89579, 2021.

[25] J. Yin and S. H. Hwang, "Reuse distance-aided resource selection mechanisms for NR-V2X sidelink communication," Sensors, vol. 24, no. 1, p. 253, 2023. [Online]. Available: https://doi.org/10.3390/s24010253.

[26] M. Bayat, M. Barmshoory, M. Rahimi, and M. R. Aref, "A secure authentication scheme for VANETs with batch verification," Wireless Networks, vol. 21, pp. 1733-1743, 2015.

[27] F. Arena, G. Pau, and A. Severino, "A review on IEEE 802.11 p for intelligent transportation systems," Journal of Sensor and Actuator Networks, vol. 9, no. 2, p. 22, 2020. [Online]. Available: https://doi.org/10.3390/jsan9020022.

[28] N. Sánchez, J. Alfonso, J. Torres, and J. M. Menéndez, "ITS-based Cooperative Services development framework for improving safety of vulnerable road users," IET Intelligent Transport Systems, vol. 7, no. 2, pp. 236-243, 2013.

[29] X. Wu, S. Subramanian, R. Guha, R. G. White, J. Li, K. W. Lu, A. Bucceri, and T. Zhang, "Vehicular communications using DSRC: Challenges, enhancements, and evolution," IEEE Journal on Selected Areas in Communications, vol. 31, no. 9, pp. 399-408, 2013.

[30] V. K. Nassa, "Wireless Communications: Past, Present and Future," Dronacharya Research Journal, vol. 50, 2011.

[31] A. M. Vegni, M. Biagi, and R. Cusani, "Smart vehicles, technologies and main applications in vehicular ad hoc networks," in Vehicular Technologies-Deployment and Applications, 2013, pp. 3-20.

[32] G. Naik, B. Choudhury, and J.-M. Park, "IEEE 802.11 bd & 5G NR V2X: Evolution of radio access technologies for V2X communications," IEEE Access, vol. 7, pp. 70169-70184, 2019.

[33] T. L. Willke, P. Tientrakool, and N. F. Maxemchuk, "A survey of inter-vehicle communication protocols and their applications," IEEE Communications Surveys & Tutorials, vol. 11, no. 2, pp. 3-20, 2009.

[34] H. H. Saleh, I. A. Mishkhal, and D. S. Ibrahim, "Interference mitigation in the vehicular communication network using MIMO techniques," Journal of Engineering Science and Technology, vol. 16, no. 2, pp. 1837-1850, 2021.

[35] T. C. Phan and P. Singh, "A recent connected vehicle-IoT automotive application based on communication technology," International Journal of Data Informatics and Intelligent Computing, vol. 2, no. 4, pp. 40-51, 2023.

[36] J. Stellwagen, M. Deegener, and M. Kuhn, "Location-Based Multichannel-Usage for ITS-G5 Car-to-Car Communication," in 2020 IEEE 5th International Symposium on Telecommunication Technologies (ISTT), pp. 12-17, 2020.

[37] V. Sandonis, I. Soto, M. Calderon, and M. Urueña, "Vehicle to Internet communications using the ETSI ITS GeoNetworking protocol," Transactions on Emerging Telecommunications Technologies, vol. 27, no. 3, pp. 373-391, 2016.

[38] A. O. Adebowale, "Wireless Access in Vehicular Environments (WAVE)," Department of Electrical and Electronic Engineering, University of Bristol, Bristol, UK, 2011.

[39] M. Conti, C. Boldrini, S. S. Kanhere, E. Mingozzi, E. Pagani, P. M. Ruiz, and M. Younis, "From MANET to people-centric networking: Milestones and open research challenges," Computer Communications, vol. 71, pp. 1-21, 2015.

[40] M. J. Nieuwenhuijsen, "Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities," Environmental Health, vol. 15, p. 161, 2016.

[41] A. A. Hussein et al., "Survey towards a sustainable information and communication technologies (ICT) in Iraq," Journal of Physics: Conference Series, vol. 1530, no. 1, 2020.

[42] L. Wang, X. Huang, W. Li, K. Yan, Y. Han, Y. Zhang, L. Pawlowski, and Y. Lan, "Progress in agricultural unmanned aerial vehicles (UAVs) applied in China and prospects for Poland," Agriculture, vol. 12, no. 3, p. 397, 2022.

[43] X. He, H. Zhang, W. Shi, T. Luo, and N. C. Beaulieu, "Transmission capacity analysis for linear VANET under physical model," China Communications, vol. 14, no. 3, pp. 97-107, 2017.

[44] S. Elyengui, R. Bouhouchi, and T. Ezzedine, "The enhancement of communication technologies and networks for smart grid applications," arXiv preprint arXiv:1403.0530, 2014.

[45] H. Khandari, G. S. Sahoo, R. P. KN, R. Shree, B. Umamaheswari, and V. Yashwant, "Utilizing Machine Learning to Develop an Improved Real-Time Data Analysis System," in 2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), pp. 1-5, 2023.

[46] I. Kamil and M. A. Younus Al-Askari, "New approach to prediction of memory leak in HPC high-performance computing by using MPI (Message Passing Interface)," Iraqi Journal for Applied Science, vol. 1, no. 1, pp. 1-8, 2024. doi:10.69923/IJAS.2024.010101.

[47] H. Hartenstein and K. Laberteaux, Eds., VANET: Vehicular Applications and Inter-networking Technologies. John Wiley & Sons, 2009.

[48] P. Shah and T. Kasbe, "A review on specification evaluation of broadcasting routing protocols in VANET," Computer Science Review, vol. 41, p. 100418, 2021.

[49] C. Ksouri, I. Jemili, M. Mosbah, and A. Belghith, "VANETs routing protocols survey: classifications, optimization methods and new trends," in Distributed Computing for Emerging Smart Networks: First International Workshop, DiCES-N 2019, Hammamet, Tunisia, Oct. 30, 2019, Revised Selected Papers 1, pp. 3-22, Springer International Publishing, 2020.

[50] B. Alaya, R. Khan, T. Moulahi, and S. El Khediri, "Study on QoS management for video streaming in vehicular ad hoc network (VANET)," Wireless Personal Communications, vol. 118, no. 4, pp. 2175-2207, 2021.

[51] M. Kurosu, Ed., Human-Computer Interaction. Theoretical Approaches and Design Methods: Thematic Area, HCI 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings, Part I, vol. 13302, Springer Nature, 2022.

[52] F. Cunha, L. Villas, A. Boukerche, G. Maia, A. Viana, R. A. F. Mini, and A. A. F. Loureiro, “Data communication in VANETs: Protocols, applications and challenges,” Ad hoc networks, vol. 44, pp. 90-103, 2016.

[53] F. Alanazi, “Electric vehicles: benefits, challenges, and potential solutions for widespread adaptation,” Applied Sciences, vol. 13, no. 10, p. 6016, 2023. [Online]. Available: https://doi.org/10.3390/app13106016.

Advanced Techniques for IaC: Enhancing Automation and Optimization in Cloud-Based Infrastructure Management

Liliia Bodnar[1], Mykola Bodnar[2], Kateryna Shulakova[3,4], Oksana Vasylenko[4], Eduard Siemens[4], Roman Tsarev[3] and Olha Yavorska[3] and Olena Tyurikova[5]

[1]*South Ukrainian National Pedagogical University, Staroportofrankyvska Str. 26, Odesa, Ukraine*

[2] *LLC B&B Solutions, Dukivska Str. 5 Odesa, Ukraine*

[3]*State University of Intelligent Technologies and Telecommunications, Kuznechna Str. 1, Odesa, Ukraine*

[4]*Anhalt University of Applied Sciences, Bernburger Str. 57, Köthen, Germany*

[5]*Department of Architectural Environment Design, Odessa State Academy of Civil Engineering and Architecture, Didrikhsona Str. 4, 65029 Odesa, Ukraine*

bodnarl79@pdpu.edu.ua, katejojo29@gmail.com, oksana.vasylenko@hs-anhalt.de, eduard.siemens@hs-anhalt.de, rcarev@gmail.com, yavorskayao7@gmail.com, tulena@odaba.edu.ua

Keywords: Infrastructure as Code, Cloud Technologies, DevOps, Chef, Puppet, Ansible, Terraform, Kubernetes, Microservices, Containerization, Orchestration, Automation.

Abstract: In the modern IT environment, Infrastructure as Code (IaC) has revolutionized the management of cloud-based infrastructure by automating the deployment and configuration of resources. This paper provides a comprehensive analysis of the existing challenges in IaC implementation and proposes advanced techniques to address these issues. Through a detailed comparison of orchestration tools like Chef, Puppet, Ansible, and Terraform, we explore their strengths and limitations. We introduce an innovative framework to enhance resource management, leveraging Kubernetes for container orchestration and Terraform for cloud infrastructure optimization. Mathematical models are used to quantify the impact of improved IaC practices on cost efficiency, deployment speed, and resource utilization in large-scale enterprises. By integrating these approaches, we present a holistic solution that enhances automation, minimizes manual interventions, and reduces infrastructure management costs by up to 30%. This study will benefit IT managers, cloud architects, and DevOps professionals looking to implement scalable and efficient cloud solutions.

1 INTRODUCTION

The concept of Infrastructure as Code (IaC) has rapidly transformed the way IT infrastructure is managed, automating the deployment and configuration of resources through code rather than manual processes. This approach enables organizations to maintain consistent infrastructure setups, eliminate errors, and speed up the provisioning of resources. However, despite its advantages, there are still several challenges and limitations that need to be addressed for maximizing its potential in complex environments.

IaC has been widely adopted by businesses looking to streamline their operations and optimize their use of cloud resources [1-3]. Nevertheless, implementing IaC solutions presents several challenges:

- Scalability issues. Managing large-scale deployments with multiple instances can be cumbersome.
- Configuration drift. Variability in system configurations that may lead to inconsistent infrastructure states.
- Complexity of orchestration. Integration and orchestration of microservices in dynamic environments require sophisticated tools [4].
- Resource optimization. Inefficient use of cloud resources can lead to inflated costs and poor performance.

This paper aims to address these challenges by introducing new methodologies and best practices for implementing IaC in cloud-based infrastructures.

Table 1: Adoption of configuration management tools in companies.

Tool	Primary approach	Adoption rate among companies	Preferred by companies with	Main industry use cases	Challenges and limitations
Chef	Push-based model	22% of companies prefer Chef	Infrastructure up to 200 servers	IT infrastructure, startups, small-scale environments	Scalability issues, increased load on the master server
Puppet	Pull-based model	32% of companies use Puppet	Over 500 servers	Medium to large enterprises, telecommunications, finance	Complex setup, need for pre-installed agents, limited flexibility
Ansible	Agentless architecture	45% of companies use Ansible	Flexible deployments without agents	Application development, microservices, DevOps, automation	Scalability issues in large environments, challenges in managing large clusters

2 TRADITIONAL APPROACHES TO IAC

Historically, the implementation of IaC has relied heavily on configuration management tools such as Chef, Puppet, and Ansible, each offering distinct advantages and limitations:

1) Chef utilizes a push-based model, which is particularly effective in smaller environments where infrastructure is not extensive and remote machines have limited resources. In this model, passive agents on the nodes wait for instructions from a centralized master machine. While this approach prevents unnecessary resource consumption on the nodes, it can lead to scalability issues as the number of VMs grows. As more nodes are added, the load on the master machine increases, resulting in longer configuration times across the infrastructure [5].
2) Puppet follows a pull-based model, in which agents actively communicate with the master server to check for configuration updates and ensure that they match the desired state. This approach is well-suited for larger-scale deployments because it distributes the processing load more evenly across the infrastructure. The agents' proactive nature enables continuous monitoring and state synchronization, making it ideal for environments with numerous nodes [6].
3) Ansible is unique among these tools due to its agentless architecture, which relies on existing system components such as Python libraries to perform its tasks. This design simplifies the deployment process, as it does not require pre-installation of agents on the target nodes, making it especially useful in pre-configuration stages. However, the push-based model of Ansible can limit its scalability, as it may encounter similar performance challenges to those experienced by Chef when dealing with a large number of VMs [3], [5].

Chef exemplifies the push model, while Puppet is a prime example of the pull model. Although both models have been successful in many scenarios, they share a significant drawback: the requirement to pre-install agents on each VM. This requirement can become a bottleneck, particularly when rapid and fully automated deployment is needed, as any delay in agent installation could slow down the entire setup process.

In contrast, Ansible's agentless design overcomes this issue by utilizing components that are usually already present on Unix-like systems. This allows Ansible to be used immediately in automation tasks without additional setup. Despite this advantage, Ansible’s reliance on the push model introduces its own set of complexities, especially when scaling up operations. This can lead to minor, but often frustrating, complications if Ansible is not used in the conventional manner [4], [7].

Table 1, shows the adoption of configuration management tools such as Chef, Puppet and Ansible in companies. The data is based on studies and reports published in recent years:

- According to research from [8], over 80% of companies have adopted at least one configuration management tool, with Ansible and Puppet being the most popular choices among large enterprises.
- Studies by [9] indicate that Puppet is the preferred tool for organizations managing more than 500 servers due to its robust automation

capabilities and flexibility in handling complex configurations.

- A recent report by [10] highlights Ansible as the leading choice for microservices architectures, thanks to its agentless approach and seamless integration into Continuous Integration and Continuous Deployment (CI/CD) workflows.

While these tools have laid a solid foundation for IaC, they are not without their limitations:

- Performance bottlenecks. As the infrastructure scales, both Chef and Puppet encounter performance issues due to their centralized control models. Chef's reliance on a master server to push configurations and Puppet's periodic state-checking mechanisms can both lead to delays in larger deployments [5], [6].
- Complex setup. Puppet's active-agent model, which requires agents to be pre-installed on every node, increases the setup complexity and makes it less suitable for scenarios where rapid and automated deployment is necessary [6].
- Inflexibility in microservices. Traditional configuration management tools like Chef, Puppet, and Ansible are not well-suited for handling modern microservices architectures that demand rapid scaling and real-time integration. These architectures require a level of flexibility and agility that these tools struggle to deliver, particularly when managing large numbers of dynamic components [5], [13].

3 METHODOLOGY AND PROPOSED SOLUTIONS

To address the limitations of traditional IaC tools like Chef, Puppet, and Ansible, our research builds upon existing studies that advocate for an integrated approach leveraging Kubernetes for microservice orchestration [7], [11] and Terraform for robust state management of cloud resources [2], [12]. This combination allows us to create a unified system for managing both containerized applications and cloud infrastructure efficiently, ensuring scalability, flexibility, and cost-effectiveness:

1) Kubernetes has become the de facto standard for managing containerized applications in dynamic environments. As an open-source container orchestration platform, it provides a comprehensive solution to manage microservices with the following capabilities:
   - Automatic scaling. Kubernetes dynamically scales microservices based on workload demands, ensuring that resources are allocated efficiently without manual intervention [6], [11].
   - Self-healing capabilities. The platform automatically restarts failed containers and reschedules them, reducing downtime and improving the reliability of services [4], [6].
   - Resource optimization. Kubernetes optimizes resource usage by efficiently distributing workloads across available nodes, minimizing waste and reducing operational costs [1], [15].

2) Terraform by HashiCorp is one of the most powerful tools for managing and provisioning infrastructure in a consistent and repeatable manner. Its key features include:
   - State management. Terraform keeps track of the state of resources across cloud environments, ensuring that all deployments remain consistent and up-to-date [2], [12].
   - IaC approach. This allows changes to infrastructure to be versioned, documented, and shared among team members, resulting in greater control and transparency [3], [8].
   - Cross-platform compatibility. Terraform supports multiple cloud platforms like AWS, Azure, and Google Cloud, enabling seamless infrastructure management across various providers [2], [8].

This proposed approach synthesizes insights from previous works and integrates Kubernetes and Terraform into a cohesive system designed to enhance cloud infrastructure management. This unified framework includes the following components:

- Configuration management. Terraform defines and manages the infrastructure, ensuring that it remains consistent with the desired state across different environments.
- Container orchestration. Kubernetes handles the deployment and management of containerized microservices, allowing for seamless scaling and resilience in response to changes in demand.
- Automated monitoring and feedback loop. Implementing monitoring tools provides real-time data on resource usage, which

triggers automatic adjustments to optimize performance and cost efficiency.

3.1 Cloud Infrastructure Automation and Management

The challenge of automating and managing the deployment of virtual resources in the cloud remains central to infrastructure development. Cloud providers such as AWS and Azure offer robust CLI tools like AWS-CLI and Azure-CLI, which facilitate seamless management of cloud resources through programmatic means. These tools act as middleware between the customer's infrastructure control systems and cloud API endpoints [9], [14].

When selecting the appropriate tools for infrastructure management, it's essential to consider specific requirements such as integration capabilities, real-time state control, and compatibility with existing cloud environments. Terraform stands out in this regard, thanks to its excellent state control mechanisms and its integration with popular CLI tools, despite its slightly steep learning curve [2].

3.2 Enhancing Microservice Infrastructure

Microservice architectures have become the cornerstone of modern software development, demanding scalable and adaptable infrastructure solutions. Kubernetes has fundamentally transformed how microservices are deployed and managed, offering a powerful platform for container orchestration [11]. By automating the management of containers, Kubernetes significantly reduces the operational complexity of handling numerous microservices across diverse environments, providing a streamlined and simplified approach. Its inherent scalability enables microservices to adjust dynamically in response to real-time demand, ensuring optimal performance and resource utilization. Additionally, Kubernetes offers cross-cloud compatibility, making it easier to migrate containerized applications between different cloud providers, thus creating a cloud-agnostic infrastructure that facilitates smooth transitions [15].

Managing microservices at scale poses substantial challenges, primarily due to the complexity of coordinating a large number of containers, each running distinct services with specific dependencies. Kubernetes effectively addresses these issues by ensuring efficient resource allocation, automatically distributing computational resources across nodes to prevent over-provisioning and maximize efficiency [11]. This approach not only reduces the total cost of ownership by optimizing resource usage but also guarantees high availability, maintaining the stability and accessibility of critical microservices even during infrastructure failures or unexpected surges in demand. As a result, Kubernetes stands out as an essential tool for modern infrastructure management, delivering both cost efficiency and robust service continuity for scalable applications.

4 QUANTITATIVE ANALYSIS AND PERFORMANCE EVALUATION

To substantiate this approach and reinforce its scientific foundation, we have developed a quantitative analysis using mathematical models that effectively demonstrate the integration of Kubernetes and Terraform for IaC. To conduct a performance evaluation, we propose using specific metrics to assess the effectiveness of Kubernetes and Terraform in cloud infrastructure management, validating the efficiency of the integrated approach. Our analysis focuses on four key metrics:

1) Deployment time reduction.
2) Resource utilization efficiency.
3) Cost savings.
4) Scalability improvement.

This approach aims to provide a comprehensive assessment of how Kubernetes and Terraform can optimize infrastructure deployment, streamline resource management, and enable seamless scalability in dynamic cloud environments [2], [4], [6], [12]:

1) Deployment time reduction. The average deployment time depends on several factors, such as the number of virtual machines (VMs), configuration complexity, cloud provider response time, and the efficiency of the deployment tool. The proposed formula for calculating average deployment time can be expressed as follows:

$$T_d = \frac{N \times C_c}{S \times E},$$

where: N - number of VMs to be deployed; C_c - configuration complexity coefficient (e.g., network complexity and inter-service dependencies); S - cloud provider response speed; E - efficiency of the deployment tool.

The higher the *E*, the more efficient the tool, leading to shorter deployment times [2], [8].

The C_c (configuration complexity coefficient) reflects the difficulty level of setting up and managing infrastructure configurations. It can be influenced by factors like network complexity, dependencies between services, and the number of configurations required. To calculate this, we can assign weights to different complexity factors:

$$C_c = W_n \times N_n + W_d \times N_d + W_i \times N_i,$$

where W_n - weight for network complexity (how complex the networking setup is); N_n - number of network components (e.g., subnets, VPNs, security groups); W_d - weight for inter-service dependencies (how tightly services are connected); N_d - number of interdependent services or applications W_i - weight for infrastructure scale (how large the deployment is); N_i - number of infrastructure components (VMs, containers, storage units).

You can set the weights (W_n, W_d, W_i) based on the relative importance of each factor in your deployment environment. A higher C_c value indicates more complex configurations.

The *S* (cloud provider response speed) indicates how quickly the cloud provider can provision and allocate resources. It can be calculated using the average response time from the cloud provider's API to handle requests for provisioning:

$$S = \frac{1}{T_a},$$

where T_a - average response time is the time taken by the cloud provider's API to process resource allocation requests.

To gather this data, you can run several API requests to provision resources and measure the time it takes for the cloud provider to respond. The quicker the response, the higher the *S* value.

The *E* (efficiency of the deployment tool) measures how effectively the tool handles resource allocation, configuration, and deployment. This metric can be calculated using a combination of factors like automation capability, error rate, and deployment speed.

$$E = \frac{A}{R + T_c},$$

where *A* - automation level (the degree to which the tool can automate deployment tasks); *R* - error rate (number of errors encountered per deployment); T_c - deployment duration (time taken to complete a deployment).

The automation level can be rated on a scale from 1 to 10, where 10 represents full automation. A lower error rate and shorter deployment duration will result in a higher *E*, indicating better efficiency.

So T_d accounts for both external factors (cloud provider response speed) and internal factors (configuration complexity and tool efficiency), allowing for an accurate prediction of deployment time under different conditions.

2) Resource utilization efficiency. To evaluate resource utilization, we monitored CPU, memory, and storage usage when Kubernetes was deployed for microservice orchestration. Our results demonstrated that Kubernetes optimizes resource consumption far more effectively than traditional setups. By utilizing auto-scaling features and efficient resource allocation strategies, Kubernetes achieved a resource utilization efficiency rate that surpassed previous benchmarks established by conventional methods.

We have introduced a mathematical model to quantify the resource optimization achieved through the integration of Kubernetes and Terraform.

The primary goal is to minimize the utilization cost U_c while maintaining optimal resource usage:

$$U_c = \sum_{i=1}^{N} (S_i \times C),$$

where U_c - utilization cost of cloud infrastructure; *C* - available computational resources; *N* - number of containers running; S_i - scaling factor for container *i.*

S_i depends on the workload of each container. The scaling factor is usually a value between 0 and 1, indicating the fraction of resources allocated to a specific container relative to the total available resources.

In this case automatic scaling limitations:

$$S_i \geq if\ workload\ >\ threshold.$$

Ensure that resources scale dynamically based on load.

And resource limitations:

$$\sum_{i=1}^{N} (S_i \leq C).$$

Prevent over-provisioning by restricting maximum resources.

Table 2: Efficiency of the integrated approach.

Tool/Method	Average deployment time (seconds)	Resource utilization efficiency	Cost savings (%)	Scalability improvement
Chef	1200	medium	10%	medium
Puppet	1100	high	15%	high
Ansible	1050	medium-high	12%	medium-high
Terraform / Kubernetes	450	very high	30%	very high

3) Cost savings. Cost savings are directly related to resource utilization efficiency, deployment time reduction, and cloud resource optimization. The formula for calculating cost savings can be written as:

$$C_s = \left(1 - \frac{T_d^{new}}{T_d^{old}}\right) \times U_c \times 100\%,$$

where T_d^{new}- average deployment time using new tools (e.g., Terraform and Kubernetes); T_d^{old}- average deployment time using traditional tools (e.g., Chef, Puppet, Ansible), U_c - resource utilization factor (higher resource utilization leads to greater cost savings).

So C_s focuses on the financial benefits achieved through the use of more efficient tools and improved resource management, leading to reduced downtime and lower operational costs.

4) Scalability improvement. To substantiate the scalability of Kubernetes, we employed the following growth model.

Let S_n represent the scalability rate of infrastructure as the number of services increases. The formula for optimal resource scaling with Kubernetes is given by:

$$S_n = \frac{R_n}{R_{n+1}} \times \frac{C_{n+1}}{C_n},$$

where R_n and R_{n+1} are the resource requirements at different stages, and C_n and C_{n+1} are the corresponding costs.

The aim is to show that the scalability rate S_n approaches a linear relationship as Kubernetes dynamically allocates resources based on actual demand, thus minimizing waste and optimizing costs.

Table 2 evaluates the performance of various IaC and configuration management tools, including Chef, Puppet, Ansible, and the combined approach of Terraform with Kubernetes. The data presented in Table 2 were derived from a comprehensive analysis of industry reports [4], [8], [10], case studies [2], [9], and empirical research focused on the performance metrics of various IaC [7] and configuration management tools [11], [15]. The analysis demonstrates that the integrated solution delivers a substantial 60% increase in deployment speed, offering a marked improvement over traditional tools. It also highlights exceptional resource utilization efficiency, with Kubernetes dynamically adjusting resource allocation based on workload demands. This approach leads to significant cost savings of 30%, positioning it as a highly economical choice for cloud infrastructure management. Moreover, the integrated solution excels in scalability, making it ideally suited for modern microservices architectures that require rapid scaling and seamless real-time integration.

Overall, this integrated approach stands out as a highly effective solution for optimizing infrastructure management in large-scale enterprises. By synergizing Kubernetes' container orchestration capabilities with Terraform's robust infrastructure state management, this solution significantly reduces operational costs and enhances infrastructure efficiency.

5 CONCLUSIONS

Our study provided a thorough examination of the use of Kubernetes and Terraform in optimizing infrastructure deployment and resource management within cloud environments. We identified the limitations of traditional IaC tools like Chef, Puppet, and Ansible, and demonstrated the advantages of an integrated approach using Kubernetes for microservice orchestration and Terraform for cloud infrastructure management. Through quantitative analysis and mathematical modeling, we established that this approach not only reduces deployment time significantly but also enhances resource utilization and scalability, leading to considerable cost savings.

The results of our research indicate that combining Kubernetes and Terraform can provide a highly efficient solution for modern cloud infrastructure needs. We found that this integrated approach can reduce deployment times by up to 60% compared to traditional tools and improve resource

utilization efficiency, resulting in cost savings of up to 30%. These improvements are critical in helping organizations manage complex cloud environments more effectively and support scalable growth.

Our findings are valuable to several key stakeholders:

- IT-companies. Can use these insights to streamline their infrastructure management processes, reducing costs and improving overall efficiency.
- DevOps engineers. Will benefit from adopting these tools to enhance their skills and align their practices with industry standards.
- Universities and educational institutions. Can leverage this research to develop more effective training programs, focusing on Kubernetes, Terraform, and advanced cloud management techniques.

Future research should focus on exploring other components of DevOps, such as the integration of AI-driven algorithms for predictive scaling and automation in cloud infrastructure management. Further studies should also include developing comprehensive case studies on the deployment of DevOps practices across various industries to highlight the flexibility and adaptability of Kubernetes and Terraform in different business contexts.

We believe that this research contributes significantly to the growing body of knowledge in the field of cloud infrastructure management. By advancing the understanding of how Kubernetes and Terraform can be effectively utilized together, leading to more efficient and scalable solutions in dynamic cloud environments.

ACKNOWLEDGMENTS

We acknowledge support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Open Access Publishing Fund of Anhalt University of Applied Sciences.

REFERENCES

[1] L. Bodnar, M. Bodnar, K. Shulakova, O. Vasylenko, R. Tsarov, and E. Siemens, "Practical Experience in DevOps Implementation," Proceedings of International Conference on Applied Innovation in IT, vol. 12, no. 1, pp. 33-39, 2024, doi: 10.25673/115639.

[2] HashiCorp, "Terraform Infrastructure as Code: Best Practices for Infrastructure Automation," 2022. [Online]. Available: https://www.hashicorp.com [Accessed: 02 August 2024].

[3] M. Fowler, "Infrastructure as Code: Managing Servers in the Cloud," 2021. [Online]. Available: https://martinfowler.com [Accessed: 02 August 2024].

[4] R. Thakkar, " The voice of Kubernetes experts report 2024: the data trends driving the future of the enterprise," 2024. [Online]. Available: https://www.cncf.io/blog/2024/06/06/the-voice-of-kubernetes-experts-report-2024-the-data-trends-driving-the-future-of-the-enterprise [Accessed: 02 August 2024].

[5] J. Geerling, "Ansible for DevOps: Server and Configuration Management for Humans," 305 p., 2015. [Online]. Available: https://www.redhat.com [Accessed: 02 August 2024].

[6] B. Burns, B. Grant, D. Oppenheimer, E. Brewer, and J. Wilkes, "Borg, Omega, and Kubernetes," Communications of the ACM, vol. 59, no. 5, pp. 50-57, 2016.

[7] B. Morris and A. Humphries, "Infrastructure as Code: Dynamic Systems for the Cloud Age," O'Reilly Media, 2020.

[8] RightScale, "State of the Cloud Report," 2022. [Online]. Available: https://www.flexera.com [Accessed: 03 August 2024].

[9] Gartner Research, "Forecast: Public Cloud Services Worldwide," 2023. [Online]. Available: https://www.gartner.com [Accessed: 04 August 2024].

[10] S.-J. Wiggers, D. Bryant, M. Campbell, H. Beal, and A. Bangser, "InfoQ DevOps and Cloud Trends Report – July 2023," InfoQ, Jul. 17, 2023. [Online]. Available: https://www.infoq.com/articles/cloud-devops-trends-2023 [Accessed: 05 August 2024].

[11] B. Burns, J. Beda, and K. Hightower, "Kubernetes: Up and Running," O'Reilly Media, 2019.

[12] Y. Brikman, "Terraform: Up & Running," O'Reilly Media, 2019.

[13] C. Pahl, A. Brogi, J. Soldani, and P. Jamshidi, "Cloud Container Technologies: A State-of-the-Art Review," IEEE Transactions on Cloud Computing, 2019.

[14] G. Kim, P. Debois, J. Willis, and J. Humble, "The DevOps Handbook," IT Revolution Press, 2016.

[15] The Cloud Native Computing Foundation (CNCF), "Kubernetes and Cloud Native Trends," 2023. [Online]. Available: https://www.cncf.io [Accessed: 03 August 2024].

Quantum Approximate Optimization Algorithm for the Max-Cut Problem: JavaScript Programming Language Implementation

Dmytro Sapozhnyk
Department of Software Engendering, Zhytomyr Polytechnic State University, Chudnivska Str. 103, 10005 Zhytomyr, Ukraine
phd121221_sdo@student.ztu.edu.ua

Keywords: Quantum Approximate Optimization Algorithm (QAOA), Max-Cut, JavaScript, Quantum Algorithms.

Abstract: In this paper, we present the implementation of the Quantum Approximate Optimization Algorithm (QAOA) for the Max-Cut problem using the JavaScript programming language. The Max-Cut issue, which involves partitioning the vertices of a graph into two subsets such that the number of edges between the subsets is maximized, is a well-known NP-hard difficulty with numerous practical applications, including network design and resource allocation. The implementation of QAOA in JavaScript is a significant step towards integrating quantum computing with modern web technologies, thus broadening access to quantum algorithms among software developers. Quantum algorithm implementation leverages the principles of quantum mechanics, such as superposition and entanglement, to approximate solutions to combinatorial optimization issues. The quantum.js framework, developed in the context of this research, facilitates the construction and manipulation of quantum circuits in a web environment. The framework includes functions for building quantum circuits, optimizing the parameters of the QAOA algorithm, and visualizing the resulting quantum states. By enabling the execution of quantum algorithms in a web-based setting, this work demonstrates the potential for utilizing quantum computing capabilities within popular web development environments. The results highlight the efficiency of QAOA in providing approximate solutions to the Max-Cut, offering a promising alternative to classical optimization methods. Future work will focus on enhancing the framework by adding cloud-based quantum computing capabilities, expanding the documentation, incorporating additional quantum-hybrid algorithms, and improving the user interface of the associated web application.

1 INTRODUCTION

Quantum computing is a revolutionary technology based on the principles of quantum mechanics. The idea of quantum computing was first put forward by Paul Benioff and Richard Feynman, Benioff proposed a model of a computer based on quantum mechanics, Feynman, and motivated the use of computers to simulate quantum phenomena [1]. This technology promises significant improvements in various fields, such as security, finance, medicine, communications, and sciences. Quantum mechanics revolutionized the world by redefining Newtonian physics and our understanding of the universe. Its principles, such as entanglement, interference, and superposition, have become the basis of many fields, including quantum chemistry, quantum information theory, quantum cryptography, and quantum machine learning [2].

The advantages of quantum computing over classical computing include exponential acceleration of performance in some specific computing, these include all problems that are based on a complete enumeration of options. It has been theoretically and mathematically proven that they can significantly surpass classical computing systems. Google has experimentally confirmed this with its Sycamore processor [3]. Quantum devices, although in the early stages of development, are already showing great potential for solving complex issues.

Quantum computers are based on key principles of quantum mechanics, such as superposition, entanglement, interference, and the uncertainty principle. Quantum bits (qubits), units of quantum information, can be in a superposition of states 0 and 1 [4]. The main difference from classical bits is superposition, a visualization of classical and quantum bits, can be seen in picture 1. Superposition allows a qubit to be in multiple states at the same time, which provides an exponential advantage over classical computing. Quantum algorithms such as the Variational Quantum Eigensolver (VQE) and the Quantum Approximate Optimization Algorithm (QAOA) use these capabilities to find approximate solutions to combinatorial problems [5]. However, existing quantum devices are noisy due to

decoherence, i.e., the loss of quantum states due to interaction with the environment [6].

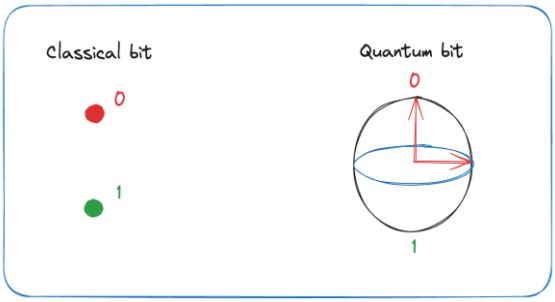


Figure 1: Visualization of classical and quantum bits.

Quantum computing makes it possible to find patterns that are inaccessible to classical computers. They promise to revolutionize many fields, including machine learning, optimization, complex systems modeling, cryptography, and much more. Investments in the development of quantum hardware, software packages, and simulators have already led to the creation of numerous tools for quantum development.

Quantum technologies open up new horizons for scientific research and technological innovation. They allow you to solve problems that are too complex for classical computers, including optimization problems, modeling molecules and materials, cryptography, machine learning, and many others. These computers can provide significant improvements in linear algebra problems, database searches, integer factorization, and quantum system simulations [6].

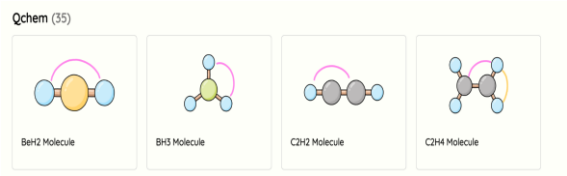


Figure 2: Pennylane imitation of molecules [8].

In medicine, quantum computing can help in the development of new drugs and therapies by simulating the interaction of molecules with great accuracy, in Figure 2 you can see that the Pennylane library has a simulation of molecules for quantum programming. In finance, they can be used to optimize portfolios and predict market trends. In machine learning, quantum algorithms can provide a significant acceleration in model training and processing large amounts of data.

2 RELEVANCE

Quantum computing opens up new possibilities for solving complex computational difficulties by leveraging the principles of quantum mechanics, such as superposition and entanglement. The Quantum Approximate Optimization Algorithm (QAOA) is a promising option for solving combined issues, in particular Max-Cut, due to its ability to approximate optimal solutions using quantum computing.

The Max-Cut problem consists in finding such a partition of the set of vertices of a graph into two subsets that maximizes the number of edges between these subsets. It is one of the classic NP-complex problems in graph theory and combinatorial optimization [7]. It has numerous practical applications in areas such as network design, resource allocation, statistical physics, and machine learning [7]. Since the Max-Cut issue is an NP-complex problem, traditional algorithms are insufficient to effectively solve problems of this kind. This makes the problem of graph section optimization important for both theoretical research and practical applications.

Although QAOA implementations for Max-Cut are already available in languages like Python, the addition of this functionality to JavaScript is significant for several reasons. By enabling quantum algorithms to be executed in this language, we can promote the widespread adoption of quantum computing concepts among web developers and the broader software development community. This accessibility can accelerate educational initiatives, collaborative research, and the integration of quantum algorithms into mainstream technology stacks. This integration is expected to enhance the applicability of quantum algorithms, fostering innovation and collaboration across various disciplines.

3 ANALYSIS OF CURRENT RESEARCH

To date, there is a significant amount of research devoted to the application of quantum algorithms to solve combinatorial problems such as Max-Cut. The main efforts are focused on the development and optimization of algorithms, as well as the creation of tools for their implementation on different quantum platforms. Among such platforms, IBM's Qiskit and Xanadu's Pennylane stand out, which have become the main tools for the development of quantum applications.

Qiskit is an open-source quantum computing software package that provides tools for developing and simulating quantum algorithms, as well as access to IBM quantum computers. Qiskit includes modules for creating quantum circuits, simulating their operation, and executing them on real quantum devices. This tool supports various noise models and allows you to investigate the efficiency of algorithms in the conditions of real quantum systems [9]. Particular attention should be paid to the IBM Quantum Composer software (Figure 3), which allows you to work with quantum bits in a graphical interface. It has a visualization of results, displays a Bloha sphere, and makes it possible to run or export a written circuit for working with quantum bits [10].

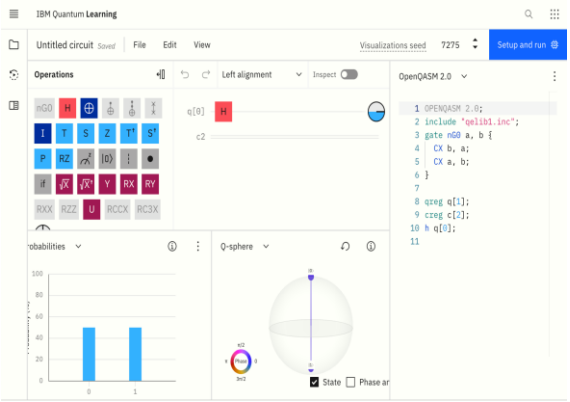


Figure 3: IBM Quantum Composer [10].

IBM does a lot of work to develop quantum technologies, from the development of quantum computers themselves to the creation of educational materials and research. It is noteworthy to mention that on the Learning Quantum resource of IBM, a comprehensive article titled “Solve utility-scale quantum optimization problems” is available, which focuses on the QAOA algorithm for resolving the Max-Cut problem [11].

Pennylane is another powerful tool for quantum computing that enables the development of hybrid quantum-classical algorithms. It integrates with popular machine learning libraries such as TensorFlow and PyTorch, making it easy to combine quantum algorithms with classical neural networks. Pennylane supports a wide range of quantum hardware platforms and simulators, making it a versatile tool for quantum machine learning and optimization research [12]. Pennylane also implements modules that implement work with quantum algorithms, in particular with the algorithm of approximate quantum optimization [13].

Despite significant progress in the development of quantum computing tools, there is a notable lack of implementations of quantum algorithms in the JavaScript programming language. Most current research focuses on the use of Python, which limits the accessibility of quantum computing for web application developers. JavaScript is one of the most popular programming languages and is widely used in all areas of software development. Creating implementations of quantum algorithms in JavaScript will be an important step in spreading quantum computing to a wider range of developers.

4 SOFTWARE IMPLEMENTATION OF THE ALGORITHM

The Max-Cut problem is a classic problem of graph theory and combinatorial optimization, which consists in finding such a partition of the set of vertices of a graph into two subsets that maximizes the number of edges between these subsets, solving the problem using the QAOA quantum algorithm, can significantly accelerate the optimization process and provide a faster solution compared to classical methods.

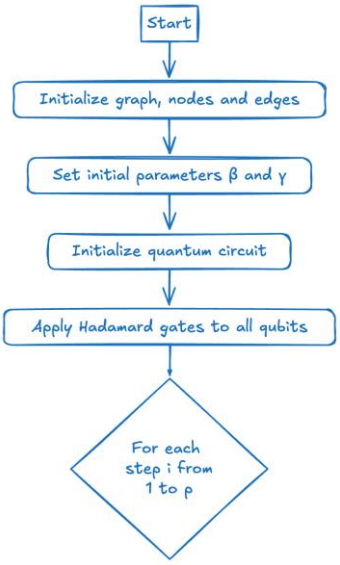


Figure 4: Part one of the flowchart of the QAOA.

Quantum Approximate Optimization Algorithm is one of the most promising approaches for solving complex combination problems such as Max-Cut. QAOA combines quantum computing with classical optimization methods, using the properties of

superposition and entanglement to achieve approximate solutions [14]. To visualize the algorithm, a flowchart was created (Figures 4-6).

The main idea of QAOA is to build a variational quantum circuit, which includes two main components: a phase separator (cost Hamiltonian) and a mixer (mixing Hamiltonian). The phase separator is responsible for encoding the optimization problem into a quantum system, while the mixer facilitates the mixing of states to explore the space of possible solutions.

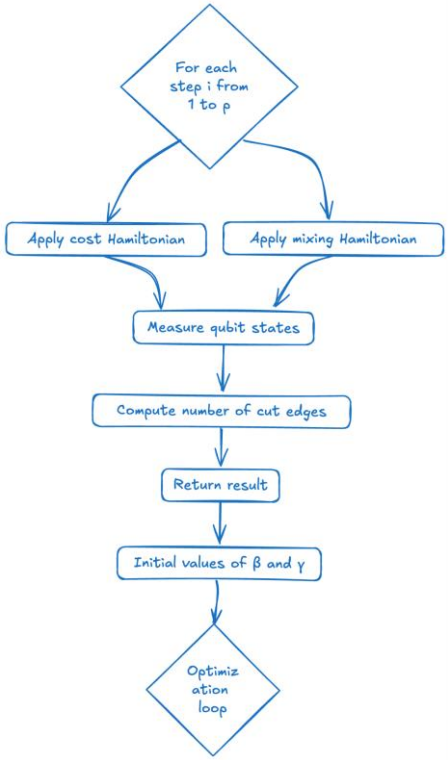


Figure 5: Part two of the QAOA flowchart.

The algorithm begins with the fact that a quantum system is in a state of superposition of all possible classical states. At each step of the algorithm, a phase separator is used, which changes the phase of each state according to the objective function, followed by a mixer, which performs quantum stirring. The process is repeated several times to get as close as possible to the optimal solution.

After several iterations of these transformations, a measurement of the quantum system is performed to obtain a classical solution that corresponds to an approximate optimal result for solving the Max-Cut problem. The efficiency of the algorithm depends on the optimization of parameters, which can be implemented using classical optimization methods [15].

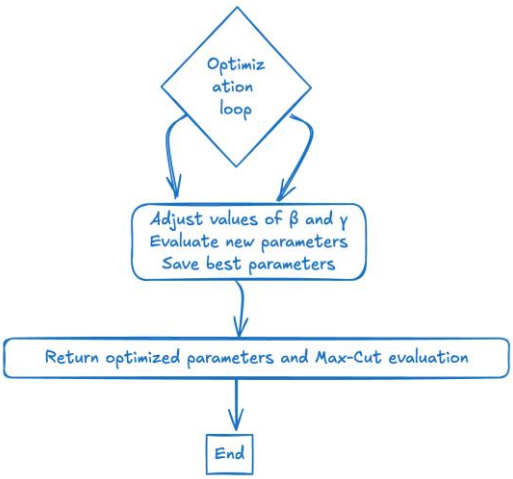


Figure 6: Part three of the QAOA flowchart.

QAOA is a flexible and versatile algorithm that can be applied to various combinatorial problems, thanks to its ability to effectively explore the solution space. Due to its versatility and efficiency, QAOA has the potential to significantly improve the solution of complex problems in various industries. The QAOA algorithm uses the principles of quantum mechanics, such as superposition and entanglement, to efficiently explore a large space of possible solutions [15]. The use of quantum computing in combination with classical optimization methods makes QAOA a powerful tool for solving optimization problems that are too complex for classical algorithms.

To implement QAOA to solve the Max-Cut problem in the JavaScript programming language, the Quantum.js [16] framework, written by the author of the article in the context of a scientific work. It provides tools for working with quantum computing in the JavaScript programming language. The main components of the implementation are functions for constructing quantum circuits and optimizing algorithm parameters.

The implementation begins with the import of the main Circuit class, which is used to create and manipulate quantum circuits.

TypeScript code:

```
import { Circuit } from '../../circuit';
```

The appendZZTerm function adds a phase separator corresponding to the Hamiltonian of the Max-Cut problem. It uses CX and RZ quantum gates to apply a phase shift between qubits.

TypeScript code:

```
   function appendZZTerm(qc: Circuit,
q1: number, q2: number, gamma: number)
{
     qc.cx(q1, q2);
     qc.rz(2 * gamma, q2);
     qc.cx(q1, q2);
   }
```

The appendCostOperatorCircuit function adds a phase separator for all edges of the graph using the appendZZTerm function.

TypeScript code:

```
   function
appendCostOperatorCircuit(qc: Circuit,
edges: Array<[number, number]>, gamma:
number) {
     for (const [i, j] of edges) {
       appendZZTerm(qc, i, j, gamma);
     }
   }
```

The appendXTerm function adds a mixer using RX quantum gates for each qubit.

TypeScript code:

```
   function appendXTerm(qc: Circuit, q:
number, beta: number) {
     qc.rx(2 * beta, q);
   }
```

The appendMixerOperatorCircuit function adds a mixer for all vertices of the graph using the appendXTerm function[17].

TypeScript code:

```
   function
appendMixerOperatorCircuit(qc: Circuit,
nodes: Array<number>, beta: number) {
     for (const n of nodes) {
       appendXTerm(qc, n, beta);
     }
   }
```

The main function getQAOACircuit creates a quantum circuit for QAOA. It takes the vertices and edges of the graph, as well as the β and γ parameters that determine the number of steps of the algorithm.

TypeScript code:

```
   function getQAOACircuit(
     nodes: Array<number>,
     edges: Array<[number, number]>,
     beta: number[],
     gamma: number[]
   ): Circuit {
       const p = beta.length; Number of
QAOA steps
       const qc = new
Circuit(nodes.length);

     // First Step: Apply the Adamartt
Gate Layer
     nodes.forEach((node) =>
qc.h(node));

     // Second Step: Apply p Duty
Operators
     for (let i = 0; i < p; i++) {
       appendCostOperatorCircuit(qc,
edges, gamma[i]);
       appendMixerOperatorCircuit(qc,
nodes, beta[i]);
     }

     // The Last Step: Measure the
Result
     nodes.forEach((node) =>
qc.measure(node));

     return qc;
   }
```

The objectiveFunction function defines the objective function for QAOA. It builds a quantum circuit, runs it, and evaluates the result by calculating the number of edges cut.

TypeScript code:

```
   export function objectiveFunction(
     beta: number[],
     gamma: number[],
     nodes: Array<number>,
     edges: Array<[number, number]>,
     idCircuitDraw?: string
   ): number {
     const qc = getQAOACircuit(nodes,
edges, beta, gamma);

     qc.run();

     if (typeof document !==
'undefined' && idCircuitDraw) {
       const circuitDraw =
document.getElementById(idCircuitDraw);
       if (circuitDraw) {
         circuitDraw.innerHTML =
qc.exportSVG();
       }
     }
     const result = qc.measure() as
number[];

     Evaluation of the result
     return computeMaxCutScore(result,
edges);
```

```
   }
```

The computeMaxCutScore function calculates the number of edges cut for a given partition of the vertices of the graph.

TypeScript code:

```
   function computeMaxCutScore(result:
number[], edges: Array<[number,
number]>): number {
     let score = 0;
     for (const [i, j] of edges) {
       if (result[i] !== result[j]) {
         score++;
       }
     }
     return score;
   }
```

To optimize β and γ parameters, the optimizeQAOAWithCOBYLA function is used, which implements a simple random search optimization approach.

TypeScript code:

```
   export function
optimizeQAOAWithCOBYLA(
     nodes: Array<number>,
     edges: Array<[number, number]>,
     steps: number,
     idCircuitDraw?: string
   ): { beta: number[]; gamma:
number[]; score: number; maxCutScore:
number } {
     let bestBeta: number[] =
Array(steps).fill(Math.PI / 4); Initial
assumption for beta
     let bestGamma: number[] =
Array(steps).fill(Math.PI / 4); Initial
assumption for gamma
     let bestScore =
objectiveFunction(bestBeta, bestGamma,
nodes, edges, idCircuitDraw);
     let bestMaxCutScore = bestScore;

     const maxIterations = 100; Maximum
number of iterations
     const randomStepScale = 0.01;
Scale of Random Change

     for (let iter = 0; iter <
maxIterations; iter++) {
       for (let i = 0; i < steps; i++)
{
         Const Newbeta = [...
Bestbeta];
         constant newgamma = [...
Bestgamma];
         // Changing the beta and gamma
values to a small random increment
         newBeta[i] += (Math.random() -
0.5) * randomStepScale;
         newGamma[i] += (Math.random()
- 0.5) * randomStepScale;

         const newScore =
objectiveFunction(newBeta, newGamma,
nodes, edges);

         if (newScore > bestScore) {
           bestBeta = newBeta;
           Bestgamma = Neugamma;
           bestScore = newScore;
           bestMaxCutScore = newScore;
         }
       }
     }

     return { beta: bestBeta, gamma:
bestGamma, score: bestScore,
maxCutScore: bestMaxCutScore };
   }
```

This implementation demonstrates the integration of the QAOA quantum algorithm with JavaScript, allowing the power of quantum computing to be harnessed on a web page or in a Node.js environment. The use of quantum computing in JavaScript contributes to the spread of quantum technologies among web developers and opens up new opportunities for optimizing complex problems.

5 CONNECTING THE FRAMEWORK TO A WEB APPLICATION

To implement the use of the algorithm, it was decided to build a web application using modern web development technologies [17]. The interaction of the application with the user (Figure 7) can be described as follows:

1) Open the page;
2) Enter graph edges;
3) Click the draw graph button;
4) Click the start calculation button;
5) Get the result.

The architecture of the application is built based on the following components[17]:

1) App.tsx: The main component of React, containing the core logic and rendering of the application;

2) Form.tsx: A form component that provides user input required to run the QAOA algorithm;
3) graph.ts: contains a Graph class that implements basic functionality for creating graphs, such as adding nodes, edges, and graph visualization;
4) useEdges.ts: A React hook for controlling the state of edges in a graph, provides functions for adding, removing, and updating edges;
5) main.tsx: The main login file for the application, which includes setting up and initializing the main components;
6) App.css: A style file for the add-on containing CSS rules for the main components.

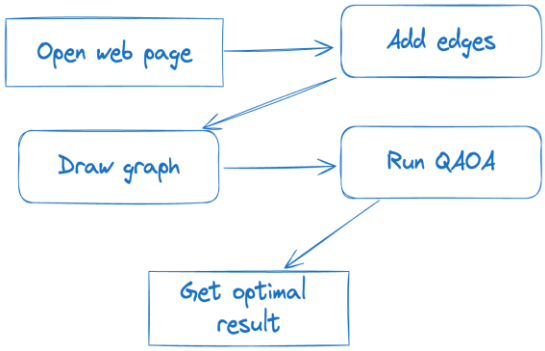


Figure 7: Scheme of the application.

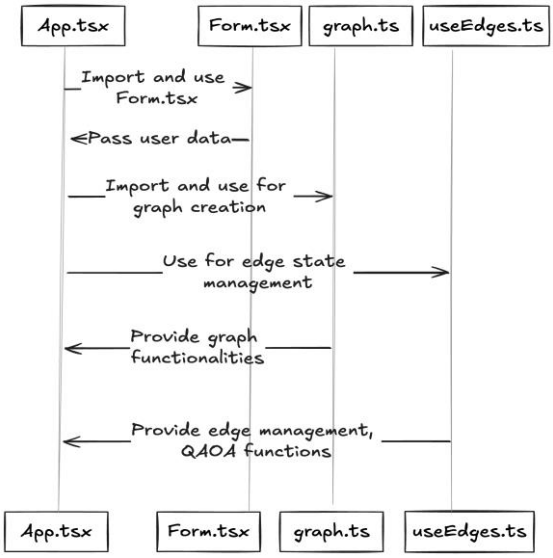


Figure 8: Application component interaction.

The interaction of the components is described in Figure 8. The components interact as follows [17]:

1) Initialization (main.tsx): The main.tsx file initializes the application by connecting the main component of App.tsx;
2) Main component (App.tsx): The App.tsx component imports and uses Form.tsx for user input, graph.js for graph visualization, and other helper components and hooks;
3) Data Entry Form (Form.tsx): Form.tsx is responsible for the user entering parameters to run the QAOA algorithm. This data is transmitted to App.tsx for further processing;
4) Graph (graph.ts): provides an implementation of the Graph class, which is used to create and manage graphs;
5) Edge Management and QAOA (useEdges.ts): The useEdges.ts Hook is used to manage the state of edges in graphs, providing functions for adding, removing, and updating edges, connecting the quantum.js library, and running QAOA algorithm calculations;
6) Styles (App.css): App.css contains CSS rules for the design of the main components of the application.

The Graphical Interface (Figure 9) of the appendix can be divided into 4 zones [17]:

1) Graph construction;
2) A form for constructing a graph;
3) Graph of quantum interaction between qubits;
4) Display of results.

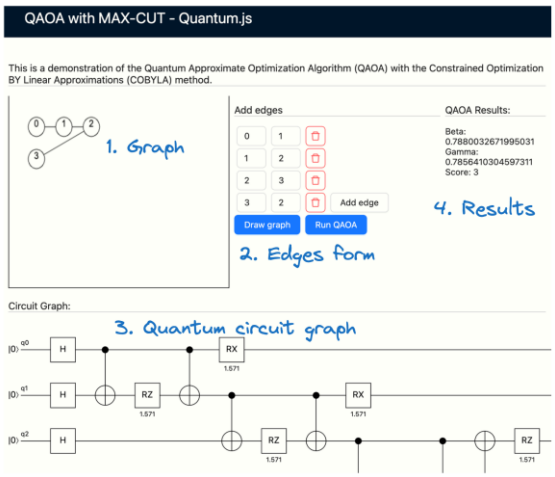


Figure 9: Application GUI.

6 CONCLUSIONS

As a result of the study, a quantum.js framework for high-level interaction with quantum bits was built and an implementation of the QAOA quantum hybrid algorithm was added to solve the Max-Cut problem.

By disabling the quantum.js framework to the web application, a graphical interface was built that allows you to: build and visualize graphs, use the QAOA algorithm in the browser, draw a quantum scheme of interaction between qubits used in the QAOA algorithm. For future improvement of the framework, the following points quantum.js highlighted:

- connect to quantum computers in the cloud;
- add other quantum-hybrid algorithms;
- build the functions of the framework for solving real problems (the task of optimizing resources, solving logistic problems).

To improve the performance of the web application, it is worth paying attention to the user's interaction with the application, improving styles, and adding validation for entering edge values.

The implementation demonstrated the possibility of effective use of quantum algorithms in the web environment, which contributes to the widespread use of quantum computing among web developers. The main achievements were the creation of the quantum.js framework, which provides high-level interaction with quantum bits and the integration of quantum bits computing with web technologies through a modern graphical interface. It has been shown that the QAOA quantum algorithm can significantly improve the optimization process for complex combinatorial problems compared to classical methods. The use of quantum computing in combination with classical optimization methods makes QAOA a powerful tool for solving NP-complex problems. Prospects for further development include connecting to quantum computers in the cloud, expanding documentation, adding new quantum-hybrid algorithms, and improving the user interface of the web application.

REFERENCES

[1] R. Feynman, [Online]. Available: https://en.wikipedia.org/wiki/Richard_Feynman.

[2] A. Almajid, "Summary of Real Quantum Mechanics," 2024. 10.13140/RG.2.2.20837.61928.

[3] “Google’s Sycamore: Exploring the Power of Google’s Quantum Computer,” [Online]. Available: https://medium.com/the-quantastic-journal/googles-sycamore-exploring-the-power-of-google-s-quantum-computer-266374339d5.

[4] “Qubit,” [Online]. Available: https://en.wikipedia.org/wiki/Qubit.

[5] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. Booth, and J. Tennyson, “The Variational Quantum Eigensolver: A review of methods and best practices,” Physics Reports, vol. 986, pp. 1-128, 2022. 10.1016/j.physrep.2022.08.003.

[6] H. Sahu and Dr. Gupta, "Quantum Computing Toolkit From Nuts and Bolts to Sack of Tools," 2023.

[7] “The Max-Cut Problem,” [Online]. Available: https://www.cs.cmu.edu/afs/cs/academic/class/15854-f05/www/scribe/lec02.pdf.

[8] “Pennylane datasets,” [Online]. Available: https://pennylane.ai/datasets/.

[9] “Qiskit,” [Online]. Available: https://docs.quantum.ibm.com/guides.

[10] “IBM Quantum Composer,” [Online]. Available: https://quantum.ibm.com/composer.

[11] “Solve utility-scale quantum optimization problems,” [Online]. Available: https://learning.quantum.ibm.com/tutorial/quantum-approximate-optimization-algorithm.

[12] V. Bergholm et al., "PennyLane: Automatic differentiation of hybrid quantum-classical computations," 2018. arXiv:1811.04968.

[13] “Intro to QAOA,” [Online]. Available: https://pennylane.ai/qml/demos/tutorial_qaoa_intro/.

[14] M. X. Goemans and D. P. Williamson, "Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming," Journal of the ACM, vol. 42, no. 6, pp. 1115-1145, 1995.

[15] E. Farhi, J. Goldstone, and S. Gutmann, "A Quantum Approximate Optimization Algorithm," 2014. arXiv:1411.4028.

[16] “Quantum.js Framework.” [Online]. Available: https://github.com/EarlOld/quantum.js.

[17] “QAOA-quantum,” [Online]. Available: https://github.com/EarlOld/QAOA-quantum.js.

Context-Defined Model of Open Systems Interaction for IoT Cybersecurity Issues Study

Victor Tikhonov[1], Eduard Simens[2], Yevhen Vasiliu[1], Valery Sitnikov[3], Abdullah Taher[4], Olena Tykhonova[1], Kateryna Shulakova[1,2] and Serhii Tikhonov[3]
[1]*Department of Information Technologies and Cybersecurity, State University of Intelligent Technologies and Telecommunications, Kuznechna Str. 1, 65023 Odesa, Ukraine*
[2]*Anhalt University of Applied Sciences, Bernburger Str. 57, 06366 Köthen, Germany*
[3]*Department of Computer Systems, Odesa Polytechnic National University, Shevchenko Avenue 1, 65044 Odesa, Ukraine*
[4]*Department of Electronic and Communication Engineering, College of Engineering, Al-Qadisiyah University, Babilon Str., 29, 58002 Al Diwaniyah, Iraq*
victor.tykhonov@gmail.com, olena.tykhonova@suitt.edu.ua, ye.vasiliu@gmail.com, sitnvs@gmail.com, abdallahqays@gmail.com, katejojo29@gmail.com, od.sergii.tihonov@gmail.com

Keywords: Internet of Things, Open Systems and Networks, General System Theory, Cybersecurity, Context-Defined Model.

Abstract: The article considers the issues of IoT systems and networks privacy. The complex researches on general system theory state of the art has been surveyed, as well as cybersecurity aspects of IoT technologies and models analyzed over the past decades with respect to the Industry 4.0 concept and industrial IoT architecture. A comprehensive decomposition formalism proposed on the base of M. Mesarowic system theory for presenting the complex cybersecurity object of study as a set of its parts. An original context-defined model of open systems interaction S-CDM has been constructed for cybersecurity issues study with the use of J. Neumann classification and A. Uemov system triad. The S-CDM model is shaped in 3-layer hierarchical graph. The first S-CDM layer displays the cybersecurity problem as a set of relations between the key players of information market – resources, clients and agents. The second S-CDM layer adds the network technologies impact on the relations between the key market-players. The third S-CDM layer includes a "degree of trust" factor into the key market-player relationships. The introduced context-defined model enables to reduce the overall problem complexity of the IoT cybersecurity study to a number of less complex and easy-handled partial tasks. As an example, a particular local version of the S-CDM model is considered in the form of six independent tasks for AI-aided methods of cybersecurity provision. The results of the work intend to contribute general system researches in the sphere of network system modelling and IoT privacy.

1 INTRODUCTION

Many challenges in modern society require a deep understanding of systems characterized by a large number of components, for which the constituent's concerns (atoms, cells, devices, individuals, organizations) are accompanied by the issues of their interactions study. Typical examples are nuclear physics, knowledge organization, global computer networks and Internet of things (IoT).

Holistic approach to such objects implies construction some relevant model of open system parts intercaction, dependent on researcher subject objectives. Because of that, different concurrent models of similar objects have right to exist and implement.

For instance, the Standard Model of particle physics provides a uniform theory to electromagnetic, weak, and strong interactions. Modern knowledge categorizing systems use three models: hierarchical, faceted, and enumerative. The contemporary Internet model has gone from initial OSI reference model and the TCP/IP protocol suite up to the ultimate Industrial Internet architecture framework.

One of the main goals of constructing a model of a complex object is to reduce the overall problem complexity when tearing it into several relatively independent and more simple tasks. This is not yet a problem itself solution, but the first step to get it.

An important large object of systematic study in recent years is the IoT architecture in the context of information privacy.

This work focuses the IoT privacy issues as a complex cybersecurity problem, and intends to find its comprehensive decomposition on the base of general system theory and Internet of things architecture analysis, in order to convert down the problem into a set of interconnected partial tasks of less complexity.

In the scope of this work, an object **O** model is considered as a subject' **S** personal point of view on the object of his study understood as a "closed-open substance" connected both to open subject **S** and open environment **W**. It means, that an abstract model of a real object is supposed to some extent subjective, while its relevance can be proven empirically.

So, an "open system" is defined as a subject-view function **S(O, W)**, where object **O** consist of a set of things $\{\mathbf{T_n}\}$ interacting to each other, or somehow with its open environment **W** through the subject **S**, and possibly directly with **W**.

The Internet itself is conventionally studied as a large open network formed by open sub-networks interacting via common protocols and interfaces. The global Internet project was created as a union of autonomous systems, networks and independent manufacturers of network equipment, without centralized system administration, built-in security mechanisms and real-time big data management.

Adaptation to new requirements often followed the way of "patching holes" when protocols` amendments and changes superimposed on their previous modifications, which ultimately led to cybersecurity issues due to unbridled Internet complexity increase and unclear consequences of AI rapid advances.

In the process of network technologies evolution, various reference models of Internet open systems interaction have been proposed, including the first well-known 7-layer OSI/ISO reference model and the ultimate 3-layer reference IoT model developed by Industrial Internet Consortium (IIC); these models reflect different visions of Internet architecture, and to one degree or another, continue to be used.

Section 2 of this work surveys researches on complex systems and networks.

Section 3 formulates motivations and objectives of the work.

Section 4 analyses the cybersecurity aspects in the IoT technologies and models with respect to "Industrial IoT" (IIoT) specifications as a key component of the Industry 4.0 concept.

Section 5 constructs a context defined model of open system interaction with the use of J. Neumann classification and A. Uemov system triad for the IoT cybersecurity issues study.

Sections 6 and 7 summarize the work's discussion and conclusions.

2 COMPLEX SYSTEMS AND NETWORKS RESEARCHES. STATE OF THE ART

Large-scale project development, such as those involving the Internet, often requires breaking down complex tasks into sequential stages. Systems, whether natural or artificial, typically follow a life cycle with stages like initiation, development, and conclusion. In system theory, this cyclic evolution is seen as a fundamental phenomenon where the end of one phase often leads to the beginning of another.

In contrast to biological systems, technical systems often evolve through incremental improvements, gradually adapting to new conditions or needs, which can eventually lead to increased complexity and potential instability. This process highlights the importance of structured design principles to manage the growth of these systems.

To ensure compatibility and coherence between various components of a project, it's crucial to establish a unified system architecture. This architecture formalizes the properties and interaction protocols of the different elements. Employing a systematic approach with a focus on modular design enables the decomposition of complex systems into simpler, interrelated components.

Modern system analysis methodologies, such as object-oriented tools and modelling techniques, support this structured approach. By focusing on phases like planning, analysis, design, and implementation, these methodologies aim to create new value through systematic development and optimization [1].

In his work "Systemology. General Theory of Systems" A.E. Kononyuk notes the importance of forming a professional language of systems analysis, including the choice of an adequate system of concepts and relevant terminology (names), considering the peculiarities of their intuitive interpretation in various natural languages (German, English, etc.). It is proposed to widely use terms in their established meaning, which can only be clarified; orientation towards the terminology of fundamental sciences (primarily mathematics); use, if possible, of international terminology [2].

The attribute "general" of system theory itself is clarified as "the most essential, characteristic of all

the foundations of something". Two main classes of problems identified: system analysis and synthesis. Emphasized, that rapid growth of information would be catastrophic for the further development of human cognition if no generalized knowledge occurred along with the expansion and deepening of sciences [2].

In recent years, systems analysts have increasingly focused the methods of big data processing, artificial intelligence (AI) and machine learning (ML) when using the systems theory as a math-superstructure to the theory of learning (both machine and human), connecting them with model-based system design practice. The set-theory seen as a formal framework for general system knowledge with the details of learning theory [3].

The enterprise model as an open system is considered by K. Lang & Co, aiming to create a framework for digital transformation: a system is self-contained; consists of subsystems, inputs, and outputs; and is in constant change. Systems arise through the determination of a barrier, which enables the distinction between the system and its environment. The elements inside a system, the organization comprising structures and processes, people and technologies or equipment all influence each other mutually, and the system only has meaning through the interaction of the elements. The system includes 6 items: business model, input data, supplier interface, company, services, client interface [4].

In October 2023, the USA President's Council (PCAST) published report "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence" which assessed the AI capabilities and its open human-environment to address the major global challenges. When supervised properly and responsibly by human experts, trained on high quality data, and verified using reliable scientific techniques, "AI tools can become engines of innovation that can supercharge the ability of scientists and policymakers to address such challenges". To achieve these goals, advanced models, data sets, and benchmarks need to be broadly available to the scientific community.

Though, AI can create new challenges (distilling errors and biases embedded in skewed training data, enormous energy for computation, the possibility that faulty science could be unwittingly generated, and the ease with which nefarious actors could use new powerful AI technologies for malicious purposes. The ideal future of AI-enabled science will require continued attention in three areas: empowerment of human scientists; responsible use of AI tools; sharing of basic AI resources. The experts formulated basic recommendations: expand efforts to secure access AI resources to federal data sets for approved critical research needs; support basic and applied AI researches across academia, industry, national labs and federal agencies; adopt responsible and transparent AI-use throughout all stages of the scientific research process; encourage innovative approaches to integrate AI-assistance into scientific workflows [5].

3 MOTIVATION AND OBJECTIVES OF THE WORK

In the modern era of Internet of Things (IoT) and global instability, information security became one of the most pressing issues, that includes various aspects of storing, accessing and transmitting large volumes of confidential data. It belongs to the objects of complex systems and networks study surveyed above in Section 2.

To date, solid results have been achieved in systems analysis and design. Based on G. Kron diakoptics, M. Mesarovic method and L. von Bertalanffy general systems theory (GTS), various models of system-decomposition can be constructed to reduce the initial object's complexity as a first step to approach a problem. However, this is not a trivial engineering task but somehow a creativity art.

One of the long-known principles of hierarchical decomposition and ordering of complex things is indexing classes of things by integers, e.g., Universal Decimal Classification (UDC). Though, the non-trivial matter is – which base of the number system is better to use (10, 8, etc.), and how to define semantic classes at each level of hierarchy.

J. Neumann gave a rigorous answer to the first question: ternary integer numbering for abstract things ordering is optimal in terms of information capacity. The Mesarovic method lets formally obtain relations sets on the given classes of things in accordance with A. Uemov system triad "Things, properties, relations".

Thus, the core non-trivial task in complex object decomposition is the relevant definition of interconnected semantic classes of things with respect to given problem study. This motivation idea underlies our objectives and discussions in this work.

The work aims system decomposition of the complex IoT cyber security problem in multiple fewer complex tasks as a first step to approach the problem.

To achieve this goal, the following objectives are set in the work:

1) Analysis the cyber security aspects of basic IoT technologies and models with particular focus on the USA NIST SP 800-183 specification and Industrial IoT Connectivity Framework.

2) Construction a context-defined model of open systems interaction for the Internet of things cyber security problem study on the base of J. Neumann ternary classification and A. Uemov system triad.
3) Interpretation the obtained context-defined model in a particular case-study of the artificial intelligence-based cybersecurity provision.

4 CYBERSECURITY ASPECTS IN THE IOT TECHNOLOGIES AND MODELS

The problem of cybersecurity is largely due to the rapid development of sensor network technologies and the growth of the IoT segment in the overall Internet infrastructure [6].

Germany played a leading role in foundation of basic IoT technologies and development the concept of future industry – "Industry 4.0". In 1980s began the rise of classic Fieldbus technology with proprietary serial protocols from some German manufacturers, aiming to reduce the installation and maintenance effort of industrial machines and systems. It was an industrial network system for real-time distributed control to connect instruments in a manufacturing plant. A Fieldbus works on networks of diverse topologies (daisy-chain, star, ring, branch, tree etc.). In 1987, the Process Field Bus (ProfiBus) was launched through a publicly funded project by companies and institutes [7].

Based on these advances, one of the first IoT technologies M-Bus was developed in 1991 by H. Ziegler from Paderborn Univ., in coop with Texas Instruments GmbH and Techem GmbH for remote reading of utility meters (water, gas, electricity and heat). M-Bus was initially standardized in EN1434 for thermal energy meters and then defined in EN13757x specifications family issued by the EC committee CEN since 2002: EN 13757-1 (2002-2022) – generic descriptions for communication systems; EN 13757-2 (2004-2018) – wired OSI PHY and DL layers; EN 13757-3 (2005-2018) – dedicated OSI APP layer. The PHY layer supports galvanic-type asynchronous interface on 2-core copper cable. In 2003 the wireless M-Bus was deployed (EN 13757-4, 2008-2019) [8].

In 1998, the Connectivity Standards Alliance (CSA, formerly Zigbee Alliance), after the Bluetooth (1994) and Wi-Fi (1997) technologies emerged, created ZigBee-technology for automated data collection from IoT devices around 10-100 m at ISM (Industrial, Scientific, Medical) radio frequencies 900/868 MHz and 2.4 GHz. It has low power consumption and data transfer rate (20-250 Kbit/s), but a high level of data protection (AES with 128-bit symmetric key); still, it took off only in 2002 when more interested companies joined the Zigbee Alliance, [9]. Related standard IEEE 802.15.4 was firstly adopted in 2003 and ratified by ZigBee in 2004; then it was revised to Zigbee_Light_Link library (2006) and ZigBee Pro/2007 revision [10].

In 2012, the *first* IoT reference 4-layer model of ITU (Y.400/Y.2060) was issued [11]. In 2014, the *second* reference 7-layer IoT architecture was adopted by the IoT World forum [12], [13].

In June 2016, the ISO specification 20922 was published for Message Queuing Telemetry Transport protocol (MQTT), that has become the de-facto standard for Industry 4.0 concept (manufacturing modernization projects and digital transformation of the field, real-time decision making, enhanced productivity, flexibility and agility) [14]. In July 2016, the *third* IoT reference 4-layer model was issued by the USA National Institute of Standards and Technologies (NIST SP 800-183) with particular accent on the IoT security issues [15].

The MQTT standard, along with the NIST SP 800-183 IoT-architecture (2016), opened a new stage in the development of IoT technologies known as "Industrial IoT" (IIoT) – a key component of the Industry 4.0 concept. The history of the IIoT begins with the invention of the programmable logic controller (PLC) by R. Morley (1968), which was used by General Motors in automatic transmission manufacturing. These PLCs allowed for fine control of individual elements in the manufacturing chain.

The term "IIoT" was coined in the USA, it covers parts of the overall Industry 4.0 concept ([7]). Now, the "IIoT" means extension the IoT paradigm into industrial settings and applications via internet IP-connectivity to integrate advanced sensors, software, and machinery aimed for collecting, analyzing and processing upon vast amounts of data. This data-driven approach enables real-time decision-making and predictive analytics, leading to enhanced operational efficiency, reduced costs, and improved product quality [17]. In accordance with that, in 2017, the 802.15.4 ZigBee standard was extended up to over IP running version by the Dotdot (//:) library ([18], [19]); its current revision (2020) supports wireless personal area network (WPAN) of ad-hock mesh topology [20].

In 2022, the Industry IoT Cons. (aka Industry Internet Cons. IIC™, now integrated into the Digital Twin Cons. DTC™) published "The Industrial IoT Connectivity Framework" (IIC-model), where the IIoT interoperability presented by 3 layers (technical

end devices, syntactic IIoT platform, semantic human/AI users), [21], [22]. This document presents the currently *ultimate (number four)* 3-layered reference model of the IoT. One of the popular today IIoT development platform is a virtual EMQX broker, that was designed by EMQ company; it supports protocols MQTT (3.1, 3.1.1, and 5.0), HTTP, QUIC, and WebSocket for up to 100 million concurrent IoT devices per cluster with 1 million messages per second and a millisecond latency [23].

Consider the cybersecurity context of IoT, take in account the following related aspects. 1) The IoT networks operate vast amounts of sensitive data, ranging from personal information to critical business data. 2) Many IoT-segments are imbedded in critical infrastructure like power grids, water supply and transportation. 3) The IoT systems widely use vulnerable wireless links [24]. E.g., ZigBee standard permits re-use of link keys for re-joining the network; so an attacker can clone the legitimate device and spoof the network layer of Trust Centre by pretending to be previously connected device wanted to re-join the network [25]. 4) Recently, AI has played an increasingly important role in Big Data processing around the IoT segments; on the other hand, this opens up negative opportunities for malicious use of new AI-researches.

The above ITU model of IoT (2012) defines 4 layers: devices, network transport, service platform and applications. Two kinds of security capabilities declared: generic ones (authorization, authentication, data integrity protection, access control etc.) and IoT-specific options (without concrete details).

The next spoken above IoT/WF model (2014) defines 7 layers: (IoT devices, connectivity, edge cloud computing, data accumulation, data aggregation, applications, business processes). Those years exhibited an extensive growth of IoT segment, so, this model had primarily warred about technological aspects of Big Data management. The IoT privacy had not attracted enough attention then; as a result, serious problems arose.

To address emerging IoT privacy concerns, NIST introduced an IoT reference model (SP 800-183) in 2016. It defines 4 abstract layers in the context of privacy provision: sensor clusters with individual impact weights d_i; aggregators that implement summation Sum(d_i); communication channels (e.g., USB, wireless, wired, verbal etc.) incl. those offering as a service (AAS); external AAS-utilities (eUtilities) where a human may be viewed as an eUtility-service.

Data supplied by an eUtility can be weighted. Non-human eUtility may have device ID, which can be crucial for identification/authentication. The key role in this model play the so-called "decision-making triggers" (DTs) acting at any layer and formalizing the primary goal of the IoT on the upper layer (eUtility). The overall IoT network framework upon this model is similar to a spatially distributed artificial neural network. Herewith, the complex system problem of IoT cybersecurity is functionally decomposed on four distinct tasks of reduced complexity according to defined 4 layers of the model.

The 3-tier IIC-model (2022) spoken above addresses the “Industrial Internet Reference Architecture” IIC-IIRA [26] and “Industrial Internet Security Framework” IIC-IISF [27]; these three documents currently most fully and systematically describe the overall IIoT-architecture and the general concept of cybersecurity in modern digital world.

The IIC-model follows OSI/ISO and TCP/IP: OSI layers L1÷L4 have been merged in a transport **T** tier for delivering TCP/UDP messages (technical interoperability **TchIO**); layers L5÷L7 merged/split for data framework **F** (files of state or events formed by messages – syntactic interoperability **SinIO**); document **D** – the context-interpretation of **F** (Semantic interoperability **SemIO**).

TchIO means information exchange (bits/bytes, e.g. pencil scribbles), assuming that info-exchange structure (e.g. pencil and paper) established. SinIO means info-exchange in a common data structure (e.g., in a known language grammar). SemIO is unambiguous context-appropriate interpretation of exchanged data.

The IIC-IIRA (2022) comprises 3 tiers: edge (IoT access network), Internet-based service platform and enterprise API/human ([25], page 44). It is close to previously discussed 4-tier NIST SP 800-183 model, if communications combined with service platform. This breaks a typical IoT-domain in 3 interacting parts with 3 individual responsibilities (edge, platform, API) and 3 interoperability ones: “edge-platform”, “edge API”, “platform-API”; these responsibilities include security aspects (both IT- and OT-types).

The IIC-IISF (2016) provides guidance for trustworthy systems, security characteristics, technologies and techniques to be applied, and how to gain assurance that the appropriate mix of issues have been addressed to meet stakeholders' expectations. The core term “Trustworthiness of an IIoT System” is defined as a unit of 5 entities (resilience, reliability, safety, privacy, security) which are exposed to 4 external influences: environment disruption, system faults, attacks, human errors ([25], p. 23).

Key factors of IIoT security are focused therein: convergence of IT and operational technology OT (p.

24); evolution of IT and OT security ([25], p. 25); brownfield (new + legacy) deployments in OT (p. 26); in a typical IIoT system, cloud computing is supposed a critical point of vulnerability as using a shared third-party service-providers creates a number of trust boundaries affecting security and privacy ([25], p. 27); functional breakdown noted for security model and policy ([25], p. 59).

5 CONSTRUCTION A CONTEXT DEFINED MODEL FOR CYBERSECURITY STUDY

The Internet has been around for about half a century, and as its technologies have evolved, various reputable organizations have introduced at least 7 different functional models of the Internet open systems interaction, starting with OSI reference model and TCP/IP protocol suite in late 1970s.

It is important to note that each newly emerging model of network architecture did not reject the previous ones, but rather complemented and developed them in the context of new specific objectives. Therefore, all these models, one way or another, continue to be used by specialists. The seven-layer OSI/ISO reference model defines commonly popular professional terms of an abstract open system, while the four-layer TCP/IP stack formulates detail interfaces of real network objects interaction. The three-layer NGN/ITU model maps any open packet-based network (incl. entire Internet) in a pretty simple view: transport infrastructure (lower level) and application domain (upper level), which counteract via IP protocol (middle level).

The last four models (2012, 2014, 2016, 2022) have been developed for IoT-networks architecture; among them, NIST SP 800-183 (2016) model was a response to the cybersecurity challenge, and the ultimate IIC-model (2022) together with IIC-IIRA and IIC-IISF, form a holistic vision of the modern Industry 4.0 concept.

Thus, it is concluded, that a holistic approach to a complex problem needs construction a specific model of this problem in the context of the objectives posed by the researcher.

Let's introduce a context defined model (CDM) of open system interaction between clients **C**, resources **R** and agents **A** via network **N**, Figure 1.

The object **R** in Figure 1 means a set of *information resources* which are commonly accessible on the market; the object **C** symbolizes a set of *clients* of the info-resources market; **A** is a set of *broker-agents* on the info-resources market.

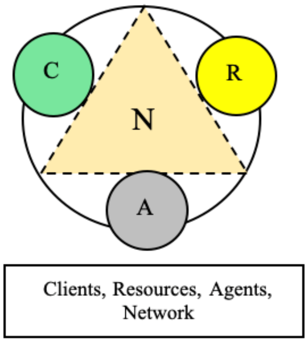


Figure 1: "Client-Resource-Agent" context defined model of open system interaction (CDM).

Objects (**R**, **C**, **A**) may interact directly in pairs (**R**,**C**), (**R**,**A**), (**C**,**A**) or/and via the *network* communication media **N**. Each of the objects (**R**, **C**, **A**) is a unit of two subsets: $\mathbf{R}=(\mathbf{R_T}, \mathbf{R_F})$; $\mathbf{C}=(\mathbf{C_T}, \mathbf{C_F})$; $\mathbf{A}=(\mathbf{A_T}, \mathbf{A_F})$, where ($\mathbf{R_T}$, $\mathbf{C_T}$, $\mathbf{A_T}$) are authorized (*true*) objects; ($\mathbf{R_F}$, $\mathbf{C_F}$, $\mathbf{A_F}$) – unauthorized (*fake*) objects.

At the first two steps of security-problem decomposition, a set **S** of six relationships can be constructed:

$$\mathbf{S}:=\{\mathbf{S_i}\} = \{[(\mathbf{R},\mathbf{C})_1, (\mathbf{R},\mathbf{A})_2, (\mathbf{C},\mathbf{A})_3], [(\mathbf{R},\mathbf{C})_{N4}, (\mathbf{R},\mathbf{A})_{N5}, (\mathbf{C},\mathbf{A})_{N6}]\}, \quad (1)$$

where relationships like (x, y) in (1) may be defined either commutative (symmetric) or non-commutative (asymmetric). The symmetric case of relationships **S** in (1) presents the adjacency matrix graph in Figure2.

S	R	C	A	N
R		1	1	1
C	1		1	1
A	1	1		1
N	1	1	1	

Figure 2: Symmetric adjacency matrix graph of S.

At the third step, the following commutative or noncommutative Cartesian products can be defined:

$$\begin{cases} (\mathbf{R},\mathbf{C})_1 \rightarrow (\mathbf{R_T},\mathbf{C_T})\times(\mathbf{R_F},\mathbf{C_F}); \\ \dots\dots\dots\dots \\ (\mathbf{C},\mathbf{A})_3 \rightarrow (\mathbf{C_T},\mathbf{A_T})\times(\mathbf{C_F},\mathbf{A_F}); \\ (\mathbf{R},\mathbf{C})_{N4} \rightarrow (\mathbf{R_T},\mathbf{C_T})_N \times(\mathbf{R_F},\mathbf{C_F})_N; \\ \dots\dots\dots\dots \\ (\mathbf{C},\mathbf{A})_{N6} \rightarrow (\mathbf{C_T},\mathbf{A_T})_N \times(\mathbf{C_F},\mathbf{A_F})_N. \end{cases} \quad (2)$$

The bipartite graph of a commutative Cartesian product for $(\mathbf{R},\mathbf{C})_1$ for direct relationships member in (2) presents Figure 3.

S_1	R_F	C_F
R_T	(R_T, R_F)	(R_T, C_F)
C_T	(C_T, R_F)	(C_T, C_F)

Figure 3: Bipartite graph of Cartesian product for $(R,C)_1$.

In the scope of this work, the CDM model (Fig. 1) along with decomposition formalisms (1) and (2) is accepted as the *cyber Security Context-Defined Model* (S-CDM).

The security context-defined model S-CDM can be presented by a 3-layer hierarchical graph shown in Figure 4.

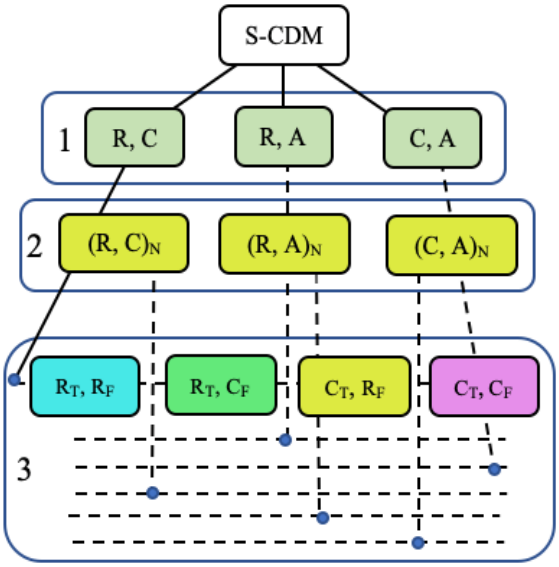


Figure 4: The graph of security context-defined model.

The *first layer* in Figure 4 (denote it S-CDM1) displays the cybersecurity problem decomposition as a set of three types of relationships between the three key objects: **R** (information resources on the market), **C** (customers of the information resources), **A** (intermediary agents, or brokers, on the information resources market). Here the S-CDM1 is understood in two different aspects: a) in a *narrow sense*, as a system of regulatory and legal relations between three key players on the information market (**R**, **C**, **A**) – with a direct mechanism of interaction (ignoring the network issues, illegal or criminal objects impact, hacker attacks, etc.); b) in a *broad sense* (considering the influence of various counteracting or harmful factors).

The *second layer* in Figure 4 (S-CDM2) adds network media (N) as a significant operational factor in relationships between the key players (R, C, A) on the information market. Joint accounting of S-CDM1 and S-CDM2 is considered next as a partial extension of the S-CDM1 model (in its narrow sense) with respect to cybersecurity problem. Herewith, the S-CDM2 layer splits each of the S-CDM1 tasks in two related distinct tasks which consider or not the network interaction impact.

The *third layer* in Figure 4 (S-CDM3) includes the logical category "degree of trust" into 6 relations of layer 2 between the key players (**R**, **C**, **A**) on the information market.

Thus, each of the six S-CDM2 layer relations splits into 4 classes at the S-CDM3 layer according to (2) and Figure 3. The four of S-CDM3 classes split by $\mathbf{S_1}=(\mathbf{R},\mathbf{C})_1$ are explicitly shown in Figure 4 as $\{(R_T, R_F), (R_T, C_F), (C_T, R_F), (C_T, C_F)\}$; the rest five ones are outlined implicitly. As a result, the context-defined model S-CDM (Fig. 4) of cybersecurity problem decomposition concludes 24 elementary tasks at its bottom layer.

The introduced above S-CDM-model for hierarchical decomposition of relationships between the key information market players (**R**, **C**, **A**) enables setting various cybersecurity relevant tasks of systems analysis and synthesis by presentation the S-CDM model in different structural bases. Below, two examples of such structural bases are exhibited.

Basis 1. Regulatory/operational cybersecurity. The cybersecurity problem is presented by three orthogonal structural primitive branches (Fig. 5). There are three independent tasks that can be raised up on this S-CDM model; each of them includes two nested sub-tasks. This basis is convenient for studying the cybersecurity problem ignoring the third parties aggressive influence.

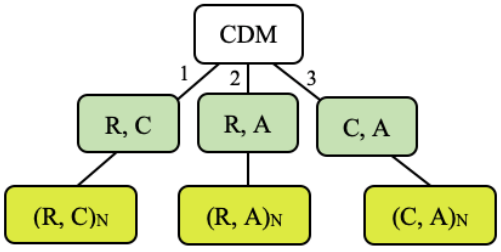


Figure 5: Regulatory/operational cybersecurity model.

Basis 2. Fake-agents on the info-market.

The cybersecurity problem is presented by two orthogonal primitive branches – independent tasks (Fig. 6); each of them includes four nested sub-tasks.

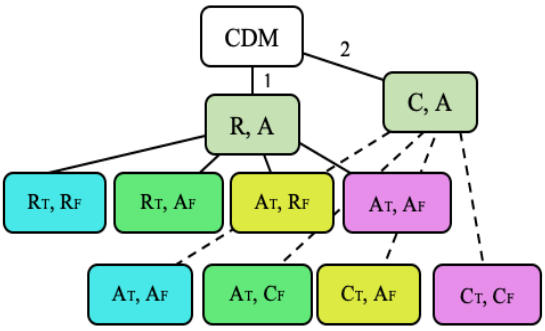


Figure 6: Fake-agents on the info-market model.

The S-CDM-approach can be used to approach the cumbersome issue of AI-application in the IoT cybersecurity provision through decomposition the object of study into less complex partial tasks.

Let "**( )**" be common regulatory/operational (RO) tools; "**( )$_{AI}$**" – particular AI-aided RO-tools; "**ID**" – *identifications*; "**AU**" – *authorizations*; "**VR**" – *verifications* (e.g. authentications).

Let define a local cybersecurity model:

$$\text{S-CDM}_{CS} := [\mathbf{(\,)}, \mathbf{(\,)}_{AI}] \times [\mathbf{C}, \mathbf{R}] \times [\mathbf{ID}, \mathbf{AU}, \mathbf{VR}]. \quad (3)$$

It results in 2×2×3=12 partial relations; the first 3 of them (**C**,**ID**), (**C**,**AU**), (**C**,**VR**) refer to common client identification/authorization/authentication methods; next three ones extend the **ID**, **AU**, **VR** notions into the sphere of info-resources management.

The last 6 members of product (3) define a local AI-aided cybersecurity model in the form of six independent tasks

$$\text{S-CDM}_{AI\text{-}CS} := \mathbf{(\,)}_{AI} \times [\mathbf{C}, \mathbf{R}] \times [\mathbf{ID}, \mathbf{AU}, \mathbf{VR}]. \quad (4)$$

The Cartesian-product member [**( )$_{AI}$**, **C**, **VR**] in (4) defines a cluster "AI-aided methods and operational tools for client verification (authentication)", which can be taken as a next step of researches.

Consider a comprehensive AI-to-cybersecurity survey [27] (2023) with 247 references. In includes 6 "identity-authentication" AI-related items: "Bio-signal classification for human identification based on convolutional neural networks"; "Improving the security and QoE in mobile devices through an intelligent and adaptive continuous authentication system"; "Securing smart offices through an intelligent and multi-device continuous authentication system"; "An approach to detect user behavior anomalies within identity federations"; "Web user authentication using chosen word keystroke dynamics"; "Keystroke identifier using fuzzy logic to increase password security".

Thus, it is possible to identify a compact area of knowledge and related publications for systemic research as part of a larger problem. Herewith, specific relevant facts can be fixed, e.g.: human identification systems use signal acquisition, signal pre-processing, and feature extraction/classification extracted from the images based on convolutional neural networks (ref. 90); permanent authentication of mobile devices owner provides user profile generation and real-time comparison of the current mobile usage with typical user's behavior using by ML techniques (ref. 91).

6 DISCUSSION

The 20th century evidences formation a new fundamental vision of things and problems around us. Many challenges require a deep insight on complex objects study with a large number of components, for which the constituent's issues are accompanied by their interaction's concerns.

An efficient instrument of general system theory (GST) is hierarchical decomposition a given large object of study into a set of interconnected partial tasks. But, implementation of this instrument is somehow a creativity art rather than a trivial engineering task. In particular, a difficult task is relevant definition of interconnected semantic classes of things with respect to given problem study.

A critical problem today is Internet of things systems and networks privacy. In spite of significant progress in both wired and wireless IoT-technologies, new tasks and challenges emerge, e.g. in the context of advanced not clear predicted artificial intelligence capabilities.

This work focuses the IoT privacy object of study as a complex cybersecurity problem. As a first approach to this problem solution, a GST-decomposition formalism proposed, that converts down the overall cybersecurity object into a set of 24 interacting partial tasks of less complexity. The objects components are clients, resources and agents of the information market, along with network communication means.

Within the scope of this work, an original cyber security context-defined model S-CDM of open systems interaction constructed on the base of J. Neumann ternary classification and A. Uemov system triad for the IoT privacy issues study.

The S-CDM-model is shaped in 3-layer hierarchical graph. The S-CDM1 layer displays the cybersecurity problem as a set of relationships between the clients, resources and agents.

The S-CDM2 layer adds network technologies impact on the relations between the key market-players.

The S-CDM3 layer includes the “degree of trust” factor in relationships between the key market-players.

The introduced S-CDM model provides decomposition of a complex cybersecurity problem to a number of less complex and easy-handled partial tasks.

7 CONCLUSIONS

This article addresses the problem of cybersecurity within the IoT architecture, focusing on privacy protection amidst the challenges of digital transformation and rapid advancements in artificial intelligence. The analysis revealed that IoT is characterized by a complex structure requiring a systematic approach to solve issues related to data protection, identification, authentication, and authorization.

The primary method proposed involves decomposing a complex problem into a number of interconnected sub-problems on the base of general systems theory. A novel context-defined model of open system interaction S-CDM was developed to structure the IoT cybersecurity problem into less complex tasks. This model provides a foundation for building localized and AI-enhanced models aimed at the identification, verification, and authorization of clients and resources.

The research also demonstrated that artificial intelligence plays a pivotal role in securing IoT through mechanisms such as continuous authentication, biometric signal classification, and anomaly detection in behavior. Special emphasis was placed on integrating IoT into industrial systems (IIoT) within the framework of Industry 4.0.

The results of the work intend to contribute general system researches in the sphere of network system modelling and IoT privacy.

ACKNOWLEDGMENTS

We acknowledge support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Open Access Publishing Fund of Anhalt University of Applied Sciences.

REFERENCES

[1] A. Dennis et al. “System analysis and design”, 5th ed., John Wiley & Sons, 2012, 563 p. [Online]. Available: https://www.google.com/search?client=safari&rls=en&q=Dennis+A.+Wixom+B.H.+System+Analysis+and+Design+5+Ed++pdf&ie=UTF-8&oe=UTF-8.

[2] A.E. Kononuk "Systemology. General System Theory", Book 1. Basics, Kiev, 2014, 568 p. [Online]. Available: https://z-lib.io/book/16404592f.

[3] A. Dennis et al., System Analysis and Design, 5th ed., John Wiley & Sons, 2012, 563 p. [Online]. Available: https://www.google.com/search?client=safari&rls=en&q=Dennis+A.+Wixom+B.H.+System+Analysis+and+Design+5+Ed++pdf&ie=UTF-8&oe=UTF-8.

[4] A. E. Kononuk, Systemology. General System Theory, Book 1. Basics, Kiev, 2014, 568 p. [Online]. Available: https://z-lib.io/book/16404592f.

[5] T. Cody et al., “Motivating a Systems Theory of AI,” ResearchGate Special Feature, vol. 23, no. 1, pp. 37-40, 2020. [Online]. Available: https://www.researchgate.net/publication/340652980_Motivating_a_Systems_Theory_of_AI.

[6] K. Lang et al., “A systems theory-based conceptual framework for holistic digital transformation,” 2021, 24 p. [Online]. Available: https://www.researchgate.net/publication/352020735_A_systems_theory-based_conceptual_framework_for_holistic_digital_transformation.

[7] Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, President’s Council of Advisors on Science and Technology, 2023, 36 p. [Online]. Available: https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence.

[8] M. Abomhara and M. Koien, “Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks,” Journal of Cyber Security, vol. 4, pp. 65–88, 2015. [Online]. Available: https://www.researchgate.net/publication/349642687_Cyber_Security_and_the_Internet_of_Things_Vulnerabilities_Threats_Intruders_and_Attacks.

[9] M-Bus und M-Bus Wireless Grundlagen. [Online]. Available: https://www.wachendorff-prozesstechnik.de/technologie/feldbus/m-bus/.

[10] H. Ziegler, Serial BUS Systems. [Online]. Available: https://m-bus.com/assets/downloads/MBDOC48.PDF.

[11] S. Payne, “Zigbee Technology: Everything You Need to Know for IoT,” 2023. [Online]. Available: https://www.telit.com/blog/zigbee-iot-guide/.

[12] ZigBee PRO/2007 Layer PICS and Stack Profiles, Rev. 08, 2018. [Online]. Available: https://infocenter.nordicsemi.com/pdf/Zigbee_PRO_Layer_PICS_and_Stack_Profile_8.pdf.

[13] “Y.2060: Overview of the Internet of Things,” 2012. [Online]. Available: https://www.itu.int/rec/T-REC-Y.2060-201206-I.

[14] “IoT Reference Model Introduced at IoT World Forum 2014.” [Online]. Available: https://ytd2525.wordpress.com/2014/10/22/iot-reference-model-introduced-at-iot-world-forum-2014/.

[15] "IoT World Forum Reference Model," 2014. [Online]. Available: https://www.researchgate.net/figure/IoT-World-Forum-Reference-Model_fig2_323525875.

[16] "What is Industry 4.0?". [Online]. Available: https://www.ibm.com/topics/industry-4-0/.

[17] J. Voas, Networks of Things, NIST Special Publication 800-183, 2016. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-183.pdf.

[18] "IIoT Explained: Examples, Technologies, Benefits and Challenges." [Online]. Available: https://www.emqx.com/en/blog/iiot-explained-examples-technologies-benefits-and-challenges.

[19] "Dotdot was a huge hit at CES 2017." [Online]. Available: https://www.symmetryelectronics.com/blog/dotdot-was-a-huge-hit-at-ces-2017/.

[20] "What is Dotdot?". [Online]. Available: https://en.dsr-corporation.com/news/what-is-dotdot.

[21] IEEE 802.15.4 Standard for Low-Rate Wireless Networks, 2020. [Online]. Available: https://store.accuristech.com/standards/ieee-802-15-4-2020.

[22] R. Joshi et al., The Industrial Internet of Things Connectivity Framework, 2022. [Online]. Available: https://www.iiconsortium.org/wp-content/uploads/sites/2/2022/06/IIoT-Connectivity-Framework-2022-06-08.pdf.

[23] C. Baudoin et al., "Global Industry Standards for Industrial IoT," 2022. [Online]. Available: https://www.iiconsortium.org/pdf/IIC_Global_Standards_Strategy_Whitepaper.pdf.

[24] "EMQX Overview." [Online]. Available: https://docs.emqx.com/en/emqx/latest/.

[25] M. A. Razzaq et al., "Security Issues in the Internet of Things (IoT): A Comprehensive Study," International Journal of Advanced Computer Science and Applications, vol. 8, no. 6, pp. 383-388, 2017.

[26] "Zigbee Security 101. Architecture And Security Issues," 2023. [Online]. Available: https://payatu.com/blog/zigbee-security-101-architecture-and-security-issues/.

[27] The Industrial Internet Reference Architecture, 2022. [Online]. Available: https://www.iiconsortium.org/wp-content/uploads/sites/2/2022/11/IIRA-v1.10.pdf.

[28] Industrial Internet of Things. Volume G4: Security Framework, 2016. [Online]. Available: https://www.iiconsortium.org/pdf/IIC_PUB_G4_V1.00_PB.pdf.

[29] R. Kaur, D. Gabrijelčič, and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," Elsevier, Information Fusion, vol. 97, 2023.

Digital Transformation of Linguistic Pedagogy

Tetiana Hromko[1], Oksana Vasylenko[2], Nataliya Chepelyuk[1] and Yevheniia Hromko[1]
[1] *Department of English Philology and Translation, Odesa Polytechnic National University, Shevchenko Avenue 1, 65044 Odesa, Ukraine*
[2]*Departments of Electrical Engineering, Mechanical Engineering and Engineering Management, Anhalt University of Applied Sciences, Bernburger Str. 57, 06366 Köthen, Germany*
hromkotv@gmail.com, oksana.vasylenko@hs-anhalt.de, natalach@gmail.com, yevheniiahromko@gmail.com

Keywords: Interactive Application, Gamification, Personalized Education, Language Acquisition, User Engagement, Social Integration, Data Protection.

Abstract: The article investigates technical and linguistic innovations in interactive applications for learning English, including the implementation of artificial intelligence, natural language processing (NLP) and gamification. The development of an innovative interactive English language learning application, EnApp, which merges linguistic innovation with technological advancements to enhance the learning experience, is presented in the article. EnApp employs a comprehensive approach, combining personalized learning, gamification, and multimedia elements to address the challenges of modern language acquisition. The application incorporates interactive dictionaries, thematic units, and multimedia content to support vocabulary, grammar, and listening comprehension. The study utilizes a variety of research methods, including user needs analysis, comparative analysis, and gamification techniques. The application's architecture, based on Unity and C#, ensures a seamless user experience. Key features include user registration, dictionary management, interactive exercises, and a progress tracking system. Through extensive testing and feedback analysis, EnApp has demonstrated its effectiveness in improving English language skills. Users reported increased motivation, engagement, and proficiency in vocabulary, grammar, and pronunciation. The application's adaptive learning algorithms and personalized feedback mechanisms contribute to its success. By integrating linguistic innovation with technological solutions, it provides a flexible, engaging, and effective tool for learners of all ages and abilities.

1 INTRODUCTION

In today's world, English is increasingly recognized as a universal medium of communication, crucial for accessing global information and cultural resources. Its role extends beyond international discourse to impact daily life for millions [1, 2]. Learning English has become integral to education, aiding individuals' integration into a global knowledge society [3].

A key aspect of contemporary English acquisition is the synergy of linguistic and technological skills. Research indicates that interactive educational applications enhance motivation and efficiency in language learning through gamification and personalized methods [4, 5]. These tools adapt to user needs, highlighting the necessity of combining linguistic and technical expertise.

The scientific merit of this work lies in its binary approach to developing English learning applications, merging linguistic learning with technical solutions. This innovative strategy addresses modern educational demands for accessibility, personalization, and interactivity. The growing emphasis on linguistic and technological competencies for language educators further supports integrated development approaches [6,7].

Traditional English teaching often encounters challenges such as low motivation and monotonous exercises, which hinder language acquisition. Interactive applications leveraging game mechanics, AI, and multimedia can effectively enhance engagement and overcome these obstacles. They offer accessibility, adaptability, and personalization, boosting learning outcomes for diverse proficiency levels. Such tools facilitate language skill development through tailored, multichannel information delivery, making them essential in modern education.

The study aims to create an interactive application EnApp that fuses innovative linguistic and technological strategies to enhance English

learning. Key objectives include developing a flexible, personalized system responsive to users' knowledge levels and needs, while incorporating gamification and multimedia elements. Artificial intelligence will be utilized for task selection and progress monitoring. Ultimately, the application will support users of all ages in overcoming traditional learning challenges and improving communication skills.

2 RELEVANCE AND ANALYSIS OF THE TOPIC AREA

Knowledge of English is crucial for Ukraine's economic and socio-cultural development amid globalization. According to the EF English Proficiency Index (EF EPI), Ukraine ranks in the middle of the global proficiency rankings, trailing behind most Western European nations. Challenges in learning English in Ukraine include regional disparities in knowledge, limited practice opportunities, and low motivation [8].

Users of modern language apps increasingly demand flexible, accessible, and personalized learning experiences tailored to their goals. The availability of apps for on-demand learning is vital for those with busy schedules or limited access to traditional courses. Additionally, interactive elements enhance engagement and provide essential feedback, helping to reduce barriers to language learning.

Effective English language learning applications must address various linguistic requirements, including vocabulary, grammar, listening, pronunciation, and contextual awareness. Vocabulary learning requires multi-level approaches, while grammar instruction should be systematic and integrated into practical tasks. Interactive features like voice recognition can support pronunciation and listening skills.

A review of applications like Duolingo, Quizlet, and My Vocabulary reveals functional gaps, particularly in developing integrated skills such as listening and speaking, and in offering personalized options for diverse user levels. Integrating advanced multimedia components can enhance user experience and meet the needs of today's learners.

Duolingo provides basic language skills through interactive exercises and a rewards system. While its game format boosts motivation, it is limited in addressing complex grammar and is best suited for beginners [9]. Quizlet enables users to create flashcards and engage in various active learning modes, proving effective for vocabulary acquisition, but lacking in support for listening and speaking skills [10, 11]. My Vocabulary allows users to curate their vocabulary but lacks gamification and interactive features, limiting its effectiveness for multimedia learning.

3 METHODOLOGY

The main purpose of an interactive English language learning – EnApp is to create an effective learning environment that combines linguistic innovation with personalization and gamification. This approach not only increases user engagement, but also stimulates the development of comprehensive language skills, including vocabulary, pronunciation, and listening comprehension.

3.1 Personalization and Gamification as Methods of Increasing Engagement

One of the central concepts of the application's linguistic functionality is personalization, which is realized through the adaptation of tasks to the user's level of knowledge and interests. For example, at the early stages of learning, a student will be able to work with words related to everyday topics such as "family" or "leisure," while advanced users will receive content on more complex topics such as "professional activities" or "politics." In this way, the app not only supports the individual needs of users, but also gradually increases the complexity of the learning material, which contributes to effective memorization and use of new words.

Gamification increases engagement through a system of rewards and motivational elements, such as difficulty levels and achievements for each completed stage. To expand vocabulary, a series of game exercises have been developed, such as "Build a Word", where the user makes a word from individual letters, or "Guess the Translation", which involves choosing the correct translation for a given word. This allows users to feel progress and improve their skills in using words in different contexts. In addition, pronunciation training exercises include tasks where the user pronounces words or phrases and receives feedback on the correctness of the pronunciation. For example, the pronunciation trainer uses samples, such as the

pronunciation of the sounds /θ/ and /ð/ in the words “think” and “this”, which allows you to immediately see mistakes and gradually achieve the correct sound.

3.2 Description of the Main Components of the Linguistic Content

To ensure effective learning, the new app incorporates key language content components:

1) Interactive Dictionary: Users can view definitions, synonyms, example sentences, and listen to audio recordings by native speakers. The feature allows users to add new words to their "personal vocabulary" for repetition, aiding in vocabulary reinforcement.
2) Thematic Units: Learning materials are organized into themes, such as Travel, Food, and Business English, allowing users to select units that match their needs. For instance, travelers can focus on air travel and hotel reservations, while professionals can learn business vocabulary.
3) Multimedia Elements: The app utilizes video, audio, images, and real-world usage examples to support memorization. Short videos featuring native speakers demonstrate vocabulary in context, and comprehension tasks enhance listening and understanding skills.

The app's linguistic functionality combines personalized tasks with multimedia components for effective English learning. Personalization and gamification increase learner interest, while a structured content organization promotes comprehensive language competency development.

Key requirements for the app include user registration and authentication to save progress and settings, along with the ability to create and manage personalized dictionaries. Users can add, organize, edit, and delete vocabulary entries, enhancing their learning experience [12].

An achievement system is essential for motivation and feedback, with interactive exercises and tests covering vocabulary, grammar, and listening. This enables systematic tracking of progress and task adaptation to learner needs [13].

To facilitate these functions, mechanisms for sorting and editing vocabularies and sharing them among users will be implemented. This encourages social engagement and collaborative learning. Offline access is also crucial, allowing users to learn anytime without a network connection, ensuring the app's availability [14].

3.3 Approach to Research

The main methods include user needs analysis, which is carried out through questionnaires and focus groups, to identify the goals and requirements of potential users. The next step is a comparative analysis of existing applications, such as Duolingo, Quizlet, and My Vocabulary, which helps to identify their strengths and weaknesses, contributing to the development of unique features of EnApp.

An important component is gamification methods, which involve the use of game elements to increase user motivation and engagement, including achievements, levels, and rewards that promote active learning. Linguistic methods also play an important role in the study, as they use the principles of lexicology and grammar to create content that meets modern language learning requirements.

The final stage is testing and approbation of the developed application in focus groups, which allows us to get feedback and analyze the results to identify possible shortcomings and eliminate them. This comprehensive methodology ensures a high level of functionality and efficiency of the developed application, meeting the needs of modern users in learning English.

The proposed application has the following functionality:

- User registration.
- User authorization.
- Ability to create your own dictionary.
- Achievement system.
- Learning words in a game-oriented way.
- Searching the dictionary.
- Editing words.
- Editing dictionaries.
- Shuffle words in the dictionary.
- Sort words.
- Use other people's dictionaries.
- Add pictures to the dictionary.

3.4 Architecture and Technical Support of the Application

The architecture of the interactive **EnApp** application is developed using IDEF0 diagrams, which provide a structured view of functional elements and their interconnections. The IDEF0 diagrams help visualize processes such as user registration, dictionary management, interactive

exercises, and statistical data collection. This offers a clear picture of system operation, facilitating further analysis and improvement. The focus is on ensuring a seamless flow of information between the app's components, which is crucial for achieving high performance and ease of use.

The application uses an SQLite database (Table 1, Figure 1) to store data, which is a lightweight and efficient solution for mobile applications. SQLite organizes data storage for user information, dictionaries, test results, and usage statistics. This technology ensures quick data access and reduces device load, as all data is stored locally. For example, information about words and their definitions is stored in a Vocabulary table, allowing users to quickly access necessary data without delays.

Table 1: Database structure.

Table	Description
Users	User information (ID, name, password, email)
Vocabulary	Dictionary entries (ID, word, definition, part of speech)
Achievements	Achievement system (ID, user ID, achievement type, date)
Exercises	Interactive exercises (ID, exercise type, associated word, user ID)

Figure 1: Application architecture.

In designing the architecture of EnApp, the focus is on implementing a system for the Android platform using Unity technology and the C# programming language. A central component of the architecture is the IDEF0 diagram, which illustrates the app's functionality, including an interactive offline mode that ensures independence from internet access (Table 2).

The app's operation involves several key stages. First, users can enter words or definitions into their dictionaries for future study, serving as input data for the system. Second, after learning words, users earn achievements, gaining in-app currency that they can use to customize their avatars. Third, the developer ensures the system's functionality and support while users edit dictionaries, learn new words, and achieve learning milestones.

The app's management process includes spell-checking to identify grammatical errors and reward verification to determine which user achievements are eligible for rewards. The application implements three primary processes: searching for and adding new words, studying materials, and monitoring progress.

In the first process, users input new words into their dictionaries for future study. Once added, the system checks their accuracy to ensure correct information. The second process involves studying existing material, where user performance data can be used to tailor the learning process to their needs.

Monitoring user progress provides feedback on their achievements. This enables the system to not only store data but also analyze it, which in turn enhances user motivation. The app also allows users to choose dictionaries created either by themselves or other participants, encouraging collaborative learning.

SQLite, a lightweight embedded database management system, was chosen for data storage. This system efficiently stores information about users, dictionaries, and achievements while providing quick access to essential data. It supports various operations, such as creating, reading, updating, and deleting records.

3.5 Software Implementation

Unity is an integrated development environment (IDE) designed for creating games and interactive applications, known for its power, flexibility, and accessibility, which has made it popular among developers worldwide. The environment provides a comprehensive set of tools for developing both 2D and 3D projects. With support for programming languages such as C# and JavaScript, Unity is accessible to a wide range of developers. Additionally, Unity benefits from an active user community, facilitating the development process through knowledge sharing and resource exchange.

The app management system follows a structured hierarchy of objects, consisting of several primary components. This project hierarchy includes interface components, logic, functionality, and auxiliary elements (Table 3).

Table 2: Application architecture scheme for english learning app.

Component	Description	Functions	Connection to other components
Database	Stores user information, progress, vocabulary, test results, profiles.	User data storage and updates	Works with learning content, analytics, feedback modules
Learning Content	Includes lessons, interactive materials (videos, audio), exercises for all proficiency levels and learning topics.	Access to various learning materials	Integrated with adaptive algorithms
Adaptive Algorithms	Adjust content difficulty based on user performance, enabling personalized learning.	Evaluate results, adjust tasks	Interacts with database, analytics, learning content
User Interface (UI)	Provides convenient access to lessons, progress management, app navigation, and visual hints.	Content management, access to statistics	Connects with motivation system, learning content
Analytics	Tracks user progress, calculates statistics, provides data to improve content and adaptive algorithms.	Report generation, graph building	Receives data from database, works with adaptive algorithms
Feedback Modules	Provides instant corrections and explanations for errors, hints, and tracks problem areas.	Error explanations, hints	Integrated with adaptive algorithms, analytics
Motivation System	Implements rewards, achievements, rankings, and daily goals to keep users engaged.	Score tracking, achievement notifications	Works with user interface and analytics
Authentication System	Responsible for secure user access, profile protection, and data security.	Login, profile management	Works with database and user interface
Social Media Integration	Allows users to share achievements, invite friends, and encourage active learning through peer support.	Achievement sharing, user invitations	Connects with motivation system, user interface
Multilingual Subsystem	Enables learning of different languages, supporting multiple languages for interface and learning content.	Translation, multilingual content support	Integrated with database, learning content
Notification System	Sends reminders, notifications about new achievements, and alerts about important events and content updates.	Sends push notifications, emails	Interacts with motivation system, user interface
Knowledge Assessment Subsystem	Monitors user proficiency, creates personalized tests to determine language proficiency level.	Test creation, progress assessment, task adaptation	Works with database, analytics, and adaptive algorithms
Support System	Provides users with assistance, answers questions, and resolves technical issues.	Error alerts, query responses	Integrated with user interface, database
Advanced Settings	Allows users to manage learning parameters, set desired difficulty level, and adjust personal preferences.	Learning pace adjustment, feedback settings	Interacts with user interface, adaptive algorithms
Admin Panel	Allows administrators to manage learning content, update materials, view app usage statistics, and configure algorithms.	Content management, access to statistics	Works with analytics, database, learning content

Table 3: Components description.

Component	Description
Object Hierarchy	Structured database of the app's components
Scene	Work environment for UI, adapted for app customization
Project Hierarchy	Directory of all project files and resources
Directories	Basic structural elements for storing app files

Integration functions are implemented through code methods such as HideImages() and ShowImages(), which manage the display of application windows. Methods for managing registration and authorization ensure control over form completion and user data validation. The app functionality includes a sequence of checks for registration field completion and navigation management between on-screen elements.

3.6 Testing

EnApp testing was conducted manually to verify functionality, reliability, and adherence to the planned usage scenarios. Each test scenario was evaluated on a 5-point scale, covering various aspects of registration, authorization, and game task completion, confirming the app's functionality and stability. This Table 4 highlights the main technological and software components used in developing the interactive app, along with the functionalities that make it user-friendly and engaging.

Table 4: Main technological and software components.

Component	Description
Integrated Development Environment (IDE)	Unity was used as the development base, providing a comprehensive set of tools for creating interactive elements and customizing the app interface, with support for both 2D and 3D.
Programming Languages	C# and JavaScript for interface and game component integration. Unity supports these languages, offering flexibility and accessibility for developers.
Multimedia Resources	The app utilizes images, sound effects, and animations to enhance engagement and immersion in the learning process.
Gamification Elements	Includes various mini-games (Flash Cards, Flying Tiger, Jumping Tiger), rewards, and in-app currency (ABCCoins), which increase interest in learning.
User Interface	Intuitive design with accessible, easy-to-perceive colors and a main character (tiger) that appeals to users of all ages.
Registration and Authorization Functions	Methods for checking field completion, registering new users, and password reset ensure an easy and secure login experience.
Testing	Manual testing of app functionality conducted by a group of students, covering various scenarios such as registration, authorization, and performing game tasks.

3.7 Comparison with Competitors

One of the key features of EnApp that significantly differentiates it from its competitors is gamification with real-life interactive scenarios that simulate real-life situations that users face in their daily lives or professional activities. For example, business communication or travel tasks provide an authentic language experience, allowing participants to practice vocabulary and grammar in a context close to real-world conditions. Such approaches contribute to more effective memorization and use of language material.

An analysis of Duolingo's limitations shows that it has poor support for features to improve pronunciation and contextual use of complex grammatical structures. For example, the platform doesn't provide any opportunities to get instant feedback on pronunciation, limiting itself to general exercises to choose the correct answer. EnApp solves these problems by using voice recognition technology and interactive exercises to repeat complex phrases with detailed analysis of mistakes made. This allows users to not only develop grammar skills but also improve phonetic accuracy.

What makes EnApp's interactive dictionary unique compared to Duolingo, where users only have access to fixed sets of words, is the ability to allow users to add their own entries, creating a personalized lexicon that greatly extends the platform's functionality. For example, EnApp users can integrate specific terms relevant to their professional activities or store unique idioms. Combined with multimedia components, such as native speaker audio, this feature provides a multichannel approach to learning vocabulary.

The third key element is adaptive algorithms that adjust the user's learning path according to their individual needs and level of knowledge. While Duolingo has a linear approach to presenting learning material, EnApp takes into account mistakes, exercise time, and repetition rate to dynamically adjust the difficulty of tasks. For example, if a user demonstrates poor accuracy in using conditional constructions, the program automatically adds additional exercises to practice this aspect of grammar.

An analysis of Duolingo's limitations indicates that it has poor support for features to improve pronunciation and contextualize the use of complex grammatical structures. For example, the platform doesn't provide any options to get instant feedback on pronunciation, limiting itself to general exercises to choose the correct answer. EnApp solves these problems by using voice recognition technology and interactive exercises to repeat complex phrases with detailed analysis of mistakes made. This allows users to not only develop grammar skills but also improve phonetic accuracy.

In terms of efficiency, preliminary tests of EnApp have shown that users reach a new level of knowledge 20% faster compared to using traditional methods or similar applications. In addition, the system automatically corrects up to 85% of grammatical errors in real time, providing appropriate explanations and recommendations. This is evidence of the high efficiency of EnApp's adaptive approach focused on individual user needs.

3.8 Analysis of EnApp Testing Results for Different Age Groups and Levels of Knowledge

The study of the effectiveness of the interactive application EnApp was conducted among students, faculty and technical staff of the National University "Odesa Polytechnic". The test participants were divided into three focus groups by age categories: young people (17-25 years old), adults (25-45 years old) and senior audience (45+ years old). The analysis included an assessment of the average progress in learning English, the level of motivation and interaction with the app for each group.

Participants in the Youth group (17-25 years old) showed the highest average progress in English learning, increasing their vocabulary by 35% in four weeks of using EnApp. Motivation remained consistently high, which is attributed to the interactive elements of the app, such as gamified scenarios and the achievement system. 92% of respondents in this category rated the app as user-friendly, effective, and providing more opportunities to develop active language skills, including pronunciation and contextual use of grammatical structures.

In the group of adults (25-45 years old), the progress rate was 27%, which is slightly lower than among young people. The reason for this is likely to be the greater employment of respondents, which limited their time for study. Nevertheless, the level of motivation remained high, especially among academic staff who used the program to improve their professional English. Adults responded positively to the personalized vocabulary and thematic exercises that allowed them to focus on professionally relevant vocabulary. EnApp provided an integrated approach covering pronunciation, grammar and listening.

The average progress in the older audience group (45+ years) was 20%. Although the older audience faced technical difficulties in the initial stages of using the app, the integration of offline mode and the simple interface helped to overcome these barriers. Participants' motivation largely depended on their professional need to improve their English for work or travel. 78% of respondents said that the interactive dictionary feature made learning new words much easier. Compared to other platforms such as My Vocabulary, EnApp provided more personalized learning.

3.9 Summarizing the Results

The test results confirmed the effectiveness of EnApp in different age categories. The app adapts most effectively to the needs of young people due to its interactive and gaming features, while adults and older audiences appreciate personalization and thematic focus. The data are summarized in the Table 5.

Table 5: Effectiveness of EnApp across age categories.

Category	Average Progress	Motivation Level	Benefits of EnApp
Youth (16–25)	35%	High	Gamification, active language exercises
Adults (25–45)	27%	Stable	Thematic exercises, personalization
Seniors (45+)	20%	Moderate	Simple interface, offline mode

Thus, EnApp has proven to be effective among users of different age groups, especially in the context of personalization and an integrated approach to learning English. Future improvements, such as adaptation for less technically skilled users, could further enhance the app's effectiveness.

EnApp's adaptive algorithms further elaborate on personalized learning by analyzing the results of tasks, accuracy of answers, number of errors, time spent on exercises, and frequency of access to support functions such as hints or an interactive dictionary. Based on this data, the program assesses the user's knowledge level, identifies strengths and weaknesses, and creates individualized tasks. For example, if a user makes mistakes in using complex conditional constructions: in forming the third type of conditional sentences (If I had known, I would have acted differently), the system analyzes these mistakes and adds exercises specifically aimed at working on this grammatical construction. Such exercises may include:

1) Recognizing the correct variant: the user is offered several sentences, among which they need to choose the ones that correctly use the third type of conditional sentence.
2) Filling in the gaps: the system provides sentences with missing verbs in the correct tense form (e.g. If I ___ (know) before, I ___ (do) it differently).
3) Practice creating sentences: the user is asked to write their own sentences using this construction, after which the program provides a detailed analysis and correction of errors.
4) Audio-visual exercises: short videos or audios where characters use the third type of conditional sentences, followed by questions to check comprehension.
5) In case of repeated pronunciation errors, EnApp offers interactive voice recognition exercises that include target phrases and provide detailed feedback.
6) The system dynamically adjusts the difficulty levels of tasks depending on performance. A long time spent on exercises can trigger a return to simpler tasks, while high accuracy of answers helps to make the material more difficult by adding more complex texts or contextual exercises. The error history allows the algorithm to focus on the basic aspects that require additional study.
7) EnApp algorithms interact with the database, analytics module, and content module. First, data on task completion is collected, which is then analyzed to determine the progress and individual needs of the user. Based on the information received, the program updates the learning path, providing flexibility in the learning process. EnApp's adaptive algorithms are a powerful tool that, through the integration of performance data, personalized exercises, and dynamic difficulty adjustment, creates an effective environment for achieving results.

3.10 Including Specific Learning Outcomes: Analyzing the Effectiveness of EnApp

The following section outlines the key findings regarding the effectiveness of EnApp, highlighting user performance, feedback, and areas for improvement:

1) The results of testing EnApp among users have shown significant improvements in language skills, which confirms the effectiveness of the app as an innovative tool for learning English. Typical gains include a 30% increase in vocabulary after four weeks of regular use. This was made possible by interactive dictionaries and gamified exercises that encourage repetition and reinforcement of new words in real-life scenarios. Significant improvements in pronunciation were also recorded among 85% of users who used the voice recognition exercises. These exercises provided detailed feedback, allowing for real-time correction of errors.
2) A comparison of performance between EnApp users and those who learned the language using traditional methods showed a marked advantage of the program. Participants who studied with EnApp showed an average increase of 25% in performance on listening and pronunciation tests compared to those who used traditional courses. One respondent noted: “Learning is faster than in real time because the system corrects my mistakes immediately and shows me the correct option.”

User feedback emphasizes the interactivity and user-friendliness of the program. Young people (16-25 years old) described the learning experience as “bright!” and “optimistic,” especially noting the gamified elements. For adults (25-45 years old), personalized topics and the ability to study “conveniently in a remote format” were the most attractive. The older audience (45+ years old) noted that the program is “still effective,” but needs some improvements in adjusting the difficulty levels.

Despite the high ratings, some users expressed a desire to improve the program's parameters. For example, the following comments were made: “It's better with a human teacher,” because ”you can ask the person if you don't understand.” This indicates the potential for integrating video conferencing features with teachers or chatbots to explain complex tasks.

Thus, EnApp is highly effective in improving key language skills. At the same time, providing opportunities for additional feedback and more flexible settings could make the program even more attractive to a wider audience.

3.11 Extending the Discussion: Analysis and Prospects for EnApp Development

The results of the EnApp testing have demonstrated numerous strengths of the program that significantly contribute to the effectiveness of English language learning. One of the most important advantages is the high user engagement provided by the gamification system. Interactive tasks, such as Flying Tiger and Jumping Tiger, combined with a system of achievements and rewards, encourage users to practice regularly. As the test participants noted: "The program makes you want to come back to it again, as the achievements motivate you to continue learning."

At the same time, the results revealed several areas for improvement. For older users (45+ years old), adapting the interface is a key recommendation. Simplifying navigation, using larger fonts, and detailed tooltips on functionality will help reduce technical barriers for this category. In addition, users of all ages expressed a desire to integrate personal feedback functions, such as video conferencing with teachers or chatbots that can explain complex tasks.

Another promising opportunity is the integration of EnApp with other platforms. For example, cooperation with Coursera or LinkedIn Learning will allow users to receive additional certificates for completing English language courses, which will increase the value of the app for professional development. Integration with collaborative learning platforms can also enhance the social component of EnApp by facilitating the exchange of experiences between users.

Expanding the functionality of EnApp could also include developing special modules for specialized English learning, such as for medical professionals, lawyers, or technical specialists. This will allow the program to meet the needs of professional users and expand its audience.

To summarize, EnApp is a powerful tool for learning English that demonstrates high efficiency thanks to innovative approaches such as gamification, personalization, and adaptive algorithms. Implementation of the recommendations for improving the interface, expanding functionality, and integrating with other platforms will allow the program to take a leading position among interactive language learning applications.

4 RESULTS AND DISCUSSION

To ensure the EnApp quality, manual testing was conducted along with feedback analysis from focus groups. A focus group of 120 students was engaged to identify the app's strengths, such as its ability to support English learning objectives, while also pinpointing areas for improvement.

Testing covered key educational features, including authorization, registration, and gamified scenarios aimed at building vocabulary, improving pronunciation, and reinforcing grammar. One of the main test scenarios was verifying the registration and authorization process, where users filled in required fields to access the platform. According to testing results, 98% of users successfully completed the registration and authorization stages without any issues**,** highlighting the reliability of the process. A reliability coefficient Cronbach's alpha $\alpha = 0.91$ (1) was calculated for the registration process, indicating high consistency in user responses.

Cronbach's Alpha (α) – Measures internal consistency (reliability) among test items:

$$\alpha =\cdot \frac{N}{v + (N-1) * c'} \qquad (1)$$

where:

- N is the number of items,
- c is the average covariance between item pairs,
- v is the average variance of each item.

This coefficient ranges from 0 to 1, where values above 0.7 indicate good reliability.

Gamified scenarios, which included translation tasks, sentence building, and spelling, were a crucial testing element. For instance, in the Flash Cards mode, users view an English word and either select the correct translation or pronounce it, immediately receiving feedback on accuracy. Users successfully identified and pronounced words with correct translations **85%** of the time, showing this mode's effectiveness in developing both passive and active vocabulary. The app provided instant feedback, confirming translations and helping with quick word retention. A positive correlation $\boldsymbol{r = 0.78}$ (2) was found between Flash Cards usage and vocabulary retention, confirming the mode's role in supporting word acquisition.

Pearson Correlation Coefficient (r) – Measures the linear correlation between two variables, such as Flash Cards usage and vocabulary retention:

$$r = \cdot \frac{\sum(X-\overline{X})(Y-\overline{Y})}{\sqrt{\sum(X-\overline{X})^2 \sum(Y-\overline{Y})^2}}, \quad (2)$$

where:

- X and Y are the variables (e.g., time spent on Flash Cards and vocabulary test scores),
- $\overline{X}$ overline and $\overline{Y}$ are the means of X and Y.

This coefficient ranges from -1 to 1, where values closer to ±1 indicate a strong linear relationship [15].

In Flying Tiger mode, the app encouraged participants to collect letters to form a correct English word. This mode helps improve spelling and builds confidence in using new words. **82%** of users were able to form correct words using letters in this mode, demonstrating a positive impact on spelling skills. Additionally, a high inter-rater reliability $\boldsymbol{k = 0.85}$ (3) was recorded for task success in this mode, indicating consistent scoring across testers. Some participants noted the need to adjust control sensitivity to ensure precise letter placement.

Cohen's Kappa (k) – Measures inter-rater reliability, indicating the consistency among different raters:

$$k = \cdot \frac{p_o - p_e}{1 - p_e}, \quad (3)$$

where:

- p_o is the observed agreement between raters,
- p_e is the expected agreement by chance.

Values closer to 1 indicate strong agreement, while values closer to 0 suggest weak agreement.

Jumping Tiger mode involved tasks for quickly recognizing the correct word translation to progress to the next platform, helping users improve their skills in choosing the correct meaning in a specific context. During Jumping Tiger testing, **75%** of users successfully chose the correct translation in context, confirming this mode's value for building speed in decision-making based on vocabulary knowledge. A moderate correlation $\boldsymbol{r = 0.65}$ (2) between Jumping Tiger scores and vocabulary decision speed was identified, supporting the mode's effectiveness in real-time vocabulary application.

Feedback from the focus group showed a high overall level of user satisfaction. Most students, 87%, rated the app as an effective tool for independent English study**,** particularly emphasizing its functionality for vocabulary building and grammar reinforcement. Furthermore, 90% of participants stated that the reward system positively impacted their motivation**,** encouraging them to return regularly for practice. A satisfaction index of 4.5 out of 5 was recorded for the app, with a standard deviation of 0.3, showing consistently high ratings among users.

The testing results revealed a few areas that required improvement, such as interface design (for this reason, this article does not offer its appearance), interface sensitivity in Flying Tiger and Jumping Tiger modes, which was adjusted for more accurate user action recognition. Additionally, it was recommended to add explanations for certain tasks to simplify understanding of grammatical rules, especially for beginners. It was also decided to provide more flexible difficulty level adjustments to better meet the needs of users at various learning stages.

5 CONCLUSIONS

The analysis shows that innovative, interactive English learning apps are highly effective due to their flexible, personalized, and gamified features, which strongly support language acquisition. With high reliability scores (e.g., Cronbach's alpha of 0.91 for registration) and positive user satisfaction (4.5 out of 5), EnApp not only simplify learning but also maintain user motivation. Gamified elements like in-app rewards and social sharing boost engagement, while correlations in vocabulary retention ($r = 0.78$) affirm the value of such methods in memory enhancement.

Future developments could expand the app's functionality by connecting with platforms like Coursera or LinkedIn for additional resources and community support, thus offering a comprehensive learning environment. Integrating AI and NLP will further refine personalization, making this app more engaging and adaptive. Overall, interactive English learning apps have great potential to create enjoyable, effective, and socially connected learning experiences, enhancing accessibility and tailoring learning to individual needs for better mastery of English.

ACKNOWLEDGMENTS

We acknowledge support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Open Access Publishing Fund of Anhalt University of Applied Sciences.

REFERENCES

[1] D. Crystal, "English as a Global Language," 2nd ed. Cambridge University Press, 2012.

[2] D. Graddol, "The Future of English? The British Council," 2010.

[3] J. Coleman, "English-medium teaching in European higher education," Language Teaching. 2013, 46(1), pp.4-28.

[4] B. Kumaravadivelu, "The decolonial option in English teaching: Can the subaltern act?" TESOL Quarterly. 2018, 52(4), pp. 741-764.

[5] H. Reinders and P. Benson, "Research agenda: Language learning beyond the classroom." Language Teaching. 2017, 50(4), pp. 561-578.

[6] C. Chapelle and S. Sauro, "The Handbook of Technology and Second Language Teaching and Learning," John Wiley & Sons, 2017.

[7] M. Levy and P. Hubbard, "Essential competencies for English-medium instruction: A framework," Journal of English for Academic Purposes. 2020, 44, pp. 101-107.

[8] EF EPI (English First English Proficiency Index) Report 2022, [Online]. Available: https://www.ef.com/wwen/epi.

[9] R. Vesselinov and J. Grego, "Duolingo effectiveness study: Learning outcomes and user experience," Language Learning Journal, 2021.

[10] S. Montaner-Villalba, "The use of Quizlet to enhance vocabulary in the English language classroom," Asian Journal of Education and e-Learning, 2018.

[11] M. Andarab, "Effect of Quizlet on EFL vocabulary acquisition," Asian Online Journal, 2018.

[12] G.-J. Hwang and S.-C. Chang, "A review of interactive learning strategies in language education: The impact on user engagement and achievement," Journal of Educational Technology, 2021.

[13] J. Hamari, J. Koivisto, and H. Sarsa, "Does gamification work? – A literature review of empirical studies on gamification," in Proceedings of the Annual Hawaii International Conference on System Sciences, 2016.

[14] S. Kim, "Offline mobile learning apps in the education of language learners," International Journal of Educational Technology, 2018.

[15] O. Vasylenko and I. Sencha, "Mathematical and statistical methods of analysis in applied research," Textbook. Odesa: ONAT, 2011, 166 p, [Online]. Available: http://ir.nuozu.edu.ua:8080/jspui/handle/lib/319.

Comparative Analysis of the Unavailability Factors for Two Types of Optical Cable Section Repair under Conditions of Gradual and Sudden Failures

Viatcheslav Shuvalov and Irina Kvitkova
Department of Infocommunication Systems and Networks, Siberian State University of Telecommunications and Information Science, Kirov Str. 86, 630102 Novosibirsk, Russia
shvp04@mail.ru, irin.creme@yandex.ru

Keywords: Section of Optical Cable, Sudden Failure, Semi-Markov Model, Unavailability Factor, Degradation Cycle, Signal Attenuation, Degradation Failure, Repair, Service Life, Recovery after Failure.

Abstract: The article presents a comparison of unavailability factors for two optical cable section types of repair under conditions of degradation and sudden failures. The first one involves replacing a section of optical cable in case of both degradation and sudden failure caused by outside interference. In the second type, the replacement of the optical cable section is carried out only in case of degradation failure. In case of a sudden failure, the optical cable is repaired by connecting at the point of breakage. To show the degradation process and analyze the impact of sudden failures, mathematical models of the semi-Markov process are used, which allow to determine the average recovery time and time to failure, the unavailability factor, as well as to estimate the duration of the degradation cycle. The degradation cycle covers the duration of optical cable operation from the initial state to the degradation failure, including the recovery time after sudden failures. This parameter can be an estimate the service life of optical cable or its section. At the same time, it is assumed that during operation, due to the cable sections replacement, the whole optical cable will be replaced. It is shown that the degradation cycle of the section in the first type grows with increasing the sudden failures rate, and in the second type, on the contrary, decreases compared to the planned value determined in the absence of sudden failures. The more frequent sudden failures occur, the more noticeable the difference in the values of the optical cable section unavailability factor calculated for both types of repairs. These results can be used to estimate the time needed to replace sections of optical cable due to degradation, as well as to select the optimal strategy for maintenance and repair of the access network optical cable.

1 INTRODUCTION

For many years, the issues of determining the optical cable service life have been discussed in publications [1-12, etc.]. This problem has gained particular interest in recent years, since the service life of the laid optical cable has already exceeded the warranty period specified by the manufacturer. It should be noted that the service life of an optical cable is usually understood as the time that determines the need to replace the laid optical cable due to the fact that its limiting state caused by degradation processes has occurred. The limiting state is understood as the object state in which further operation is unacceptable or impractical, or restoration of its functional state is impossible or impractical [13].

Degradation processes in an optical cable can lead to its rupture, which is determined by its mechanical reliability [7], or to an increase in the signal attenuation coefficient, at which the level at the receiver input may become below the sensitivity threshold (optical reliability).

From a mechanical point of view, the durability of any object, including optical fiber, is estimated:

- the initial strength of the optical fiber;
- the presence and depth of microcracks on the fiber surface;
- the rate of optical fiber degradation, determined by the rate of microcracks growth on its surface;

- an environment that leads to a weakening of the material from which the optical fiber is made.

There are three stages of microcracks development. The first stage is the stage of microcrack nucleation; the second is the stage of microcrack depth growth, the third lasts for seconds and ends with the rupture of the optical fiber. The second stage lasts for years, and it is mainly this stage that determines the service life of the optical cable.

In [1], various scenarios for predicting the service life of optical fiber in a cable communication line are considered. The forecast is based on an expression borrowed from [14, 15]:

$$t_a = t_p \cdot \left(\frac{\sigma_p}{\sigma_a}\right)^{n_q} \cdot \left\{\left[1 - \frac{\ln(1-F)}{N_p \cdot L_{of}}\right]^{\frac{1}{m_S}} - 1\right\}, \quad (1)$$

where t_p is time to fiber failure obtained during fiber testing; σ_p is the load at which testing was carried out; L_{of} is the length of the fiber for which the service life is predicted; N_p is the number of fiber breaks during testing under load; n_q is the strength parameter of quartz glass; m_s is the Weibull distribution parameter; F is the probability fiber failure; σ_a is the load applied to the fiber during its service life.

So, in order to determine the time a fiber section damage (breakage, failure), it is necessary, according to (1), to have the results of cable testing, to know the load on the fiber and the Weibull distribution parameter m_s. In this case, the time of the fiber failure depends on the probability of failure F, which must be set. In fact, using (1), it is only possible to determine the probability of optical fiber failure during t_a, and not its lifetime, i.e. the time from the start of operation to the fiber replacement. According to [16], the value (1-F) is chosen to be 0.95. It is noted in [1] that the load on individual sections of the fiber-optic line may differ; therefore, it is proposed to calculate the so-called equivalent load, on the basis of which the service life of the fiber as a whole is predicted.

The widespread approach presented above to determining the replacement time of an optical cable caused by fiber degradation seems formal, since one-time costs of replacing an optical cable and information losses caused by the time spent on replacing an optical cable are not taken into account. In addition, it should be borne in mind that the load on the cable varies in different sections. Similar questions arise when considering the replacement period of an optical cable with an approach based on optical reliability.

In this regard, it is proposed to consider an optical cable with a length of L km as a set of n sequentially connected sections with a length of l km, where $n = L/l$, and to consider, in case of a degradation failure [13], the replacement of sections as they fail according to the rules determined by the method of maintenance and repair of these sections, taking into account the occurrence of sudden failures. During the maintenance process, cable sections can be repaired by connecting at breakage points, replacing individual cable sections, and replacing the entire optical cable. Deciding which repair will be the most profitable is a difficult technical and economic task [17]. Thus, replacing the entire optical cable is very expensive and is performed when the cost of replacing the entire cable will be less than the cost of replacing individual sections and making connections.

We will consider the service life of the sections, assuming that it is determined by degradation processes that lead to a break in the optical cable section or exceeding the signal attenuation of a critical level. So, degradation processes determine the service life of optical cable sections and the need to replace them. After a certain time all sections of the optical cable will be replaced with new ones. It is clear that it will take significantly less time to replace the section at $l << L$ than to replace the entire cable, and the costs associated with such a replacement will be stretched over time and incomparable with the costs of replacing the cable as a whole.

The article compares unavailability factors for two types of optical cable repair under conditions of gradual (degradation) and sudden failures, which differ by the influence of sudden failures on the degradation process. The first type of repair involves replacing a section of optical cable both in case of degradation failure and sudden failure [3, 4]. We assume that during operation, due to the replacement of cable sections, the whole optical cable will be replaced. Thus, only the replacement of cable sections is considered. Then the degradation cycle of the entire optical cable, characterizing its lifetime, will tend to infinity. The degradation process in [3, 4] is described by the Markov process of pure death. Sudden failures at the same time lengthen the degradation cycle, since after a sudden failure, the section is replaced with a new one. The second type involves replacing the section only in case of degradation failure. At sudden failure, which is

accompanied by a rupture of the optical cable in the section, communication is restored by welding or mechanical connection at the point of breakage.

It should be noted that both types of repair can be considered in relation to mechanical and optical reliability.

To determine the unavailability factor for the first and second types of repair, we divide the planned degradation cycle, determined in the absence of sudden failures, into n intervals of T_D duration. Each interval represents the corresponding degradation state of an optical cable section. In the last n^{th} state, a degradation failure occurs, leading to the replacement of an optical cable section. In each such condition, a degradation process takes place and a sudden failure may occur, which has an obvious character. Sudden failures are distributed exponentially with the same rate in each state. After each sudden failure, the section is restored, while during the restoration the cable is not used for its intended purpose, that is, it is inoperable. The impact of sudden failures is considered under conditions of continuous reliable monitoring.

The unavailability factor is determined by the expression [3]:

$$F_U = \frac{T_{DS}}{T_{CD}}, \qquad (2)$$

where T_{DS} is average down state time per degradation cycle; T_{CD} is average duration of the degradation cycle.

When creating mathematical models for the conditions described above, semi-Markov process state transitions diagrams are compiled, simulating the degradation process.

2 STATE-TRANSITIONS DIAGRAM OF THE OPTICAL CABLE SECTION FOR THE FIRST TYPE OF REPAIR

The type of repair involves replacing a section of optical cable because of a sudden failure. The elimination of sudden failure is carried out with recovery rate μ_1. After restoration, the transition to the initial state of degradation is carried out, as shown on Figure 1, which shows the state-transitions diagram on the degradation cycle.

There are following signs on the Figure 1: D_i – i^{th} degradation state, $i = 1, 2, ..., n$; R – recovery after sudden failure; μ_1 – recovery rate of a section after sudden failure; μ_2 – recovery rate of a section after degradation failure; p_D - probability of transition between two degradation states (probability of that a sudden failure will not occur in this state of degradation); q_D - the probability of transition to recovery after a sudden failure in one state of degradation (the probability of a sudden failure in a state of degradation) [3, 4].

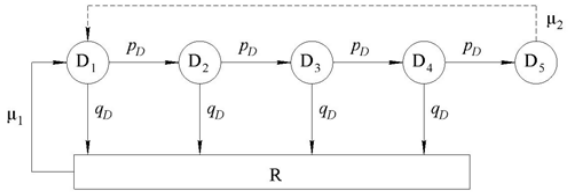


Figure 1: The state-transitions diagram on the degradation cycle for the first type of repair at $n = 5$.

In each degradation state a sudden failure may occurs. This event in the time interval T_D is characterized by two probabilities [18]:

$$p_D = \exp(-\lambda \cdot T_D);\ q_D = 1 - \exp(-\lambda \cdot T_D), \qquad (3)$$

where λ is a sudden failure rate.

The average time of down state during the degradation cycle is determined by the expression [3]:

$$T_{DS1} = n \cdot q_D \cdot \theta_{DS} + T_R = \\ = n \cdot q_D \cdot \frac{\lambda \cdot T_D - q_D}{\lambda} + \left(\frac{1 - p_D^{n-1}}{p_D^{n-1}} \cdot \frac{1}{\mu_1} + \frac{1}{\mu_2} \right), \qquad (4)$$

where θ_{DS} is the average time spent in the down state after a sudden failure in a state of degradation, determined by (5) in [3]; T_R is the average recovery time on the degradation cycle, determined by (14) in [3]; μ_1 is the cable section recovery rate after a sudden failure; μ_2 is the cable section recovery rate after degradation failure; p_D and q_D are defined by (3).

The average duration of the degradation cycle is determined by the expression [3]:

$$T_{CD1} = \frac{1 - p_D^{n-1}}{(1 - p_D) \cdot p_D^{n-1}} \cdot T_D + T_R. \qquad (5)$$

3 STATE-TRANSITIONS DIAGRAM OF THE OPTICAL CABLE SECTION FOR THE SECOND TYPE OF REPAIR

The second type of repair is characterized by the restoration of an optical cable section after a sudden failure by connecting at the breakage point.

Let's consider the i^{th} degradation state. At a sudden failure the current degradation state consists of three parts: before the failure D'_i, recovery R_i and after the sudden failure D''_i (Figure 2).

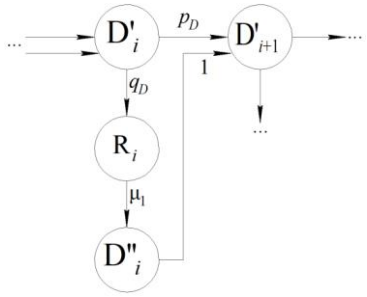


Figure 2: The state-transitions diagram on one degradation state for the second type of optical cable repair.

There are following signs on the diagram Figure 2: i − degradation state number, i = 1, 2, ..., n-1.

The degradation state is estimated by the signal attenuation parameter, which value increases from state to state. In some state, in the absence of sudden failures, the attenuation parameter reaches a limit value at which a degradation failure appears. After the occurrence of such a failure, the cable section is restored by replacing it with a new one.

The initial state of the degradation cycle is the state corresponding to the minimum value of the attenuation parameter. In the current state of degradation, the attenuation parameter increases by the value Δb in the absence of sudden failures. Raising the attenuation parameter in one state is transferred to the next degradation state when switching to it. Sudden failures can cause different changes in the attenuation parameter. This phenomenon is taken into account using the coefficient η. At η = 0 sudden failures do not affect the degradation process; at η = 1 sudden failures have the same effect as the degradation process in one state; at η > 1 sudden failures have a greater impact compared to the degradation process in one state.

The process of transitions between states in the degradation cycle is shown on Figure 3.

The transition $D_i \rightarrow D_{i+1}$ occurs at the end of T_D regardless of a sudden failure presence or absence. At the same time, from the point of view of reliability, the interval T_D consists of two parts: up (D'_i and D''_i) and down (R_i) parts.

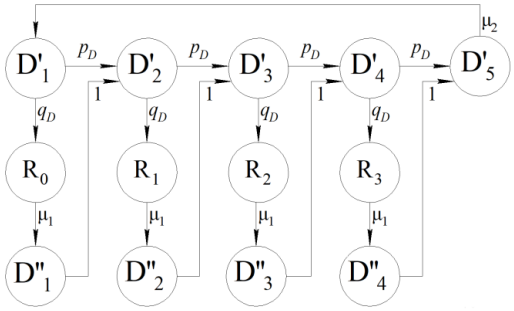


Figure 3: The state-transitions diagram on the degradation cycle for the second type of optical cable repair at n = 5.

In any case, in the state D_{i+1}, the attenuation parameter increases by the value Δb due to the degradation process in the state D_i. However, an increase in the attenuation parameter due to a sudden failure happens only if a sudden failure occurs. Since the probability of a sudden failure is q_D, this increase will be on average by the amount of $\eta \cdot \Delta b \cdot q$.

The increase in the attenuation parameter is expressed as follows:

$$b_{i+1} = b_i + \Delta b + \eta \cdot \Delta b \cdot q_D \,, \qquad (6)$$

where b_i is the parameter of signal attenuation in the i^{th} degradation state.

Taking into account (6), the calculation of the attenuation parameter in the state D_i under the condition $b_0 = 0$ is:

$$b_i = i \cdot (1 + \eta \cdot q_D) \cdot \Delta b,\ i = 1, 2, \ldots, n \,. \qquad (7)$$

In accordance with (7), at $\eta \neq 0$ and $q_D \neq 0$, the increase in the signal attenuation parameter during the transition from state to state will be greater than the planned value Δb. Degradation failure in the absence of sudden failures will occur at the value of the attenuation parameter $n \cdot \Delta b$. Degradation failure, taking into account sudden failures, will occur when the condition $b_i \geq n \cdot \Delta b$ is fulfilled, which is expressed by the inequality:

$$I \geq \frac{n}{1 + \eta \cdot q_D}, \qquad (8)$$

where I is the number of the condition with degradation failures. This number is simultaneously the number of states included in the degradation cycle. It is obvious that $I < n$.

When condition (8) is fulfilled, the degradation cycle ends in the state D_I the optical cable section is replaced, then there is its further operation with a new degradation cycle.

In each state of degradation, during the restoration of the section, the optical cable is not used for its intended purpose, i.e. it is on down state. The average recovery time on one state is q_D/μ_1, where q_D is determined by (3).

The down state average time of an optical cable during the degradation cycle, taking into account the time to replace the entire cable after a degradation failure, is determined by the expression:

$$T_{DS2} = \frac{\mu_1 + I \cdot \mu_2 \cdot q_D}{\mu_1 \cdot \mu_2}. \quad (9)$$

The average degradation cycle time is:

$$T_{CD2} = I \cdot T_D + 1/\mu_2 . \quad (10)$$

4 COMPARATIVE ANALYSIS OF THE OPTICAL CABLE SECTION UNAVAILABILITY FACTORS FOR VARIOUS TYPES OF REPAIR

The unavailability factor of an optical cable section per degradation cycle for the first type of repair, taking into account (2), is determined by the expression:

$$F_{U1} = \frac{T_{DS1}}{T_{CD1}}, \quad (11)$$

where T_{DS1} is determined by (4); T_{CD1} is determined by (5).

The unavailability factor of an optical cable section per degradation cycle for the second type of repair, taking into account (2), (8)-(10) is determined by the expression:

$$F_{U2} = \frac{T_{DS2}}{T_{CD2}} = \frac{\mu_1 + \frac{n}{1+\eta \cdot q_D} \cdot \mu_2 \cdot q_D}{\left(1 + \frac{n}{1+\eta \cdot q_D} \cdot \mu_2 \cdot T_D\right) \cdot \mu_1}. \quad (12)$$

Let's consider the behavior of the unavailability factors defined by (11) and (12) for both types of repair at different values of the sudden failures rate λ. To do this, let enter the following initial data:

- the planned number of degradation states n=30;
- duration of one degradation state T_D = 1 year = 8760 h;
- the recovery rate after a sudden failure for the first type of repair μ_1 = 1/10 1/h, for the second type of repair - 1/4 1/h;
- the recovery rate after degradation failure μ_2 = 1/10 1/h [19].

In the absence of sudden failures over the entire interval of the planned degradation cycle, an increase in the attenuation parameter due to degradation in one state can be assumed to be equal Δb = 0.2 dB with an energy reserve of 6 dB [18, 20], when a planned degradation failure occurs.

In the presence of sudden failures, restoration of the optical cable after failure increases the attenuation parameter by 0.05 dB for the welded method of connecting optical fibers and by 0.5 dB for the mechanical connection of optical fibers [5, 18]. Thus, the increase in the attenuation parameter due to a sudden failure $\eta \cdot \Delta b$ is equal 0.05 dB or 0.5 dB at Δb = 0.2 dB. Then the sudden failure impact factor η is 0.25 and 2.5, respectively. For calculations, let's take the worst case when η = 2.5.

Figure 4 (a), (b) shows the dependence of the average degradation cycle time of an optical cable section (in years) on the sudden failures rate $\lambda = 0, 10^{-9}, \ldots, 10^{-5}$ 1/ч.

According to Figure 4, it can be seen that at the first type of repair, the growth in the number of sudden failures increases the degradation cycle of the cable section, which is associated with more frequent replacement of the entire section with a new one. The second type of repair is characterized by a decrease in the duration of the degradation cycle compared to the planned value, which is associated with the appearance of additional attenuation due to the restoration of the connection at the breakage point.

Figure 5 (a), (b) shows the dependence of the optical cable section unavailability factor F_U at the degradation cycle on the sudden failures rate $\lambda = 0, 10^{-9}, \ldots, 10^{-3}$ 1/ч.

The values of the unavailability factors F_{U1} и F_{U2} at different values of λ are shown in Table 1.

At the second type of repair the values of the unavailability factor are lower than at the first type, as follows from Table 1. Moreover, the more often sudden failures occur, the more noticeable the difference. This is due to the repair time, since it takes longer to replace the entire section of the optical cable than to weld or mechanically connect the fiber breakage point.

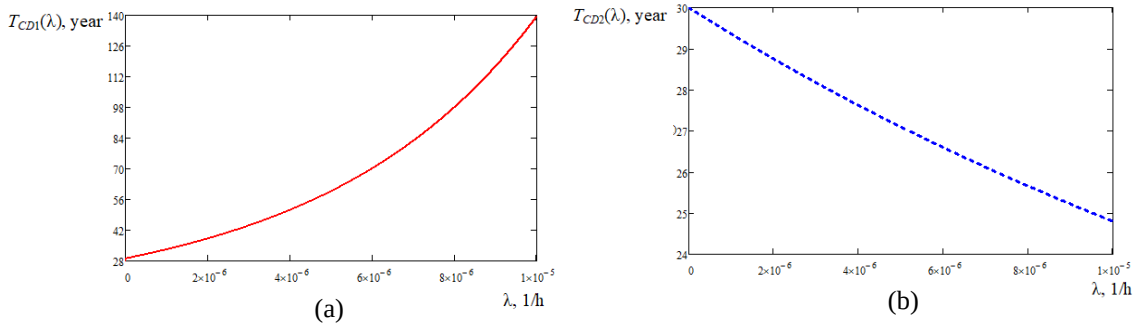


Figure 4: Dependence $T_{CD}(\lambda)$ for the first (a) and second (b) types of optical cable repair.

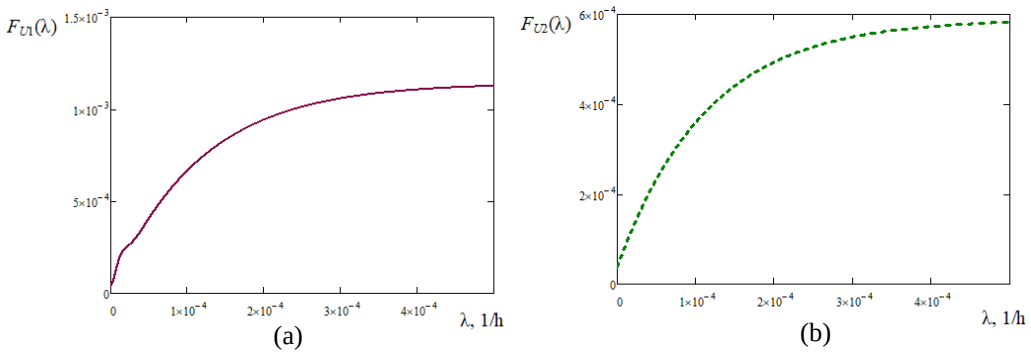


Figure 5: Dependence $F_U(\lambda)$ for the first (a) and second (b) types of optical cable repair.

Table 1: The values of the unavailability factor for two types of optical cable repair.

Parameter	Value						
λ, 1/h	10^{-9}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}
F_{U1}	$3.937\cdot10^{-5}$	$3.941\cdot10^{-5}$	$3.985\cdot10^{-5}$	$4.47\cdot10^{-5}$	$1.684\cdot10^{-4}$	$6.657\cdot10^{-4}$	$1.14\cdot10^{-3}$
F_{U2}	$3.806\cdot10^{-5}$	$3.81\cdot10^{-5}$	$3.853\cdot10^{-5}$	$4.286\cdot10^{-5}$	$8.432\cdot10^{-5}$	$3.6\cdot10^{-4}$	$5.896\cdot10^{-4}$

5 CONCLUSIONS

The models presented in the article can be used to evaluate the degradation cycle of an optical cable, which determines the period of its replacement due to cable optical fiber aging. Two types of an optical cable repair due to a sudden failure are considered.

In the first type of repair, the optical cable section is replaced in case of a sudden and degrading failure, as a result of which, after some time, the entire cable will be replaced.

In the second type of repair, the cable section is restored due to a sudden failure, the connection is made at the breakage point, which leads to an increase in signal attenuation on the optical cable section, and its degradation cycle decreases compared to the planned value of 25-30 years. Thus, it becomes possible to specify the time of replacement of an optical cable section due to degradation.

It is shown that the presence of sudden failures in the first type of optical cable repair leads to an increase in the degradation cycle, and in the second type of repair – to a decrease in the degradation cycle of the optical cable.

It is advisable to use the second type of repair for restoration at the early stages of optical cable operation, and at the end of the planned degradation cycle – the first type, which will increase the time before replacing the whole cable.

Issues related to the efficiency of using a particular type of repair require consideration not only from the point of view of reliability, but also the economic costs of cable restoration due to failure. At the same time, it is obvious that the cost of replacing an optical cable section exceeds the cost of connecting the cable at the breakage point.

The paper was carried out according to the state task 071-03-2024-008 dated 19.01.2024, Novosibirsk. Russia.

REFERENCES

[1] V. Andreev, V. Burdin, and A. Nizhegorodov, "Scenarios for predicting the service life of optical fiber in cable communication lines [Scenarii prognoza sroka sluzhby opticheskogo volokna v kabel'nyh liniyah svyazi]," (in Russian), The First Mile [Pervaya milya], vol. 4 (89), pp. 34-42, 2020.

[2] V. Burdin, V. Andreev, A. Bourdine, M. Dashkov, and A. Nizhgorodov, "Reliability and lifetime estimations for field-aged optical cable," in Proceedings of 20th International Conference, NEW2AN 2020 and 13th Conference, ruSMART 2020 St. Petersburg, Russia, August 26–28, 2020, part II, pp. 364-372.

[3] V.P. Shuvalov, B.P. Zelentsov, and I.G. Kvitkova, "On the effect of sudden failures and control errors of the first kind on the degradation cycle of an optical cable," in Proceedings of 2023 IEEE 16th International Conference of Actual Problems of Electronic Instrument Engineering (APEIE), Novosibirsk, Russia, November 10-12, 2023, pp. 550-555.

[4] B. Zelentsov, V. Shuvalov, I. Kvitkova, and D. Kachan, "Methodology for evaluating the impact of sudden failures on the reliability parameters of the optical cable damaged section," in Proceedings of International Conference on Applied Innovation in IT, Koethen, Germany, March 7-9, 2023, vol. 11, issue 1, pp. 43-48.

[5] A. Maslo, M. Hodzic, E. Skaljo, and A. Mujcic, "Aging and degradation of optical fiber parameters in a 16-year-long period of usage," Fiber and Integrated Optics, vol. 39, issue 1, pp. 39-52, February 2020.

[6] A. Maslo, M. Hodzic, A. Mujcic, and N. Goran, "Analysis of optical fibers characteristics due to different influences," in Proceedings of the International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (IAT 2019). IAT 2019. Lecture Notes in Networks and Systems, vol. 83, Springer, 2020, pp. 496-503.

[7] A. Maslo, E. Škaljo, M. Hodžić, and A. Mujčić, "Mechanical reliability and lifetime of optical fibers after 20 years of use," in Proceedings of 2023 International Workshop on Fiber Optics on Access Networks (FOAN), Ghent, Belgium, October 30-31, 2023, pp. 61-66.

[8] R.E. Abdi, R.L. Pinto, G. Guérard, and C. Capena, "Lifetime of optical fibers submitted to thermo-mechanical stresses," in: Proceedings of the International Conference on Reliable Systems Engineering (ICoRSE) - 2021. ICoRSE 2021. Lecture Notes in Networks and Systems, vol 305, Springer, 2022, pp 23–31.

[9] A.Y. Tsym, "Service life of optical cables. Analyzes. Risks [Sroki sluzhby opticheskih kabelej. Analizy. Riski]," (in Russian), Cables and wires [Kabeli i provoda], vol. 2 (382), pp. 20-26, 2020.

[10] B.L.M. Yaméogo, D.W. Charlton, D. Doucet, C.Desrosiers, M.O. Sullivan, and C. Tremblay, "Trends in optical span loss detected using the time series decomposition method," Journal of Lightwave Technology, vol. 38, issue 18, pp. 5026-5035, 2020.

[11] A.A. Voronkov and I.N. Alekhin, "Forecasting the service life of optical communication cables operating at low temperatures [Prognozirovanie sroka sluzhby opticheskih kabelej svyazi, ekspluatiruyushchihsya v usloviyah nizkih temperatur]," (in Russian), Proceedings of the Samara Scientific Center of the Russian Academy of Sciences [Izvestiya Samarskogo nauchnogo centa Rossijskoj akademii nauk], vol. 16, issue 4 (3), pp. 516-519, 2014.

[12] V.A. Burdin, V.A. Andreev, A.V. Bourdine, M.V. Dashkov, and A.O. Nizhgorodov, "The forecast of the optical cable service life. Two approaches to assessing the actual optical fibers strength of a cable line. [Prognoz sroka sluzhby opticheskogo kabelya. Dva podhoda k ocenivaniyu aktual'noj prochnosti opticheskih volokon kabel'noj linii]," (in Russian) Photon Express [Foton-ekspress], vol. 6 (174), pp. 220-221, 2021

[13] GOST R 27.102-2021 Dependability in technics. dependability of item. terms and definitions [Nadezhnost' v tekhnike: nadezhnost' ob"ekta. terminy i opredeleniya], (in Russian), Moscow: Russian Institute of Standardization [Rossijskij Institut Standartizacii], Russia, 2021.

[14] ITU-T G-series Recommendations – Supplement 59, Series G: transmission systems and media, digital systems and networks. Guidance on optical fibre and cable reliability, Feb. 2018, 21 p.

[15] IEC TR 62048: 2014 Optical fibres – Reliability – Power law theory, 2014. - 66 p.

[16] GOST R 52266-2020 Optical cables. General technical conditions [Kabeli opticheskie. Obshchie tekhnicheskie usloviya], (in Russian), Moscow: Standartinform, Russia, 2020.

[17] F.D. Riswanto, T.R. Biyanto, T.F. Iballe, A.I. Abdillah, and A.O. Adista, "Maintenance cost optimization on reliability centered maintenance based on failure rate on flash gas compression system," in Proceedings of AIP Conference, Surabaya, Indonesia, October 31 – November 2, 2019, vol. 2088, issue 1, pp. 020043.

[18] V.I. Efanov, Design, Construction and Operation of FOCL [Proektirovanie, stroitel'stvo i ekspluataciya VOLS], (in Russian), Tomsk: Tomsk State University of Control Systems and Radioelectronics, Russia, 2012, p. 102.

[19] RD 45.047-99 Fiber-optic transmission lines on the mainline and intra-zone primary communication networks of Russia [Linii peredachi volokonno-opticheskie na magistral'noj i vnutrizonovyh pervichnyh setyah VSS Rossii], (in Russian), Moscow: Ministry of Communications of Russia, Russia, 2000.

[20] C. Mazzei, M. Crescitelli, D. Fioramanti, A. Quagliarini, A. Reale, and F. Brunetti, "Technical–economic analysis to identify the acceptable maximum attenuation on PON FTTH lines for wholesale network operators", Science Report, vol. 13 (1), pp. 12327, July 2023.

On the Need to Clarify Methods for Assessing the Quality of Sound Paths Through the Use of Psychoacoustic Criteria

Nadezhda Abrosimova, Anna Belezekova, Evgeny Isenev and Maxim Shushnov
Siberian State University of Telecommunications and Information Science, Kirov Str. 86, 630102 Novosibirsk, Russia
skachok94@mail.ru, anna-belezekova@mail.ru, jon98701@mail.ru, efemerian@gmail.com

Keywords: Sound, Path, Recording, Amplification, Technique, Distortion, Psychoacoustics.

Abstract: Currently, the evaluation of the sound quality of a product or a sound path is usually based on objective physical measurements. Despite this, in real a quality of the sound path is evaluated by human hearing. The listener evaluates the sound quality of the product or sound path using his auditory system, and not a set of physical measuring devices. In this situation, psychoacoustics is a scientific field that allows bridging the gap between physical or objective and subjective assessments. In a number of previous psychoacoustic experiments, some relations were established between the physical representation of the result and the audible perceptual sensation correlated with it. Assessing the quality of audio signals is an important consideration in making high quality sound product and sound paths and various methods have been developed. In this article is review interconnection of electrical indicators of sound quality and subjective psychoacoustic criteria assessment. It is proposed to apply the psychoacoustic assessment criteria because not all qualitative indicators can be measured using objective methods.

1 INTRODUCTION

In the modern world, music, speech and other sounds surround us from all sides. Sound paths are found in television and radio broadcast studio hardware complexes, recording and mastering studios, when listening to phonograms through household devices, etc. In other words, any sound that we hear today has indirectly passed through a certain sound path - a set of devices for capturing, amplifying, processing, transmitting and reproducing sound information. It is difficult to find a person who does not appreciate and love pure and high-quality sound reproduction. But today there is no clear understanding of what the quality of sound reproduction depends on [1]-[3].

Many factors are responsible for a pleasant sound signal, but the key and important one today is an objective assessment of the quality of sound paths using standard measurement techniques. However, more and more often, questions arise about the adequacy and comprehensiveness of objective methods for assessing quality, since among professional musicians, sound engineers, music lovers and ordinary listeners, more and more attention is being given to subjective assessment, and the majority are practically not interested in data on technical characteristics.

Subjective assessment includes blind or non-blind listening to the sound path with an assessment of a number of psychoacoustic criteria for the quality of sound reproduction and provides more information about the qualitative properties of the sound path than measuring only electrical characteristics. The principles and criteria of psychoacoustic subjective assessment have been independently formed over the past 40-50 years among quality sound lovers, musicians, sound engineers, audiophiles, etc.

2 SYSTEMATIZATION OF OBJECTIVE AND SUBJECTIVE CRITERIA FOR ASSESSING SOUND QUALITY WITH THE ESTABLISHMENT OF RELATIONSHIPS

2.1 Quality Assessment Standards System

Any existing objective method for assessing the quality characteristics of analog audio devices [4]-[6] and digital paths [7]-[10] professional and household is implemented using a set of one or more measuring instruments designed to measure electrical quantities.

Summarizing the methods, we can highlight general approaches to objective assessment. Measurements of electrical quantities always affect such characteristics as the operating frequency range, unevenness of the amplitude-frequency response (AFC), non-linear distortion factor (or harmonic distortion factor), intermodulation distortion factor, the nature of the transient process under pulsed action, the signal-to-noise ratio at the output and the maximum level signal or dynamic range, channel separation (in a multi-channel system). However, none of these electrical characteristics are relevant to assessing the listener's perception of the result.

Today there is an absurd situation. Developers, manufacturers and marketers in the audio industry continue to surprise with numbers that speak of the extremely high quality of their devices, and buyers are increasingly leaning towards devices with low quality indicators, “vintage” devices, tube designs and devices without specifying their quality characteristics.

Also, a number of listeners refuse digital media, Internet streaming resources and listen to recordings from magnetic tape, on vinyl discs, although there are some adherents of the seemingly outdated CD format. In [11] it is indicated that sometimes sound engineers often deliberately reduce the quality of the sound signal during processing in order to give a pleasant sound to the created phonogram, radio program, etc.

At the same time, Appendix A of GOST IEC 61606-3-2014 [9] contains an indication of the possibility of using an alternative method for assessing electrical characteristics in order to obtain more information, but does not provide for the possibility of conducting a subjective assessment of sound quality.

2.2 Objective and Subjective Assessment Methods

Among sound engineers, music lovers, musicians and audiophiles, sound quality assessment criteria such as audibility, fullness, loudness, warmth or neutrality, timbre, tonal balance, high register, dark background, stage, micro-detail, macro-detail, fatigue or fascination are often used. There are other subjective criteria, but they are either rare or include those mentioned above.

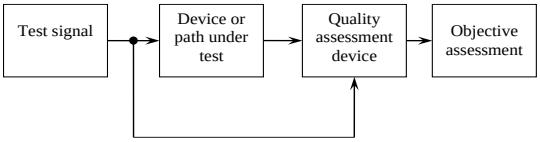


Figure 1: Scheme for objective assessment of the quality of the audio path.

Figure 1 shows a diagram of establishing the relationship between the electrical characteristics of the device and the psychoacoustic evaluation criteria. The complexity of using the scheme lies in the unclear understanding of exactly what electrical characteristics of the device under test need to be measured electrically to combine the result with the psychoacoustic assessment criteria, as well as the formation of an expert group. For this reason, it is possible to introduce new qualitative characteristics for assessing the electrical characteristics of analog audio devices and digital paths that go beyond those specified in [4]-[10]. As experts, it is necessary to involve both professionals who have experience in hearing examination or have the skill of professional assessment in the course of practical work, as well as groups of non-professionals of different age groups, since the properties of human hearing change with age. However, the requirement to divide into age groups is not mandatory. The testing scheme for subjective assessment is shown in Figure 2.

Figure 2: Scheme for subjective assessment of the quality of the audio path.

As it can be seen from Figure 2, in a subjective assessment, human perception takes place, which is based on psychoacoustics. This circuit does not take into account the response of the electrical meter - the evaluation device as in Figure 1, but takes into account the human reaction to the result of the passage of an audio signal through the path or device being evaluated.

3 PSYCHOACOUSTIC EVALUATION CRITERIA AND THEIR CONNECTION WITH THE ELECTRICAL PATH

In accordance with the division of objective quality indicators and psychoacoustic criteria, a correspondence diagram has been compiled in Figure 3. The diagram was compiled based on systematization and analysis of information available on the Internet from specialized resources, as well as based on a survey of audio equipment buyers and audio professionals, and does not take into account age groups. Solid lines show established connections, and dotted lines indicate that the relationship is possible, but not obvious.

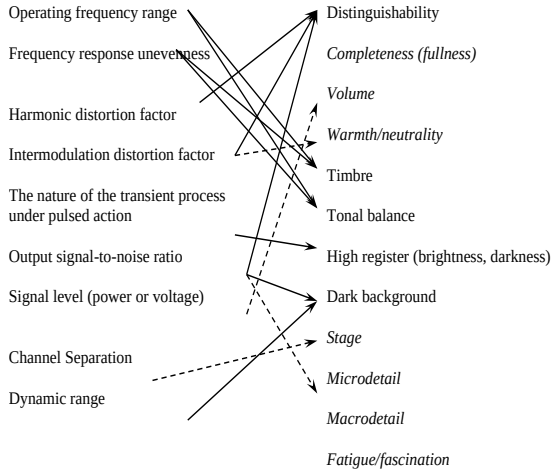


Figure 3: Correspondence of objective indicators for assessing the quality of the sound path to the subjective characteristics of psychoacoustic assessment.

The presence of uncertainties in the diagram in Figure 3 indicates an incomplete assessment of the psychoacoustic properties of sound paths using existing methods for measuring electrical characteristics.

As can be seen from Figure 3, *completeness* (fullness) of the sound picture, *scene*, *micro-detail*, *macro-detail*, *fatigue* or *fascination* are not taken into account in the objective assessment, but are considered important in the subjective assessment.

Thus, the *fullness* indicator may be associated with the effect of compression of the dynamic range when the audio signal is amplified and the effect of variation in the nature of the transient process when the input signal level changes. This assessment is not taken into account in objective electrical indicators; the measurement technique has not been developed.

Although the subjective indicator loudness is related to the signal level (output power or voltage), when listening to different sound systems with the same electrical parameters it gives a different perception.

Loudness is often associated with fatigue or fascination. So a path that sounds exciting makes you want to listen to it at a higher volume, and a path that sounds tiring – at a low volume. The use of *fatigue* and *fascination* indicators makes it even more difficult to establish clear connections with objective characteristics, since no relationship has been established for them. Thus, these criteria are the most complex, since they combine a final assessment of the sound quality of the entire sound path and must be formed based on the analysis of a number of psychoacoustic characteristics.

The subjective indicator of *warmth*/*neutrality* is usually associated with the presence of a certain spectrum of nonlinear distortions and their mutual distribution. Thus, the predominance of the second harmonic and harmonics of even orders is usually assessed as some heat. The predominance of the third harmonic and harmonics of odd orders is assessed as aggressiveness or harshness. The neutral character is usually associated with a rapidly decreasing spectrum of nonlinear distortion products. A number of electrical parameters are measured using standard methods [4]-[10], although the levels of even and odd harmonics are not indicated separately.

Assessing *stage* is quite difficult. One can often find the wording about a close stage or a distant stage. This is probably due to the enlarged presentation of the sound image in the presence of dynamic range compression, due to which quiet sounds become louder and loud sounds quieter.

Micro-detail and *macro-detail* refer to the characteristics of reproducing quiet sounds against the background of loud ones and masking quiet

sounds with loud sounds. If the wording is similar, small and large sounds are assessed separately, since the effects of psychoacoustic masking of human hearing are involved. The basis of electrical phenomena in sound paths when these indicators change are nonlinearity effects that appear when exposed to signals with a complex spectrum. Standard methods [4]-[10] provide for assessment based on the impact of a pair of stationary signals with different frequencies and a certain level ratio, which has no relation to the real sound signal.

Thus, the presence of many cross-connections between parameters or their absence does not give an exact answer to the question of which electrical parameter needs to be improved to improve sound quality in a subjective assessment.

3 USING AN ARTIFICIAL INTELLIGENCE SYSTEM TO ADJUST PSYCHOACOUSTIC EVALUATION CRITERIA

Since recently, artificial intelligence (AI) systems are beginning to be introduced into scientific and educational fields [12], [13]. An interesting possibility is to obtain more information about the evaluation methods being studied based on the analysis of the results of a survey of respondents and the synthesis of simulation models for creating (processing) reference phonograms by modifying the structure of subroutines of sound data processors (chains). The possible structure of the AI system for surveying respondents is shown in Figure 4.

The AI system contains a library of test records that can be processed by distortion chains when generating a test task for a respondent by a test generation machine. The test generation machine is a working area and, based on the initial information received from the respondent through the respondent interface subsystem, generates a test task and transmits it through an electroacoustic transducer to the respondent. The electroacoustic converter must be a reference path consisting of a digital-to-analog converter, an amplification path and stereo headphones or an acoustically prepared room and speaker systems.

During research, the respondent reports to the AI system the results of a subjective assessment of sound quality through the respondent's interface. The response acceptance subsystem can accept the respondent's answer or add or change a test task to clarify the survey result.

The knowledge base stores initial information about the research, but in the process of acquiring knowledge, the knowledge base can be changed and supplemented. Thus, it is possible to identify frequently encountered criteria for psychoacoustic assessment and compare them with distortion introduction circuits.

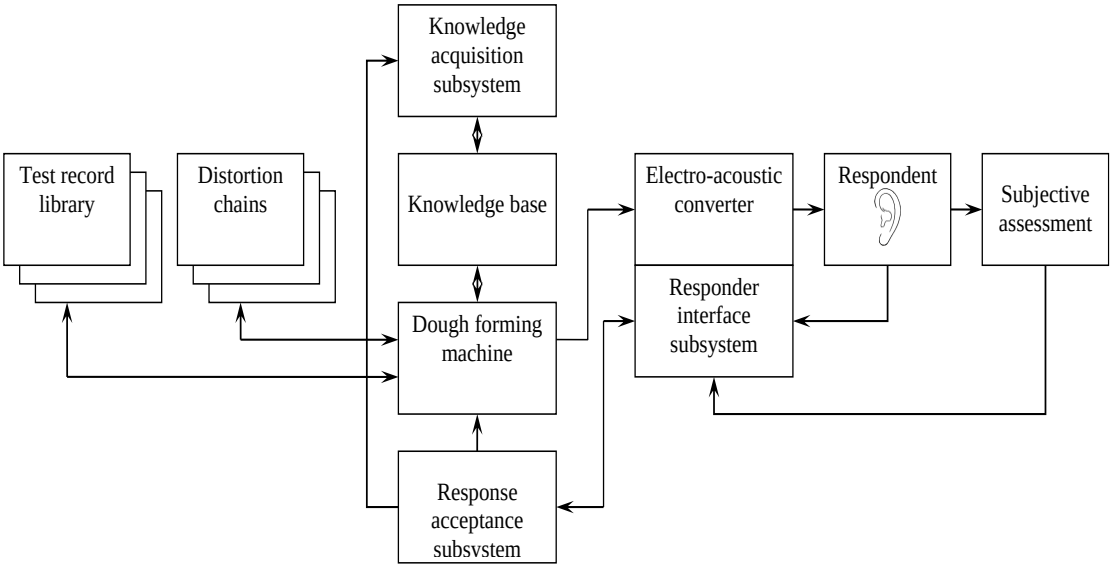


Figure 4: Scheme of using AI to survey respondents in order to clarify the psychoacoustic criteria for assessing the quality of sound paths.

The distortion injection chains can be changed during the test to clarify the degree of influence on the psychoacoustic perception of electrical characteristics.

The database of test records is selected according to the wishes of the respondent based on the respondent's interaction with the AI system. Based on feedback from the respondent, the database of test recordings can be supplemented, since recordings of different styles can change their psychoacoustic properties differently when processing distortion chains.

The result of the work of the AI system should be a knowledge base about the criteria for psychoacoustic assessment and the electrical characteristics that are compared with them.

3 CONCLUSIONS

It has become obvious that objective measurement methods [1-7] do not meet the requirements of listeners and the professional community, which requires clarification of existing assessment methods using psychoacoustic criteria.

Clarification of methods for objective assessment of sound paths by taking into account indicators related to psychoacoustic properties will improve their quality characteristics, increase listeners' interest in information, music, and educational programs by achieving a comfortable presentation.

In the audio engineering industry, the construction of sound processing devices based on a psychoacoustic assessment of their sound quality will increase the competitiveness of devices on the market compared to models that have similar functionality, but are built without taking into account psychoacoustic properties.

The use of a system based on AI at the stage of establishing sustainable mutual relations between psychoacoustic and electrical characteristics would make it possible to measure certain sets of sound device physical quantities or sound tract that would not just be numbers, but by those descriptions of the real sound quality with a psychoacoustic perception context.

It is possible that psychoacoustic criteria can be used to deliberately correct the nature of sound materials in advertising campaigns and public speaking, but this possibility requires additional research.

REFERENCES

[1] H. Fastl, "Psycho-Acoustics and Sound Quality," in Communication Acoustics, J. Blauert, Ed. Berlin, Heidelberg: Springer, 2005. pp. 157–167.

[2] H. Fastl, "Psychoacoustic basis of sound quality evaluation and sound engineering," in Proc. 13th Int. Congr. Sound Vibrat. (ICSV), vol. 1, pp. 59–74, 2006.

[3] H. Fastl and E. Zwicker, Psychoacoustics: Facts and Models, 3rd ed. Berlin, Heidelberg: Springer, 2007, 463 p.

[4] Naushniki stereofonicheskiye. Metody izmereniy. Mezhgosudarstvennyy standart GOST 28278-89 ["Stereophonic headphones. Measurement methods." Interstate Standard GOST 28278-89], Standartinform, Moscow, Russia, 2006, 23 p.

[5] Usiliteli signalov zvukovoy chastoty bytovyye. Obshchiye tekhnicheskiye. Gosudarstvennyy standart SSSR GOST 24388-88 ["Increased sound frequency signals. General technical." State Standard of the USSR GOST 24388-88], Izdatel'stvo standartov, Moscow, Russia, 1989, 11 p.

[6] Apparatura radioelektronnaya bytovaya. Metody izmereniya elektricheskikh parametrov usiliteley signalov zvukovoy chastoty. Gosudarstvennyy standart SSSR GOST 23849-87 ["The equipment is electronic household. Methods for measuring the electrical parameters of sound frequency signals." State Standard of the USSR GOST 23849-87], Izdatel'stvo standartov, Moscow, Russia, 1990, 66 p.

[7] Audio - i audiovizual'noye oborudovaniye. Komponenty tsifrovoy audioapparatury. Osnovnyye metody izmereniy zvukovykh kharakteristik. Chast' 1. Obshchiye polozheniya. Mezhgosudarstvennyy standart GOST IEC 61606-1-2014 ["Audio and audiovisual equipment. Components of digital audio equipment. Basic methods for measuring sound characteristics. Part 1. General provisions."], Standartinform, Moscow, Russia, 2016, 36 p.

[8] Audio - i audiovizual'noye oborudovaniye. Komponenty tsifrovoy audioapparatury. Osnovnyye metody izmereniy zvukovykh kharakteristik. Chast' 2. Bytovoye primeneniye. Mezhgosudarstvennyy standart GOST IEC 61606-2-2014 ["Audio and audiovisual equipment. Components of digital audio equipment. Basic methods for measuring sound characteristics. Part 2. Household application."], Standartinform, Moscow, Russia, 2016, 36 p.

[9] Audio - i audiovizual'noye oborudovaniye. Komponenty tsifrovoy audioapparatury. Osnovnyye metody izmereniy zvukovykh kharakteristik. Chast' 3. Professional'noye primeneniye. Mezhgosudarstvennyy standart GOST IEC 61606-3-2014 ["Audio and audiovisual equipment. Components of digital audio equipment. Basic methods for measuring sound characteristics. Part 3. Professional application."], Standartinform, Moscow, Russia, 2020, 42 p.

[10] Audio - i audiovizual'noye oborudovaniye. Komponenty tsifrovoy audioapparatury. Osnovnyye metody izmereniy zvukovykh kharakteristik. Chast' 4. Personal'nyy komp'yuter. Mezhgosudarstvennyy standart GOST IEC 61606-4-2014 ["Audio and audiovisual equipment. Components of digital audio equipment. Basic methods for measuring sound

characteristics. Part 4. Personal computer."], Standartinform, Moscow, Russia, 2018, 28 p.

[11] N.O. Abrosimova and M.S. Shushnov, “Sistema standartov izmereniya kachestva sovremennoy zvukovoy tekhniki” ["System of standards for measuring the quality of modern sound technology"], in Proc. Int. Sci.-Tech. Conf. Sovremennyye problemy telekommunikatsiy, Novosibirsk, Russia: SibGUTI, 2021, pp. 80–84.

[12] F. Luher, Iskusstvennyy intellekt: strategii i metody resheniya slozhnykh problem ["Artificial Intelligence: Strategies and Methods for Solving Complex Problems"], 4th ed. Moscow, Russia: Vil'yams, 2003, 863 p.

[13] I.O. Kotlyarova, “Tekhnologii iskusstvennogo intellekta v obrazovanii” ["Technologies of artificial intelligence in education"], Vestn. YUUrGU, Ser. Obrazovaniye. Pedagogicheskiye nauki, vol. 14, no. 3, pp. 69–82, 2022, doi: 10.14529/Ped220307.

Enhancement of Elgamal Cryptosystem; A Review

Paul Arhin and George Aggrey
Department of Computer Science & I.T, University of Cape Coast, Private Mail Bag, CC-123-1749, Cape Coast, Ghana
parhin@ucc.edu.gh, gaggrey@ucc.edu.gh

Keywords: ElGamal Cryptosystem, Hybrid Encryption, Hardware Enhancement, Software Enhancement.

Abstract: This article reviews a large body of research on optimizing the ElGamal cryptosystem for increased hardware and software efficiency. Areas of effort identified in the research include, hardware acceleration, hybrid systems combining ElGamal with symmetric cryptosystems, software advancements through adjustments to mathematical difficulties, and hybrids with other asymmetric cryptosystems. There is a clear demonstration that hybrid systems that combine ElGamal with symmetric encryption and software improvement through adjustments to the mathematical difficulties are highly impactful with about 95% score based on the amount of research done in the area. This indicates the significance of these domains for the further evolution of the ElGamal cryptosystem. Hybrid systems which seek to combine Elgamal with symmetric cryptosystems, such as AES, has become a very crucial area that needs a high level of concentration. This ensures both high level of security and greater enhancement to the time complexity. Through these changes, performances were improved and computational overhead were decreased by optimizing the underlying mathematical operations and combining Elgamal with other symmetric algorithms. The article proposes ChaCha20-Poly1305 as a potential for further improvement in hybrid encryption.

1 INTRODUCTION

Taher ElGamal created the ElGamal cryptosystem in 1985, and it is a foundation of public-key cryptography [1]. Because of its durability and reliance on the computationally challenging discrete logarithm issue, it is commonly used for secure key exchange and digital signatures [2]. The ElGamal cryptosystem has undergone a number of improvements throughout time to increase its security, effectiveness, and application.

These improvements, which address a variety of issues like performance optimization, attack resistance, and platform and application adaptation, have been made using both hardware and software. Specialized hardware has been utilized over the years to help improve the efficiency and security of the ElGamal cryptosystem [3].

Using Application-Specific Integrated Circuits (ASICs) and Field-Programmable Gate Arrays (FPGAs) to speed up cryptographic procedures are two notable initiatives in this field [3][4][5].

The goal of these implementations is to achieve low power consumption and high processing speed, which are essential for real-time applications and situations with limited resources, such as Internet of Things devices [6]. There has been a great advancement in the use of software to improve ElGamal cryptosystem [7]. Many of these enhancements entail applying mathematical adjustments to the fundamental algorithm, including streamlining the procedures of key creation, encryption, and decryption [8].

These improvements are meant to boost the cryptosystem's effectiveness and strengthen its defences against cryptographic assaults. The creation of hybrid encryption systems that combine ElGamal with other cryptographic algorithms to make use of their unique advantages is another example of software upgrades [9][10]. A good way of improving Elgamal efficiency has been through the use of asymmetric hybrids, which combine ElGamal with other public-key cryptosystems like RSA or ECC (Elliptic Curve Cryptography) [11][12].

The integration of various encryption layers in these hybrids improves security. A number of attempts have concentrated on merging ElGamal with symmetric encryption techniques, which provide further security advantages and fast data encryption [13]. The high computational and mathematical requirements of the Elgamal cryptosystem have significantly impacted its effectiveness. Complex

mathematical operations are required for key generation, encryption, and decryption, which can be time- and resource-consuming [14].

As a result, there is a need to increase its efficiency using a variety of techniques, such as hybrid approaches, algorithmic optimizations, and hardware acceleration. Although ElGamal has undergone substantial development, more research needs to be done in integrating alternative symmetric algorithms in conjunction with ElGamal for a more efficient cryptosystem.

2 ELGAMAL CRYPTOSYSTEM

The Elgamal cryptosystem, a widely used public-key cryptosystem scheme is dependent on the difficulty of computing discrete logarithms over finite fields [2].

It is a public key encryption scheme which means the public key is shared and used for encryption whiles the secret or private key is kept secret and used for decryption.

Like other cryptosystems, the Elgamal cryptosystem involves three stages: Key Generation, Encryption and Decryption [2][3]. Figure 1 illustrates this process.

2.1 Key Generation

The key generation process starts with the selection of a large prime number p, and then choosing a suitable generator g and then randomly selecting a private key x, and computing the public key h using the relation $h = g^x \bmod p$. The public key becomes (p, g, h) and the private key (x). This forms the foundation for secure communication during the encryption and decryption procedures.[3].

2.2 Encryption Process

In the Encryption process, the sender gets access to the recipients' public key (p, g, h) generated in the key generation stage. To encrypt the message m, the plaintext is converted into an integer such that, $0 \leq m < p$. A random integer y is chosen for the encryption. $c1$ is then computed as $c_1 = g^y \bmod p$, c_2 is computed as $c_2 = m \cdot h^y \bmod p$. The resulting ciphertext is the pair; Ciphertext (c_1, c_2) [2][3].

2.3 Decryption Process

In the decryption process, the recipient first receives the ciphertext, Ciphertext (c_1, c_2). The receiver then computes the shared secret s as $s = c_1^x \bmod p$. The modular inverse of s (s^{-1}) is computed by the recipient such that $s \cdot s^{-1} \equiv 1 \bmod p$. The recipient then goes ahead to derive the original message, $m = c_2 \cdot s^{-1} \bmod p$. The plain text is then derived.[3]

3 PERFORMANCE CHALLENGES OF ELGAMAL

Despite being widely used, the ElGamal cryptosystem has a number of significant performance issues.

The traditional Elgamal Cryptosystem is noted for its inefficiency in encrypting large message sizes. This is due to its inherent message expansion. This results in a ciphertext that is bigger than the size of the plaintext, roughly a double of the plaintext [15][16]. This will require more storage, more bandwidth and longer processing times.

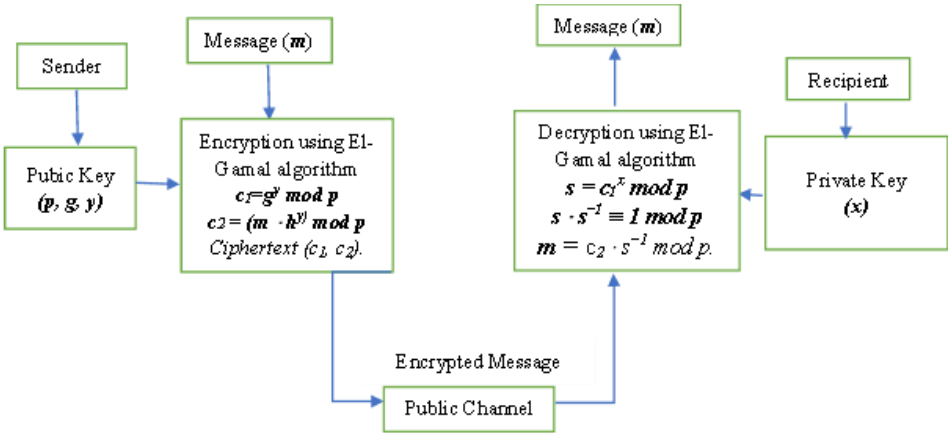


Figure 1: ElGamal Cryptosystem Procedure.

The computational complexity of key generation and the encryption/decryption procedures frequently reduces the efficiency of the cryptosystem. This is particularly troublesome for real-time applications with big databases [17].

The traditional ElGamal cryptosystem is dependent on the discrete logarithm problem. This makes it vulnerable to advancements in quantum computing capabilities. ElGamal's security presumptions might be broken by quantum algorithms [17][18].

4 ENHANCEMENTS OF ELGAMAL

According to [19], The ElGamal cryptosystem's resistance to attacks by quantum algorithms has been strengthened by the application of matrix ideas in key pair creation, making it appropriate for lower prime numbers and enhancing its security against brute-force attacks. This was done by using a square matrix as the private key to improve resistance against quantum logarithms attacks. With larger matrix sizes, there is a higher security against brute force attacks. Elliptic Curve Cryptography (ECC) was employed to upgrade ELGAMAL encryption scheme for better security.

When compared to normal ELGAMAL encryption, the suggested technique improves security and speed in both encryption and decryption duties while maintaining an extended cipher text size. It improves security and elapsed time in both encryption and decryption tasks over the conventional ELGAMAL encryption approach. A. Niemi and J. Teuhola [20] conducted a thorough study seeking to enhance Elgamal. The goal of the study is to improve the security of the Elgamal cryptosystem by using Burrows-Wheeler post-transformation with effective clustering and interpolative coding. This will boost the security of the Elgamal public-key algorithm.

An enhanced version of Elgamal was also developed by Hussein et al [15] for image encryption and decryption enhancement. The suggested method performs excellently across a range of assessment measures when tested on four distinct color photos. An improved ElGamal cryptosystem was used for image encryption, image decryption, and key pair generation in the methodology. To facilitate encryption and decryption, random bytes were created and pixels and generated bytes were subjected to an XOR process. Four color photos were used to test the approach, and the results were good.

For IoT devices, Mohan et al in 2020 [6] developed an improved Elgamal cryptosystem for secure Data Transfer in IoT networks. The updated ElGamal cryptosystem increases transmission efficiency while maintaining security levels comparable to the original. The study suggests a public-key cryptosystem (PKEIE) that is based on ElGamal and is effective in addressing issues such massive message encryption, integrity, authentication, non-malleability, and semantic security. With this approach, data of any size can be encrypted, something that the old ElGamal cryptosystem was not able to do. The suggested technique integrates integrity, authentication, and encryption using a single algorithm, making it appropriate for Internet of Things networks when compared to other ElGamal variant schemes based on security levels and performance. A. Pandey et al (2020) [21] proposed a methodology, which includes breaking down group ring elements into matrices over base rings, applying linear algebra techniques, and creating a cryptanalytic algorithm to effectively break the ElGamal-like cryptosystem put forth by Saba Inam and Rashid Ali. The security of the plan against selected ciphertext attacks (IND-CCA) and selected plaintext assaults (IND-CPA) is also examined by the authors. The authors enhanced the method of cryptanalysis, proved that Saba and Rashid's ElGamal cryptosystem was insecure, and created a cryptanalytic assault that might compromise the system.

An improved Exponential Elgamal Encryption Scheme with Additive Homomorphism was proposed by Zhou et al [22] in 2022. The process comprised evaluating the security of the established ElGamal encryption scheme, putting forth the exponential ElGamal system as a new encryption strategy, and verifying the effectiveness and security of the new scheme. When chosen plaintext was used as an attack vector, the ElGamal encryption technique was not sufficiently secure. The proposed enhanced Elgamal cryptosystem achieved a higher security.

Ranasinghe et al [23] presented a work on the generalization of the Elgamal public-key cryptosystem. The methodology is based on a proposal for an extension of the original ElGamal system that uses modular exponentiation twice during encryption and incorporates the prime factorization of the plaintext to improve the encryption process. The scheme’s encryption mechanism is enhanced.

Owolabi et al [24] also proposed a hybrid system of Elgamal and Blowfish with the focus of enhancing efficiency and security. The goal of the research is to increase encryption/decryption performance and data security of ElGamal by integrating the Blowfish and ElGamal cryptosystems. The result is the achievement of a more secure encryption technique with faster encryption and decryption speeds compared to the standalone El-Gamal algorithm.

Another work that combined Elgamal with another Symmetric algorithm, 3DES was conducted by Rachmawati et al [25] in the year 2018. The work explored a hybrid cryptographic strategy to address security concerns in data exchange over the internet by utilizing the asymmetric ElGamal algorithm and the symmetric Triple DES algorithm. Using a hybrid cryptography approach, the study combined the asymmetric ElGamal algorithm with the symmetric Triple DES algorithm. Triple DES, which requires three 56-bit keys, was utilized for both encryption and decryption. The Triple DES keys were encrypted using ElGamal, which uses a public key for encryption and a private key for decryption. Text files ending in.txt were the messages that needed to be encrypted.

The results show that, the Triple DES algorithm takes longer to encrypt and decrypt messages with larger volumes. The Triple DES and ElGamal algorithms require more time to encrypt data than to decrypt it. Still on hybrid of Elgamal with Symmetric algorithms, Rani et al [26] in their paper, “Implementation and comparison of hybrid encryption model for secure network using AES and Elgamal”, presented a hybrid cryptographic algorithm combining Elgamal and AES that enhances performance over using them alone.

This is done to improve the security, encryption/decryption time, and throughput. The new hybrid algorithm that integrates AES and Elgamal was implemented with the Java programming language, and the performance analysis conducted. The results indeed showed that, the hybrid algorithm that was designed to decrease the time of encryption and decryption and increase throughput indeed performed far better than the individual algorithms.

A proposed hybrid system that integrates the symmetric Hill Cipher 3*3 and the asymmetric Elgamal to secure instant messaging for Android was thoroughly conducted by Rachmawati et al (2019) [27].

5 RESULT AND ANALYSIS

Upon reviewing a number of articles on the improvement and enhancement of the Elgamal cryptosystem to yield better efficiency in both hardware and software, notable results were recognized in the modification of the Elgamal cryptosystem in several ways. These ways include Software enhancement through modifications to the mathematical complexities of Elgamal, Hardware enhancement, Hybrid Enhancement involving Elgamal and other asymmetric systems and Hybrid enhancement involving Elgamal and symmetric algorithms.

Software improvements are highly impactful with about 95% of research done in this area. This have mostly concentrated on altering the ElGamal algorithm's mathematical complexity and hybrid systems involving Elgamal and symmetric algorithms as well as Elgamal and other Asymmetric algorithms. in order to increase its effectiveness and security. Through these changes, performances were improved and computational overhead were decreased by optimizing the underlying mathematical operations.

These enhancements typically involve optimizing algorithms, reducing computational complexity, and improving security measures. Other studies have also looked into using hardware acceleration to increase the ElGamal cryptosystem's speed and efficiency. ElGamal is becoming more feasible for resource-constrained environments and high-throughput applications as implementations utilizing FPGAs and other hardware accelerators have shown notable increases in processing performance.

6 CONCLUSIONS

Given the benefits of symmetric algorithms for increasing efficiency, it is noteworthy that Google's high-speed symmetric algorithm, ChaCha20-Poly1305 [28], has received little to no attention thus far. A strong contender for improving the ElGamal cryptosystem, ChaCha20-Poly1305 has exceptional performance and security characteristics. Even though ChaCha20-Poly1305 has a lot of potential, the literature review points to a gap in its integration with ElGamal, indicating an area that has not yet been thoroughly investigated but has the potential to significantly increase encryption efficiency and speed.

Furthermore, the integration of AI technology is still not fully utilized in efforts to identify the greatest speed enhancements in Elgamal cryptosystem. With the development of AI, there is a significant possibility to improve security procedures, anticipate weaknesses, and optimize operations. ElGamal cryptosystem has not been fully optimized by using AI applications in hardware, including Tensor Processing Units (TPUs), and software. Should this research gap be filled, it has the potential to greatly advance the field of cryptography technologies as AI integration in software enhancements as well as hardware accelerators is lacking.

Subsequent investigations ought to concentrate on merging ChaCha20-Poly1305 with ElGamal in order to create a hybrid cryptosystem that blends ElGamal's strong security with ChaCha20-Poly1305's fast performance. It should also investigate AI methods to improve security protocols [29], optimize encryption procedures, and anticipate future flaws in the ElGamal cryptosystem [30].

It is also crucial to include AI technologies into software programs and hardware accelerators like TPUs in order to obtain better encryption efficiency and performance. By filling in these gaps, the study can make a substantial contribution to the creation of an ElGamal cryptosystem that is more effective and safer while using the most recent developments in symmetric encryption and artificial intelligence.

REFERENCES

[1] H. Hajaje et al., “CPC-H2: Convolution Power-based Cryptosystem and Digital Signature,” 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1–7, 2021, [Online]. Available: https://doi: 10.1109/CCECE53047.2021.9569033

[2] R. Ranasinghe and P. Athukorala, "A generalization of the ElGamal public-key cryptosystem," Journal of Discrete Mathematical Sciences and Cryptography, vol. 25, pp. 2395–2403, 2021. [Online]. Available: https://doi: 10.1080/09720529.2020.1857902.

[3] S. Kim and R. Kyung, "Study on Modified Public Key Cryptosystem Based on ElGamal and Cramer-Shoup Cryptosystems," 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), pp. 280–284, 2023, [Online]. Available: https://doi: 10.1109/CCWC57344.2023.10099297.

[4] A. Kurd and N. Besli, "Analysis of the Cryptography Methods for Design of Crypto-Processor," 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1–7, 2020, [Online]. Available: https://doi: 10.1109/HORA49412.2020.9152929.

[5] S. Guilley, L. Sauvage, J. Danger, T. Graba, and Y. Mathieu, "Evaluation of Power-Constant Dual-Rail Logic as a Protection of Cryptographic Applications in FPGAs," 2008 Second International Conference on Secure System Integration and Reliability Improvement, pp. 16–23, 2008, [Online]. Available: https://doi: 10.1109/SSIRI.2008.31.

[6] M. Mohan, M. Kavithadevi, and J. V., "Improved ElGamal Cryptosystem for Secure Data Transfer in IoT Networks," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), pp. 295–302, 2020, [Online]. Available: https://doi: 10.1109/I-SMAC49090.2020.9243407.

[7] H. I. Hussein and W. M. Abduallah, "An efficient ElGamal cryptosystem scheme," International Journal of Computers and Applications, vol. 43, pp. 1088–1094, 2019, [Online]. Available: https://doi: 10.1080/1206212X.2019.1678799.

[8] G. Kim and S. Li, "Decryption speed up of ElGamal with composite modulus," PLoS ONE, vol. 15, 2020, [Online]. Available: https://doi: 10.1371/journal.pone.0240248.

[9] M. Iavich, S. Gnatyuk, E. Jintcharadze, Y. Polishchuk, and R. Odarchenko, "Hybrid Encryption Model of AES and ElGamal Cryptosystems for Flight Control Systems," 2018 IEEE 5th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), pp. 229–233, 2018, [Online]. Available: https://doi: 10.1109/MSNMC.2018.8576289.

[10] D. Rachmawati, M. A. Budiman, and M. I. Wardhono, "Hybrid Cryptosystem for Image Security by Using Hill Cipher 4x4 and ElGamal Elliptic Curve Algorithm," 2018 IEEE International Conference on Communication, Networks and Satellite (Comnetsat), pp. 49–54, 2018, [Online]. Available: https://doi: 10.1109/COMNETSAT.2018.8684121.

[11] M. Enriquez, D. W. Garcia, and E. R. Arboleda, "Enhanced Hybrid Algorithm of Secure and Fast Chaos-based, AES, RSA and ElGamal Cryptosystems," Indian Journal of Science and Technology, vol. 10, pp. 1–14, 2017, [Online]. Available: https://doi: 10.17485.IJST.2017.V10I27.105001.

[12] M. Bie, W. Li, T. Chen, L. Nan, and D. Yang, "An energy-efficient reconfigurable asymmetric modular cryptographic operation unit for RSA and ECC," Frontiers of Information Technology & Electronic Engineering, vol. 23, pp. 134–144, 2022, [Online]. Available: https://doi: 10.1631/FITEE.2000325

[13] Y. Luo, X. Ouyang, J. Liu, and L. Cao, "An Image Encryption Method Based on Elliptic Curve ElGamal Encryption and Chaotic Systems," IEEE Access, vol. 7, pp. 38507–38522, 2019, [Online]. Available: https://doi: 10.1109/ACCESS.2019.2906052.

[14] P. Sharma, S. Sharma, and R. Dhakar, "Modified Elgamal Cryptosystem Algorithm (MECA)," 2011 2nd International Conference on Computer and Communication Technology (ICCCT-2011), pp. 439–443, 2011, [Online]. Available: https://doi: 10.1109/ICCCT.2011.6075141.

[15] H. I. Hussein, R. J. Mstafa, A. Mohammed, and Y. M. Younis, "An Enhanced ElGamal Cryptosystem for Image Encryption and Decryption," 2022 International Conference on Computer Science and Software Engineering (CSASE), pp. 337–342, 2022, [Online]. Available: https://doi: 10.1109/CSASE51777.2022.9759643.

[16] M. Hwang, C. Chang, and K. Hwang, "An ElGamal-Like Cryptosystem for Enciphering Large Messages," IEEE Transactions on Knowledge and Data Engineering, vol. 14, pp. 445–446, 2002, [Online]. Available: https://doi: 10.1109/69.991728.

[17] T. Prantl et al., "Performance Evaluation for a Post-Quantum Public-Key Cryptosystem," 2021 IEEE International Performance, Computing, and Communications Conference (IPCCC), pp. 1–7, 2021, [Online]. Available: https://doi: 10.1109/IPCCC51483.2021.9679412.

[18] J. Jianwei, H. Wang, H. Zhang, S. Wang, and J. Liu, "Cryptanalysis of an ElGamal-Like Cryptosystem Based on Matrices Over Group Rings," Communications in Computer and Information Science, 2018, [Online]. Available: https://doi: 10.1007/978-981-13-5913-2_16.

[19] Maxrizal, S. Irawadi, and Sujono, "Discrete Logarithmic Improvement for ElGamal Cryptosystem Using Matrix Concepts," 2020 8th International Conference on Cyber and IT Service Management (CITSM), pp. 1–5, 2020, [Online]. Available: https://doi: 10.1109/CITSM50537.2020.9268832.

[20] A. Niemi and J. Teuhola, "Burrows-Wheeler post-transformation with effective clustering and interpolative coding," Software: Practice and Experience, vol. 50, pp. 1858–1874, 2020, [Online]. Available: https://doi: 10.1002/spe.2873.

[21] A. Pandey, I. Gupta, and D. Singh, "Improved cryptanalysis of a ElGamal Cryptosystem Based on Matrices Over Group Rings," Journal of Mathematical Cryptology, vol. 15, pp. 266–279, 2020, [Online]. Available: https://doi: 10.1515/jmc-2019-0054.

[22] R. Zhou and Z. Lin, "An Improved Exponential ElGamal Encryption Scheme with Additive Homomorphism," 2022 International Conference on Blockchain Technology and Information Security (ICBCTIS), pp. 25–27, 2022, [Online]. Available: https://doi: 10.1109/ICBCTIS55569.2022.00017

[23] R. Ranasinghe and P. Athukorala, "A generalization of the ElGamal public-key cryptosystem," Journal of Discrete Mathematical Sciences and Cryptography, vol. 25, pp. 2395–2403, 2021, [Online]. Available: https://doi: 10.1080/09720529.2020.1857902.

[24] O. Y. Owolabi, P. B. Shola, and M. B. Jibrin, "Improved Data Security System Using Hybrid Cryptosystem," International Journal of Scientific Research in Science, Engineering and Technology, vol. 3, pp. 90–93, 2017. [Online]. Available: [Online]. Available: https://www.academia.edu/89014842/Improved_Data_Security_System_Using_Hybrid_Cryptosystem

[25] D. Rachmawati, A. Harahap, and R. N. Purba, "A hybrid cryptosystem approach for data security by using triple DES algorithm and ElGamal algorithm," IOP Conference Series: Materials Science and Engineering, vol. 453, 2018, [Online]. Available: https://doi: 10.1088/1757-899X/453/1/012018

[26] S. Rani and H. Kaur, "Implementation and comparison of hybrid encryption model for secure network using AES and Elgamal," International Journal of Advanced Research in Computer Science, vol. 8, pp. 254–258, 2017, [Online]. Available: https://doi: 10.26483/ijarcs.v8i3.2990.

[27] D. Rachmawati, A. Sharif, and Ericko, "Hybrid Cryptosystem Combination Algorithm Of Hill Cipher 3x3 and Elgamal To Secure Instant Messaging For Android," Journal of Physics: Conference Series, vol. 1235, 2019, [Online]. Available: https://doi: 10.1088/1742-6596/1235/1/012074

[28] J. P. Degabriele et al., "The Security of ChaCha20-Poly1305 in the Multi-User Setting," Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, 2021. [Online]. Available: https://doi: 10.1145/3460120.3484814

[29] Li, B., Feng, Y., Xiong, Z., Yang, W., & Liu, G. (2021). "Research on AI Security Enhanced Encryption Algorithm of Autonomous IoT Systems." Information Sciences, vol. 575, pp. 379-398. [Online]. Available: https://doi: 10.1016/J.INS.2021.06.016

[30] Blackledge, J., & Mosola, N. (2020). "Applications of Artificial Intelligence to Cryptography." Transactions on Machine Learning and Artificial Intelligence, vol. 8, pp. 21-60. [Online]. Available: https://doi: 10.14738/tmlai.83.8219.

SECTION 2

Data Analysis and Processing

Proceedings of the 12th International Conference on Applied Innovations in IT (ICAIIT)

NOVEMBER 2024

Classification Tree Model for Determining Society Unsafety Factors Convicted

Olha Kovalchuk[1], Nataliia Chudyk[1], Tetiana Drakokhrust[1], Valerii Kaniuka[2], Roman Ivanytskyy[3], and Larysa Khokhlova[3]

[1] *Theory of Law and Constitutionalism Department, West Ukrainian National University, 11 Lvivska Str., 46009 Ternopil, Ukraine*

[2] *Department of Security and Law Enforcement, West Ukrainian National University, 11 Lvivska Str., 46009 Ternopil, Ukraine*

[3] *Department of Physics and Mathematics, Ternopil Volodymyr Hnatiuk National Pedagogical University, 2 Maxyma Kryvonosa. Str., 46009 Ternopil, Ukraine*

o.kovalchuk@wunu.edu.ua, n.chudyk@wunu.edu.ua, t.drakokhrust@wunu.edu.ua, valeriikaniuka@gmail.com, romik_iv@ukr.net, larysa_khokhlova@tnpu.edu.ua

Keywords: Justice, Decision Trees, Classification Modeling, Society Safety, Judicial Decision-Making, Informational Support.

Abstract: The global rise in crime rates is a result of the ineffectiveness of the traditional punishment system and the need to develop new approaches to applying relevant punitive measures to criminals, taking into account their level of danger to society. This paper presents a classification Decision Trees model to identify significant factors influencing the objective level of danger a convicted person poses to society. Applied research was conducted based on real data (individual characteristics and criminal history records) of 2,052 convicts serving sentences in correctional facilities in Ukraine. It was found that the most significant predictors for determining the level of danger to society were the number of previous incarcerations and the number of suspended sentences. Assessing the level of societal danger posed by convicts is one of the key aspects of upholding the principles of fair justice. The obtained results can serve as informational support for judicial decisions, ensuring a balance between societal protection and successful offender reintegration.

1 INTRODUCTION

The escalating prevalence of criminal activities across numerous nations has sparked grave apprehensions within societal spheres and posed formidable challenges for law enforcement entities and judicial frameworks. This burgeoning trend poses a palpable threat to society's safety, adversely impacting economic progress, eroding citizens' sense of security, and culminating in an erosion of trust in law enforcement agencies and the judicial branch [1]. The current trajectory necessitates a comprehensive analysis of the underlying factors fueling such negative dynamics and the formulation of effective strategies to counteract this phenomenon at both the international and national levels.

Concurrently, the combat against criminal elements demands a multifaceted approach, which transcends merely augmenting the efficacy of law enforcement agencies but also necessitates addressing the "prison paradox", wherein an escalation in the incarcerated population does not yield a significant reduction in crime rates and incurs additional financial burdens [2].

Society must acknowledge that not all perpetrators are hardened and irredeemable. Frequently, individuals engage in illicit acts due to a confluence of circumstances, recklessness, or the influence of a deleterious environment. In such instances, it is judicious to differentiate between offenders who do not pose a substantial threat to society and entrenched criminal elements. Furnishing prospects for resocialization and rehabilitation for the former category alleviates the strain on the penitentiary system and paves the way for their reintegration into a law-abiding society [3]. The notion of affording a chance for rectification to certain categories of offenders is pertinent and justifiable.

An objective evaluation of the level of danger a suspect (convict/prisoner) poses to society is a pivotal element in upholding the rule of law, ensuring justice, striking a balance of interests, and enhancing the efficacy of the judicial system [4]. Such an assessment is predicated upon a comprehensive analysis of various individual characteristics to determine an individual's proclivity for recidivism, their social adaptability, and the viability of successful resocialization upon release.

2 RELATED WORKS

The problem of the court's imposition of a fair punishment for a crime has recently attracted the attention of many scholars and practitioners in the legal field [5, 6]. Such punishment should be proportionate to the gravity of the crime committed and take into account all the circumstances and personal characteristics of the accused [7]. The limited number of studies on these issues mostly considers the mental and psychological factors affecting the level of societal danger of convicts [8]. Researchers A. Karlsson and A. Hakansson determined that in addition to well-known risk factors such as male gender and young age, other factors increase the likelihood of offenses. These include the use of amphetamines and injectable drugs, previous convictions for violent and property crimes, as well as homelessness and psychiatric problems [9]. The authors A. Pękala-Wojciechowska et al. proved that physical and mental health is a serious obstacle to the social reintegration of those released from prison [10]. Scholars C. Webster and M. Qasim argued that social and economic problems lead men to commit criminal offenses [11]. However, such studies are partial, and the study of factors from criminal records of previous convictions that are associated with increased societal danger of convicts is extremely rare [12, 13] and requires diverse study and a multidisciplinary approach.

3 METHODOLOGY

One of the key factors in making informed judicial decisions regarding the type and duration of punishment, the possibility of applying a suspended sentence, parole, or probation participation is assessing the level of danger of convicts to society. The empirical basis of our study consisted of real data on 2,052 individuals sentenced to serve their sentences in penal institutions in Ukraine. For applied research, we used the following variables (individual characteristics of inmates and information about their previous convictions):

- Sex: 1 – male, 2 – female;
- Age: 1 – under 18 years, 2 – from 18 to 30 years; 3 – from 30 to 45 years; 4 – age 45 or older;
- AAP – age at the time of the first conviction to the imprisonment: 1 – under 18 years, 2 – from 18 to 30 years; 3 – from 30 to 45 years; 4 – age 45 or older;
- AAS – age at the time of the first conviction (suspended sentence or the imprisonment): 1 – under 18 years, 2 – from 18 to 30 years; 3 – from 30 to 45 years; 4 – age 45 or older;
- MS – marital status: 1 – single, 2 – married;
- ED – education: 0 – incomplete secondary, 1 – secondary, 2 – special secondary, 3 – incomplete higher, 4 – higher;
- PR – place of residence to the actual degree of punishment: 0 – rural area, 1 – urban area;
- TE – type of employment at the time of conviction to the imprisonment: 0 – unemployed, 1 – part-time, 2 – full-time;
- ED – availability of early dismissals: 0 – no,1 – yes;
- MD – motivation for dismissal: 0 – no,1 – yes
- and numerical variables:
- RC – number of the imprisonment;
- SC – number of suspended convictions;

RR – level of danger of convicts to society (low, moderate, high) is the dependent variable, and the other – is the independent variable.

To identify the factors that influence the level of danger to society, we utilized classification using Decision Trees [14].

Formal Decision-Making Criteria: the decision-making process is based on a classification tree method using these specific variables. The decision tree recursively partitions data based on these variables to categorize convicts into three threat levels: low, moderate, and high.

3.1. Tree-Based Predictive Modeling

Decision tree learning is a widely adopted technique in statistics and machine learning domains. The Concept Learning System laid the foundation for all decision tree induction algorithms. These algorithms construct decision trees by employing methods from regression and correlation analysis. One prominent example is the CART (Classification and Regression Trees) algorithm, which recursively partitions the data into two child branches. The splitting criterion

for each branch depends on the number of data points it represents. The split is performed based on the feature that exhibits the highest correlation with the target variable for that branch.

A decision tree has a special root node from which all other nodes can be reached. The terminal nodes are called leaves. Each level of the tree represents a decision point, where a condition is evaluated, and the branches emanating from that node correspond to the possible outcomes of that condition. Essentially, at each decision point, the data instances are sorted such that each instance belongs to exactly one branch. This recursive partitioning process continues until a leaf node is reached, resulting in a hierarchical division of the data into smaller subsets. The number of possible ways to represent decisions can be substantial. Therefore, the choice of the splitting method depends on the type of feature being evaluated and the operation used for condition testing.

This study aims to identify the significant factors for the prisoner's distribution into groups (high, moderate, low) according to the degree of the society's danger. Such studies are being conducted in Ukraine for the first time.

The classification tree method is a flexible tool for predicting the class membership of observations. It provides the ability to classify simultaneously based on multiple variables and in various ways, which facilitates analysis and increases the reliability of results.

3.2. Mathematical Formulation of the Problem

Let: $D = \{(x_1,y_1), (x_2,y_2), ..., (x_n,y_n)\}$ be the training dataset of n observations where $x_i = (x_{i1}, x_{i2}, ..., x_{im}) \in X$ is the vector of m features (individual characteristics of inmates and information about their previous convictions) $y_i \in Y = \{1,2,3\}$ is the class label (level of societal danger: low, moderate, high).

The goal is to find a decision tree function $T: X \rightarrow Y$ that minimizes the misclassification error:

$$T^* = \underset{T \in \mathcal{T}}{\operatorname{argmin}} \sum_i L(T(x_i), y_i), \quad (1)$$

where $\mathcal{T}$ is the set of all possible decision trees $L(\hat{y},y)$ is the 0-1 loss function: $L(\hat{y},y) = 0$ if $\hat{y} = y$, $L(\hat{y},y) = 1$ if $\hat{y} \neq y$.

At each node t, the optimal split s^* is chosen to maximize the reduction in impurity:

$$s^* = \underset{s \in S}{\operatorname{argmax}} \Delta I(s,t), \quad (2)$$

where $\Delta I(s,t) = I(t) - p\_L \cdot I(t\_L) - p\_R \cdot I(t\_R)$, S is the set of all possible splits $I(t)$ is the impurity measure (Gini or entropy) at node t, p_L, p_R are the proportions of samples going to left/right child nodes t_L, t_R are the left and right child nodes.

For entropy calculation at node t:

$$I(t) = -\sum_k p(k|t) \log_2 p(k|t), \quad (3)$$

where $k \in Y = \{1,2,3\}$ – class labels (levels of societal danger), $p(k|t)$ – the proportion of samples belonging to class k at node t.

Then the information gain for split s at node t is:

$$IG(s,t) = -\Sigma_i (n_i/n) I(t_i), \quad (4)$$

where n – is the total number of samples at node t, n_i – the number of samples in child node t_i, $I(t_i)$ – is the entropy of child node i.

The process continues recursively until a stopping criterion is met [15].

4 RESULTS AND DISCUSSION

We applied Tree-Based Predictive Modeling [16] to determine the most significant predictors (individual characteristics of inmates and information about their previous convictions) that influence the level (low, moderate, high) of danger posed by convicts to society. The results of the conducted classification are presented in the form of a tree structure, which branches into left and right branches, each containing 11 nodes (Table 1). The interpretation of the results is significantly simplified by using the classification tree graph (Fig. 1).

In the Node column, the node number is indicated, while in the Size of Node column, the number of objects in the corresponding node is shown.

The left branch contains nodes 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22. The right branch consists of nodes 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23. At the first node, the number of convicts with a high level of danger to society is 492, with a moderate level – 552, and 1,008 convicts are classified as having a low level of danger to society. From node 1, two branches (right and left) emerge with corresponding nodes 2 and 3.

The condition for separating convicts into nodes 2 and 3 is as follows: if the value of RC (number of imprisonments) is less than or equal to 2.5, the level of danger to society posed by the convict is low. 1,008 convicts are classified as having a low level of danger to society, 264 – moderate, and 48 – high.

Table 1: The classification tree structure.

Node No.	Tree structure Dependent variable: RR								
	Left branch	Right branch	Size of node	N in class high	N in class moderate	N in class low	Selected category	Split variable	Split constant
1	2	3	2052	492	552	1008	low	RC	2.5
2	4	5	1320	48	264	1008	low	SC	1.5
4	6	7	1105	0	97	1008	low	RC	1.5
6			796	0	0	796	low		
7	8	9	309	0	97	212	low	SC	0.5
8			212	0	0	212	low		
9			97	0	97	0	moderate		
5	10	11	215	48	167	0	moderate	SC	3.5
10	12	13	185	18	167	0	moderate	SC	2.5
12			136	0	136	0	moderate		
13	14	15	49	18	31	0	moderate	RC	1.5
14			31	0	31	0	moderate		
15			18	18	0	0	high		
11			30	30	0	0	high		
3	16	17	732	444	288	0	high	RC	4.5
16	18	19	454	166	288	0	moderate	SC	1.5
18	20	21	330	42	288	0	moderate	SC	0.5
20			219	0	219	0	moderate		
21	22	23	111	42	69	0	moderate	RC	3.5
22			69	0	69	0	moderate		
23			42	42	0	0	high		
19			124	124	0	0	high		
17			278	278	0	0	high		

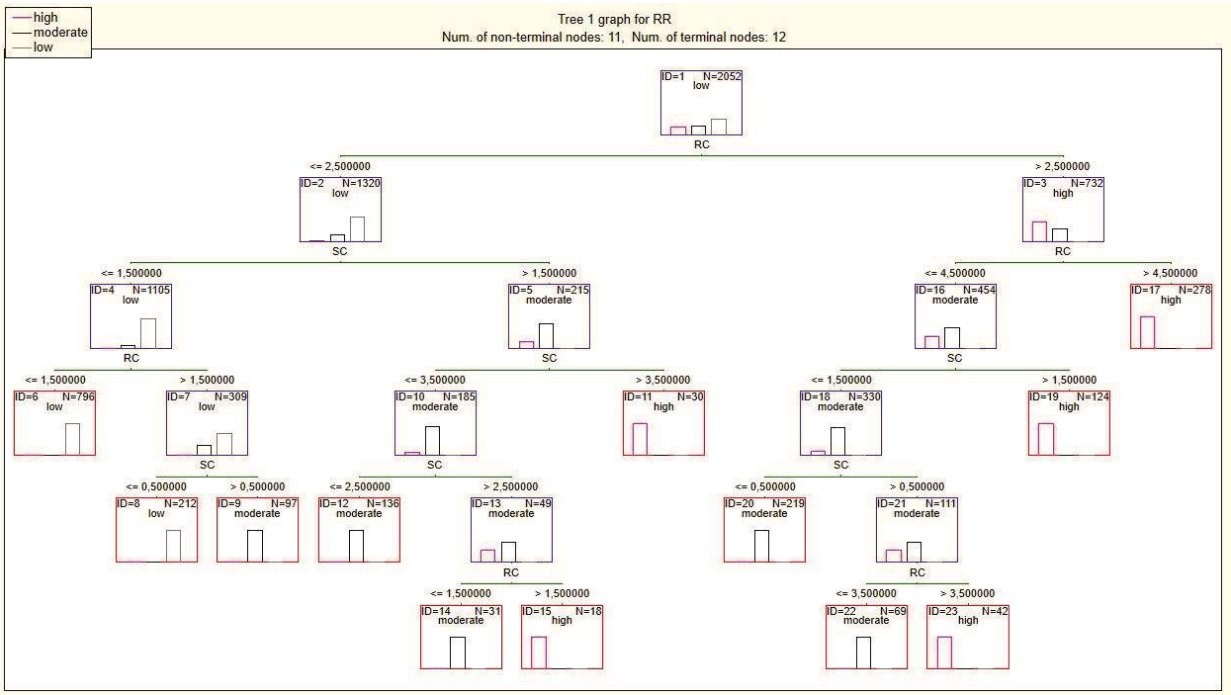


Figure 1: Classification tree.

At nodes 4 and 5, convicts are distributed as follows: if the value of SC (number of suspended convictions) is less than or equal to 1.5, the level of danger to society posed by the convict is low. The classification results at this stage: are 1,008 convicts with a low level of danger to society, 97 – with a moderate level.

For the 1,105 convicts assigned to the group with SC <= 1.5 (node 4), the distribution occurs at nodes 6 and 7 according to the following rule: if RC <= 1.5, then such convicts are finally classified as posing a low level of danger to society (node 6). Otherwise (node 7), the distribution of convicts into groups occurs as follows: if SC <= 0.5 (node 8), the convicts are finally assigned to the group with a low level of danger to society. Otherwise, they are classified as posing a moderate level of danger to society (node 9). Similarly, convicts are distributed into groups based on the level of danger to society at other nodes.

As a result of the conducted classification, out of the 2,052 convicts analyzed, 492 (23.98%) were classified as posing a high level of threat to society, 552 (26.90%) as posing a moderate level of threat, and 1,008 (49.12%) convicts were deemed to have a low level of danger to society. All objects were classified correctly. This is evidenced by the classification matrix (Fig. 2) and the constructed classification matrix diagram (Fig. 3). The classification results showed that the most significant predictors for determining the level of danger to society were the number of previous incarcerations and the number of suspended sentences.

Classification matrix
Dependent variable: RR
Options: Categorical response, Analysis sample

	Observed	Predicted high	Predicted moderate	**Predicted low**	Row Total
Number	high	492			492
Column Percentage		100.00%	0.00%	0.00%	
Row Percentage		100.00%	0.00%	0.00%	
Total Percentage		23.98%	0.00%	0.00%	23.98%
Number	moderate		552		552
Column Percentage		0.00%	100.00%	0.00%	
Row Percentage		0.00%	100.00%	0.00%	
Total Percentage		0.00%	26.90%	0.00%	26.90%
Number	low			1008	1008
Column Percentage		0.00%	0.00%	100.00%	
Row Percentage		0.00%	0.00%	100.00%	
Total Percentage		0.00%	0.00%	49.12%	49.12%

Figure 2: Classification matrix.

The obtained estimates confirm the results of our previous studies [4, 12] and scientific investigations by other researchers in this field [17, 18]. Moreover, our results correlate with contemporary criminological theories and align with existing knowledge about recidivism risk factors [19, 20].

In this study, we employed the classification tree method to analyze data on 2,052 convicts serving sentences in correctional facilities in Ukraine. The variables included in the analysis were gender, age, age at first conviction, marital status, education, place of residence, employment, parole availability, and motivation for release, as well as the number of previous incarcerations and suspended sentences. The dependent variable was the level of danger posed by the convict to society (low, moderate, high).

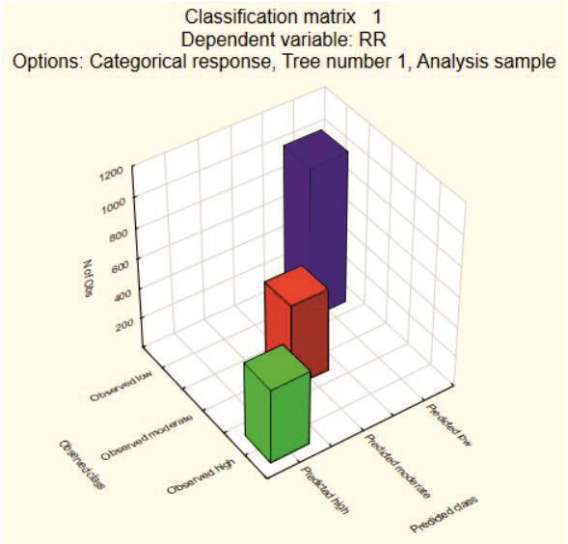


Figure 3: 3-D diagram of the classification matrix.

The integrated threat assessment criterion is primarily based on the number of imprisonments and the number of suspended sentences. The method creates a hierarchical decision tree where:

- ≤ 2–3 imprisonments suggest low threat;
- ≤ 1–2 suspended sentences indicate lower societal risk.

Combinations of these factors determine the final threat categorization.

These findings have practical implications for the digitalization of the judicial system and can be used to inform judicial decision-making regarding the type and duration of punishment, the possibility of applying a suspended sentence, parole, or probation participation based on the assessment of recidivism risk to balance societal protection and successful reintegration of offenders.

5 CONCLUSIONS

Assessing the level of danger that convicts pose to society is crucial for making informed judicial decisions regarding the type and duration of

punishment, and the possibility of applying a suspended sentence, parole, or probation participation. The study employed the classification tree method to analyze data from 2,052 convicts serving sentences in correctional facilities in Ukraine. A range of variables were utilized, including gender, age, age at first conviction, marital status, education, place of residence, employment status, availability of parole, motivation for release, as well as the number of previous incarcerations and suspended sentences. The classification results revealed that the most significant predictors for determining the level of danger to society were the number of previous incarcerations and the number of suspended sentences.

These findings have practical implications for the digitalization of the judicial system and can be used to inform judicial decisions regarding the type and duration of punishment, the possibility of applying a suspended sentence, parole, or probation participation based on the assessment of recidivism risk to strike a balance between societal protection and successful reintegration of offenders.

Promising avenues for future research in this area include incorporating additional variables, such as psychological characteristics, history of substance abuse, traumatic past events, and others, to deepen the understanding of factors influencing the risk posed by convicts to society.

REFERENCES

[1] L.J. Leonard, "Can Restorative Justice Provide a Better Outcome for Participants and Society than the Courts?," Laws, vol. 11:14, 2022.

[2] D. Stemen, "The Prison Paradox: More Incarceration Will Not Make Us Safer," New York: Vera Institute of Justice, 2017, 12 p. Available online: https://www.vera.org.

[3] O.W. Letlape, and M. Dube, "The Parole Officials' Views Concerning the Parole System in Rehabilitating Offenders: Experiences from Brits Community Correction Centre in South Africa," Soc. Sci., vol. 12:410, 2023.

[4] O. Kovalchuk, M. Kasianchuk, M. Karpinski, and R. Shevchuk, "Decision-Making Supporting Models Concerning the Internal Security of the State," International Journal of Electronics and Telecommunications, vol. 69(2), 2023, pp. 301–307.

[5] O. Arandjelović, "Crime and Punishment: A Rethink," Philosophies, vol. 8:47, 2023.

[6] K. Bullock, and A. Bunce, "The prison don't talk to you about getting out of prison': On why prisons in England and Wales fail to rehabilitate prisoners," Criminol. Crim. Justice, vol. 20, 2020, pp. 111–127.

[7] B. Agozino, "Reparative justice: The final stage of decolonization," Punishm. Soc., vol. 23, 2021, pp. 613–630.

[8] I. Testoni, I. Nencioni, M. Arbien, E. Iacona, F. Marrella, V. Gorzegno, C. Selmi, F. Vianello, A. Nava, A. Zamperini, et al., "Mental Health in Prison: Integrating the Perspectives of Prison Staff," Int. J. Environ. Res. Public Health, vol. 18:11254, 2021.

[9] A. Karlsson, and A. Håkansson, "Crime-Specific Recidivism in Criminal Justice Clients with Substance Use—A Cohort Study," Int. J. Environ. Res. Public Health, vol. 19:7623, 2022.

[10] A. Pękala-Wojciechowska, A. Kacprzak, K. Pękala, M. Chomczyńska, P. Chomczyński, M. Marczak, R. Kozłowski, D. Timler, A. Lipert, A. Ogonowska et al., "Mental and Physical Health Problems as Conditions of Ex-Prisoner Re-Entry," Int. J. Environ. Res. Public Health, vol. 18:7642, 2021.

[11] C. Webster, and M. Qasim, "The Effects of Poverty and Prison on British Muslim Men Who Offend," Soc Sci., vol. 7:184, 2018.

[12] K.M. Berezka, O.Ya. Kovalchuk, S.V. Banakh, S.V. Zlyvko, and R. Hrechaniuk, "A Binary Logistic Regression Model for Support Decision Making in Criminal Justice," Folia Oeconomica Stetinensia, vol. 22(1), 2022, pp. 1–17.

[13] C.C. Onyeneke, and A.H. Karam, "An Exploratory Study of Crime: Examining Lived Experiences of Crime through Socioeconomic, Demographic, and Physical Characteristics," Urban Sci., vol. 6, 43, 2022.

[14] M. Sharma, "Classification using Decision Trees," Medium, May 23, 2022. Available online: https://medium.com/@muditbits/classification-using-decision-trees-6b01939e01a5.

[15] C.A. Shanon, "Mathematical Theory of Communication," The Bell Systems Technical Journal, vol. 27:1948, pp. 379–423.

[16] TIBCO Statistica, "User's Guide," 2024. Available online: https://docs.tibco.com/pub/stat/14.0.0/doc/html/UsersGuide/GUID-B3E7379A-0459-4650-A614-1D8974325AB1.html.

[17] M.R. Kamaluddin, N.A. Mahat, G.A. Mat Saat, A. Othman, I.L. Anthony, S. Kumar, S. Wahab, S. Meyappan, B. Radhakrishnan, and F. Ibrahim, "The Psychology of Murder Concealment," Acts. Int. J. Environ. Res. Public Health, vol. 18:3113, 2021.

[18] S.-C. Hsu, K.-Y. Chen, C.-P. Lin, and W.-H. Su, "Knowledge Development Trajectories of Crime Prevention Domain: An Academic Study Based on Citation and Main Path Analysis," International Journal of Environmental Research and Public Health, vol. 19(17):10616, 2022.

[19] R. Yu, N. Langstrom, M. Forsman, A. Sjolander, S. Fazel, and Ya. Molero, "Associations between prisons and recidivism: A nationwide longitudinal study," PLoS ONE, vol. 17:e0267941, 2022.

[20] D. Yukhnenko, N. Blackwood, and S. Fazel, "Risk factors for recidivism in individuals receiving community sentences: A systematic review and meta-analysis," CNS Spectrums, vol. 25(2), 2020, pp. 252–263.

Development of a Textbook in the Modern Educational and Scientific Mathematical and Informatics Environment of Higher Education Students

Marianna Kovtoniuk, Olena Soia, Olena Kosovets and Ivanna Leonova
Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Ostrozki Str. 32, 21000 Vinnytsia, Ukraine
kovtonyukmm@vspu.edu.ua, soia.om@vspu.edu.ua, kosovets.op@vspu.edu.ua, ivannaleonova6@gmail.com

Keywords: Educational Space of a Bachelor of Mathematics, Digital Textbook in Higher Mathematics, Modeling of a Textbook, Visualization of Educational Content.

Abstract: The article deals with the problems of modelling and functioning of a textbook (digital textbook) as an integral part of the educational and scientific mathematical and informatics environment of higher education. It is shown that a high-quality digital textbook ensures the formation of not only the mathematical culture of a future specialist, but also the formation of such basic competencies as the ability and willingness to self-learn, apply knowledge, skills and abilities to work with computer mathematics systems to improve the efficiency of education, self-education and professional activity. The features of a modern digital textbook in mathematical disciplines in the preparation of a bachelor of mathematics are discussed: 1) the theoretical aspect of creating textbooks in the educational space of a bachelor of mathematics in the conditions of blended learning in Ukraine; 2) content of the digital textbook in accordance with the current regulatory framework and practical experience of teachers; 3) techniques of visualisation of educational material in mathematical disciplines and demonstration of some of them in the author's textbooks; 4) the results of the student survey show a strong direct relationship between the rankings of the frequency of use of textbooks in educational activities and their impact on the effective learning process, based on the interpretation of the nine R. Gagné learning events. The developed and implemented author's digital manuals meet the main criteria for their effective functioning: strategic focus, completeness of coverage, intensity, orderliness, coherence and mobility.

1 INTRODUCTION

Creating a textbook is a complex process of preparing, publishing, testing, evaluating, and implementing it in the educational process. The main role and wide range of functions of a textbook in education have always attracted constant attention to the problems of its content, quality, and creation. And this is no accident.

Firstly, a textbook, as a learning book, reflects in detail the content of education, the learning information to be learned. It conveys this information not only in the form of text, but also in illustrations, drawings, diagrams, and graphs. Secondly, an equally important function of a textbook is to manage students' cognitive activity. The apparatus for organizing the learning of educational material consists of two parts: auxiliary knowledge included in the main educational material, and tasks, exercises, questions, etc. that should ensure the process of learning and the formation of skills and abilities. That is why scientists interpret a textbook as an information model of learning, as a kind of scenario of the educational process, which is the embodiment of didactically and methodically developed and systematized educational material. From this point of view, the textbook should reflect the goals and content of learning, determine the system of cognitive actions with the material, organizational forms of learning and methods of control.

Recently, we have observed some intensification in the publication of manuals and workshops in mathematical disciplines for the preparation of a bachelor of mathematics in the specialties 111 Mathematics, 014.04 Secondary Education (Mathematics) and 014.09 Secondary Education (Informatics) [1, 2, 3, 4], as well as publications on

the problems of the content of manuals and workshops for these specialties, taking into account the needs of students with special educational needs [5], expanding the use of digital technologies in the creation of digital manuals and workshops, etc. In our opinion, the relevance of this problem was given by the blended learning of students, which has been operating in Ukraine for a long time for various objective reasons.

The purpose of the article is to analyse the development and transformation of a textbook from a classical to a digital one using the principles of foundation, visualisation of educational content and implementation of nine learning events according to R. Gagné of the sequential transition from one state of student understanding of the educational material to other states; to conduct an experimental study among students to assess the presence and significance of the relationship between the rankings of the frequency of use of textbooks.

2 ANALYSIS OF CURRENT RESEARCH

Mexican scientists M. Pineda Becerril et al. [6] note that "a digital book is a publication based on an electronic file that can be stored on various digital media and allows the inclusion of interactive and multimedia elements". The authors have developed a multimedia tutorial on probability distribution, which is integrated into the relevant virtual environments so that students can immerse themselves in images with interactive subtitles and rotate an object in 3D.

The article [7] discusses the transformation of school teaching materials in the context of the digital society in Latin America and Europe.

Spanish researchers N. Rodríguez-Regueira and J. Rodríguez-Rodríguez [8] analyzed digital textbooks for primary education in Spain (30 digital learning materials). The novelty of their study is the development and testing of a guide for analyzing this type of materials, taking into account their main technological, pedagogical and functional characteristics in connection with the evolution of the digital educational market.

Research by contemporary scholars shows that digital textbooks are particularly difficult to integrate into mathematics teaching [9] and that it is worthwhile to study in more detail how the conditions for teaching mathematics are changing due to modern digital textbooks [10] and textbooks that use artificial intelligence [11] and how learning can be developed with their help.

The use of open textbooks in higher education institutions and the financial benefits to students from this, without reducing their learning outcomes, are emphasized in the article [12, 13]. This should encourage teachers to use high-quality open textbooks in the post-pandemic world to expand digital learning opportunities in the twenty-first century.

3 MAIN RESULTS

3.1 Theoretical Aspects of Creating Textbooks in the Educational Space of Bachelor of Mathematics in the Conditions of Blended Learning in Ukraine

The need to introduce digital technologies into the educational process is demonstrated in studies conducted by the National Training Center in the United States. The results of the research are called the "Learning Pyramid" and show that when using video and audio materials, 20% more material is learned, when using demonstrations, 30% more of the material is learned than when listening to a regular lecture. The largest percentage of knowledge gain during the course of study (work in small groups, work in pairs) and the application of the acquired knowledge immediately after studying ("brainstorming", "aquarium") is 90 % [14].

Our research suggests that in blended learning, student manuals should be consistent:

$$\text{Student} \rightarrow \text{Textbook} \rightarrow \text{Teacher}$$
$$(\text{or Student} \rightarrow \text{Educational and Methodological Complex (Digital EMC)} \rightarrow \text{Teacher}).$$

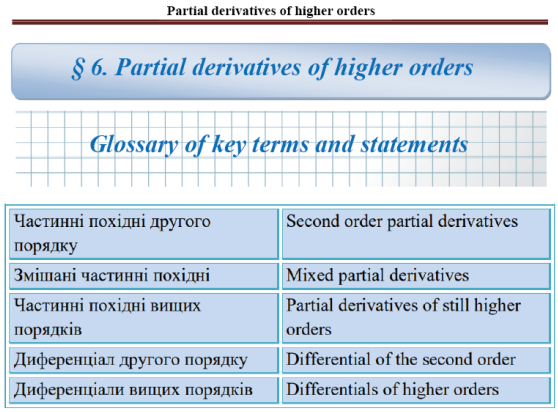
In this case, the role of the teacher is not diminished, but rather enhanced, because it is here that the teacher not only provides ready-made information but also teaches students to search for, analyze, and process new information on their own.

The textbook should be integrated into the teaching technology that the teacher designs and implements. In this case, the logic and structure of the lesson will be an element of the teacher's creativity, and he or she will be able to choose his or her own strategy and teaching methodology, and not just follow the presentation of material proposed by other authors. In a context where the idea of student-centered learning based on teacher innovation is fundamental to education, this approach to textbooks,

in our opinion, is becoming crucial in textbook development.

When creating textbooks, it is necessary to take into account that: textbooks should have a high scientific and methodological level, contain the necessary reference apparatus; textbooks should be written in an accessible form, the educational material should be related to practical tasks, the book should have close interdisciplinary connections [15].

For example, the textbook by M. Kovtoniuk, A. Klimishyn, I. Leonova, and O. Soia contains a glossary of mathematical terms in Ukrainian and their equivalents in English (Fig. 1).


Partial derivatives of higher orders

§ 6. Partial derivatives of higher orders

Glossary of key terms and statements

Частинні похідні другого порядку	Second order partial derivatives
Змішані частинні похідні	Mixed partial derivatives
Частинні похідні вищих порядків	Partial derivatives of still higher orders
Диференціал другого порядку	Differential of the second order
Диференціали вищих порядків	Differentials of higher orders

Figure 1: A fragment of a textbook on higher mathematics[1] [3].

The publication of textbooks and workbooks in Ukraine in the discipline of Mathematical Analysis has been growing over the past 35 years. In particular, students majoring in 111 Mathematics, 014.04 Secondary Education (Mathematics) and 014.09 Secondary Education (Computer Science) use the textbooks by A. Dorogovtsev (1994), I. Liashko, A. Boyarchuk, Y. Gay, A. Kalaida (1979), L. Duzhenkova, A. Kalaida (1979) (Taras Shevchenko National University of Kyiv), L. Duzhenkova, T. Kolesnyk, M. Liashenko, H. Mykhalina, M. Shkil (2003) (Mykhailo Dragomanov Ukrainian State University), N. Shunda, A. Tomusiuk, A. Shunda, A. Tomusiak, M. Kovtoniuk, A. Klimishyna, I. Leonova, O. Soia (2008, 2009, 2011, 2015, 2022, 2023) (Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, 1993), S. Gurgula, V. Moisyshyn, V. Vorobiova (2008) (Ivano-Frankivsk National Technical University of Oil and Gas), B. Kovalchuk, Y. Shipka (2002, 2004, 2006, 2010) (Ivan Franko National University of Lviv), etc.

Designing the content of vocational mathematics education is a necessary key condition for the successful construction of the discipline of the speciality. Particular attention should be paid to the development of the speciality curriculum, work programme, teaching technologies, and quality management. On the basis of these documents, educational and methodological and digital educational and methodological complexes of the discipline are designed and created, which include, in particular, author's educational and methodological developments, manuals included in the teaching technology.

We propose to deepen the theoretical and practical components of the professional education of a bachelor of mathematics on the basis of strengthening the school component of mathematics education with further foundation of knowledge at different levels. The principle of foundation in the process of teaching mathematics is understood as the process of identifying the basic educational elements of school mathematics with their subsequent theoretical generalisation, revealing the integrity, essence, transdisciplinary connections, and aimed at the intellectual development of students. The peculiarity of the foundation principle is the definition of the basis for a spiral scheme for modelling the subject competences of a bachelor of mathematics (the foundation spiral). In each mathematical discipline, it is worth highlighting the basic concepts of school mathematics with further analysis of their foundation. They will form the fundamental core of mathematical disciplines.

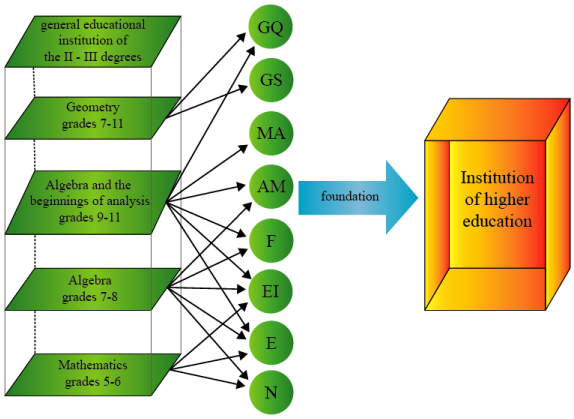


Figure 2: Source objects of funding[2].

Let's analyse this process using several examples. The educational course "Mathematics" in general secondary education is traditionally taught as the following disciplines: algebra, algebra and beginning

[1] https://kovtonyuk.com

[2] author's development

of analysis, geometry. The State Standards of School Mathematics Education define eight content areas of the school mathematics course [16]: 1) Numbers and operations on them [N]; 2) Expressions and their identical transformations [E]; 3) Equations, inequalities and their systems [EI]; 4) Functions, their properties and graphs [F]; 5) Elements of applied mathematics [AM]; 6) Beginnings of mathematical analysis [MA]; 7) Geometric shapes, their elements and properties [GS]; 8) Geometric quantities, their measurement and calculation [GQ].

The definition of content lines allows us to identify the initial objects of funding (Fig. 2).

It is crucial that all fundamental school-level knowledge is incorporated into the list of elements of the academic disciplines of bachelor's degree training in mathematics for the aforementioned specialties and is transferred from the database (formal operations in school mathematics) to the database of subject and professional competencies of the student. For example, for the concept of a derivative, it is necessary to substantiate the transition from the definition of a derivative based on the error of the difference ratio of the increment of a function to the increment of the argument to the definition of the limit transition and the delta-epsilon language and the language of sequences. Filling in these "transitions" between concepts, theorems, methods of proof, and guiding foundations of activity is one of the main tasks of the academic discipline "Mathematical Analysis".

The concepts of the primitive and definite integral, which are studied in the 11th grade of general secondary education institutions, are gradually generalized through the Riemann definite integral for a function of one variable, multiple integrals (double, triple, surface, curvilinear) for functions of many variables, and, finally, in the course of modern mathematical (functional) analysis, the Lebesgue integral on measurable sets is studied.

Having reached a certain level of abstraction of this concept, students begin to work on the methodological analysis of the concepts of the original and definite integral in the course of methods of teaching mathematics (Fig. 3). The highest level of foundation is provided during the modelling of real processes and phenomena with the help of integrals, when the student applies qualitative mathematical research methods, computer mathematics systems, and develops practical recommendations.

At each level of foundation, special attention should be paid to understanding the essence of each basic concept.

Thus, summarising the above, it can be argued that the information component of creating textbooks for bachelors of mathematics should be formed on the basis of the principle of foundation of mathematical objects.

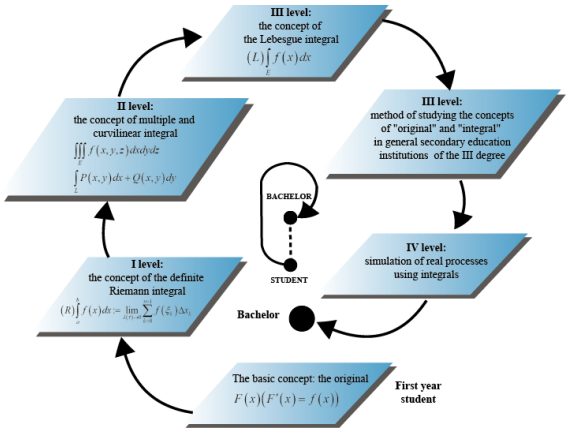


Figure 3: Scheme of foundation of the school concepts of the initial and definite integral[3].

3.2 Visualisation in Mathematics Textbooks

The term "visualisation" comes from the Latin visualis - perceived visually, visual. Visualisation is the process of presenting data in the form of an image for the purpose of maximising the ease of understanding; giving a tangible form to any object, subject, process, etc. However, such an understanding of visualisation involves minimal mental and cognitive activity of students, and visual didactic tools perform only an illustrative function [17].

Visualisation is a visualisation, creating conditions for visual observation [18].

The use of tables, diagrams, and drawings facilitates quick memorisation and comprehension of the material being studied. Taking into account modern technical capabilities, the idea of visualising information in the learning process takes on new features (Fig. 4).

Advantages of visualisation in teaching mathematical disciplines:

- helps students to organise and analyse information correctly;
- charts, diagrams, drawings, memory maps help to assimilate large amounts of information, make it easy to remember and trace the relationship between blocks of educational material;
- develops critical thinking;
- helps students integrate new knowledge;

[3] author's development

- allows you to link the information received into a holistic picture of certain phenomena or objects (Fig. 5).

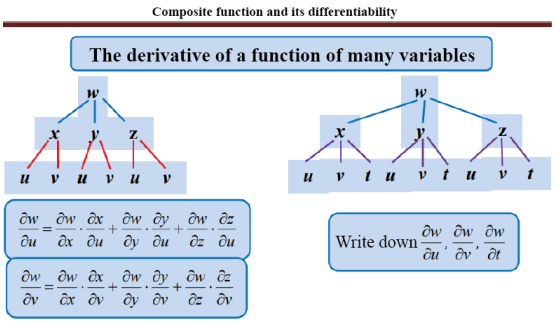


Figure 4: A fragment of a mathematics textbook[4] [3].

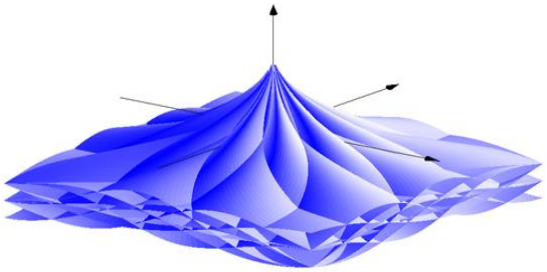


Figure 5: Visualisation of the surface "Snow Avalanche"[5].

Presentation of educational material using modern visualisation techniques is an integral part of a mathematics textbook for students of higher education institutions. With the help of visualisation, the teacher demonstrates complex abstract mathematical concepts, which helps to develop the student's imagination and memory and helps to dive deeply into the learning material.

The experimental study confirmed that the educational material presented in the textbooks sufficiently ensures the implementation of the nine learning events according to R. Gagné, a sequential transition from one state of student understanding of the educational material to other states.

4 DISCUSSION

The study considers the results of a survey of students of the Faculty of Mathematics, Physics and Computer Science of Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Ukraine.

The results of the first stage of the study showed which digital resources students use during their studies (multiple choice): textbooks (78.9 %), lecture notes (92.1 %), teachers' workbooks for laboratory and practical classes (55.3 %), individual materials of teachers for laboratory or practical classes (68.4 %), search engines (78.9 %), other open sources of information (5.2 %).

The research aims to find out whether the learning material presented in the textbooks sufficiently ensures the implementation of the nine learning events according by R. Gagné, who believed that a student should move sequentially from one state to other states of understanding the learning material. The first learning event is drawing attention to the learning material and motivating students. The ninth and final stage is the use of the acquired knowledge, skills and abilities in practice. The teacher helps the student to go through this process. Analysing a textbook according by R. Gagné makes it possible to understand how effectively the educational content is prepared, whether students have achieved the expected learning outcomes, acquired relevant knowledge, skills, mastered different ways of thinking, acquired scientific views and values that determine the ability of students to successfully socialise and conduct further professional activities [19].

The second stage of the study was conducted among students in order to assess the presence and significance of the relationship between the rankings of the frequency of using textbooks in educational activities (variable X) and their impact on the effective learning process according to the interpretation of the nine learning events by R. Gagné (variables Y_j, $j = \overline{1,9}$). The sample consisted of N=76 students aged 18-21 years (57 % girls). Parental consent was not required, as all students were of legal age at the time of the survey. The research was conducted with the permission of the faculty administration in compliance with ethical standards and with maximum respect for the privacy and confidentiality of students. The authors of the article conscientiously adhered to the ethical standards of the American Psychological Association (APA) and the recommendations of the Ethical Code of the Scientist of Ukraine [20].

The questions of the questionnaire are based on the nine events of Gagné's learning. Students ranked their answers on a scale where a score of 10 means absolute agreement – "Yes", and a score of 1 means absolute disagreement – "No". If necessary, intermediate values were also used.

Step 1. Attract attention.

(Y_1) Do study guides contribute to your interest in learning e.g. [6]?

[4] https://kovtonyuk.com

[5] author's development

[6] https://press.vntu.edu.ua/index.php/vntu/catalog/book/812?fbclid=IwAR3sdYLfk60WmS5HFzbRJtBwRisptz8AxUSxuz1YKLVsGSw7AgorYA1aPzo

Step 2. Informing students about the purpose and objectives of learning.

(Y_2) Does the use of study guides help students to understand the purpose and objectives of the learning?

Step 3. Stimulate the repetition of previous learning material.

(Y_3) Do you pay attention to the presence of revision materials, examples or questions in the textbooks to test your previous knowledge?

Step 4. Presentation of learning material.

(X) How often do you use textbooks (manuals) in your learning activities?

Step 5. Tasks, learning support, instructions for students.

(Y_4) In your opinion, are there enough practical examples in the textbooks to explain the theoretical material?

Step 6. Practice.

(Y_5) In your opinion, are there enough practical examples in the textbooks to reinforce the material?

Step 7. Feedback.

(Y_6) Do you get answers to your educational questions with the help of textbooks?

(Y_7) Is it enough for you to interact with the teacher through the tutorials?

Step 8. Assessment of learning outcomes.

(Y_8) In your opinion, how does the use of study guides contribute to self-analysis of your educational activities?

Step 9. Consolidating and applying the acquired knowledge, skills and abilities in practice.

(Y_9) In your opinion, how does the use of study guides form your ability to self-educate?

To identify the existence and significance of relationships, we chose Spearman's rank correlation coefficient r_s, which is used to determine the closeness of relationships between quantitative and qualitative attributes if their values are ranked. The rank correlation coefficient r_s is calculated by the [21]:

$$r_s = 1 - \frac{6\sum_{i=1}^{n}(x_i - y_i)^2}{n \cdot (n^2 - 1)},$$

where n – is the volume of the set of objects; $(x_i - y_i)$ – is the rank difference of the i-th object. The coefficient r_s takes values in the range from – 1 to + 1.

The sampling distribution r_s, which characterises zero correlation between two groups of ranks, is associated with the Student's t-distribution. If the value of r_s is 0 and $n > 10$, the empirical criterion for degrees of freedom $(n - 2)$ is determined by the

$$t_{emp} = \frac{r_s}{\sqrt{(1 - r_s^2)/(n - 2)}}.$$

The null hypothesis of no correlation is rejected for the significance level α at $t_{emp} > t_\alpha$.

The results of the study are presented in Table 1 for $\alpha = 0.01$ ($t_\alpha = t_{0.01}$).

Table 1. Calculations of the existence and significance of the relationship between the rankings of the frequency of using textbooks in educational activities (variable X) and their impact on the effective learning process according to the interpretation of the nine R. Gagné learning steps (variables Y_j, $j = \overline{1,9}$).

Rank correlation coefficient r_s, t-distribution	(X, Y_1)	(X, Y_2)	(X, Y_3)	(X, Y_4)	(X, Y_5)	(X, Y_6)	(X, Y_7)	(X, Y_8)	(X, Y_9)
r_s	0.994	0.997	0.997	0.994	0.997	0.997	0.997	0.998	0.998
t_{emp}	80.512	109.670	111.175	78.076	112.214	105.057	103.782	123.785	121.054
$t_{0,01}$	2.894								

Thus, since all $t_{emp} > t_\alpha$ for the significance level $\alpha = 0.01$, the null hypothesis of no correlation is rejected at the 0.01 level. The numerical value of $r_s > 0.99$ indicates a strong direct relationship between the rankings of the frequency of using textbooks in educational activities (variable X) and their impact on the effective learning process as interpreted by the nine R. Gagné learning events (variables Y_j, $j = \overline{1,9}$).

5 CONCLUSIONS

The article deals with the problems of development and transformation of a textbook (digital textbook) as an integral part of the educational and scientific mathematical and informatics environment of higher education. It is shown that a high-quality digital textbook ensures the formation of not only the mathematical culture of a future specialist, but also the formation of such basic competencies as the ability and willingness to self-learn, apply

knowledge, skills and abilities to work with computer mathematics systems to improve the efficiency of education, self-education and professional activity. The use of a digital textbook integrated into the learning technology designed and implemented by the teacher allows him or her to choose their own creative strategy and methodology for teaching students.

The features of a modern digital textbook in mathematical disciplines in the preparation of a bachelor of mathematics are discussed:

1) the theoretical aspect of creating textbooks in the educational space of a bachelor of mathematics using the principles of foundation, spiral and continuity in the context of blended learning in Ukraine;
2) content of the digital textbook in accordance with the current regulatory framework and practical experience of teachers;
3) techniques of visualisation of educational material in mathematical disciplines and demonstration of some of them in the author's textbooks;
4) the results of the student survey show a strong direct relationship between the rankings of the frequency of use of textbooks in educational activities and their impact on the effective learning process, based on the interpretation of the nine R. Gagné learning events.

The developed and implemented author's digital manuals meet the main criteria for their effective functioning: strategic focus, completeness of coverage, intensity, orderliness, coherence and mobility.

In the future, we aim to investigate the impact of using a digital textbook on the formation of mathematical, informatics and pedagogical culture of a higher education student.

REFERENCES

[1] G. Hartman, S. Fitzpatrick, A. Jordan, and C. Vollet, "APEX Calculus," 2023 [Online]. Available: https://opentext.uleth.ca/PDF/apexnovideo.pdf

[2] J. Hass, M. Weir, and G. Thomas, "University Calculus. Early Transcendentals," Second Edition, Boston, 2012.

[3] M. M. Kovtoniuk, A. Ya. Klimishyna, I. M. Leonova, and O. M. Soia, “Workshop on differential calculus of a function of many variables, “Vinnytsia: VNTU, 2023 [Online]. Available: https://press.vntu.edu.ua/index.php/vntu/catalog/book/812.

[4] R. V. Khats, "Asymptotic estimators and their application: lecture texts, practical and individual tasks," Drohobych, 2023.

[5] O. P. Kosovets, "Features of creating an electronic textbook on computer science for students with disabilities,". Computer-oriented learning systems, Issue 19, 2016 [Online]. Available: http://www.ii.npu.edu.ua/zbirnyk-kosn

[6] M. Pineda Becerril, O. García, A. Aguilar, and F. León, "Development of a Multimedia Book of the Probability Distributions Subject as Support in the Fesc Statistics Subjects," ICERI 2018 Proceedings, pp. 7550–7555, 2018 [Online]. Available: https://doi.org/10.21125/iceri.2018.0351

[7] "Introduction: From Textbooks to Digital Teaching Materials," Latin American Journal of Educational Technology – RELATEC, vol. 16 (2), pp. 9–12, 2017 [Online]. Available: https://doi.org/10.17398/1695-288X.16.2.9

[8] N. Rodríguez-Regueira, and J. Rodríguez-Rodríguez, "Analysis of Digital Textbooks," Educational Media International, vol. 59(2), pp 172–187, 2022 [Online]. Available: https://doi.org/10.1080/09523987.2022.2101207

[9] B. Pepin, G. Gueudet, and L. Trouche, "Refining Teacher Design Capacity: Mathematics Teachers’ Interactions with Digital Curriculum Resources," ZDM Mathematics Education, vol. 49 (5), pp. 799–812, 2017 [Online]. Available: http://rdcu.be/tmXb

[10] M. Utterberg Modén, "Teaching with Digital Mathematics Textbooks. Activity theoretical studies of data-driven technology in classroom practices": Doctoral Dissertation / University of Gothenburg. Gothenburg. Sweden, 2021.

[11] M. Utterberg, M. Tallvid, J. Lundin, and B. Lindström, "Artificial Intelligence and Digital Mathematics Textbooks," 12th International Conference of Education, Research and Innovation (ICERI2019), pp. 2320–2320, 2019 [Online]. Available: https://doi.org/10.21125/iceri.2019.0627

[12] J. Hilton, "Open Educational Resources and College Textbook Choices: A Review of Research on Efficacy and Perceptions," Educational Technology Research and Development, vol. 64 (4), pp. 573–590, 2016 [Online]. Available: https://doi.org/10.1007/s11423-016-9434-9

[13] H. Ma, "Empowering Digital Learning with Open Textbooks," Education Tech Research Dev, vol. 69, pp. 393–396, 2021 [Online]. Available: https://doi.org/10.1007/s11423-020-09916-9

[14] "The pyramid of learning" [Online]. Available: https://fsp.kpi.ua/ua/piramidanavchannya-edgara-dejla/

[15] "Methodological recommendations on the structure, content and scope of textbooks and manuals for higher education institutions." Scientific and Methodological Center for Higher Education of the Ministry of Education and Science of Ukraine. Protocol N 6 of 29.07.05 [Online]. Available: https://zakon.rada.gov.ua/rada/show/v0006290-05#Text

[16] "State standard of basic and complete secondary education" [Online]. Available: http://rvo-petr.ucoz.ua/pdf-doc/drzh_standard.pdf

[17] "Visualization of information" [Online]. Available: https://www.slideshare.net/obobko/ss-15090625

[18] "Visualization of educational information" [Online]. Available: http://phys.ippo.kubg.edu.ua/?page_id=662

[19] R. Gagné, L. Briggs, and W. Wager, "Principles of Instructional Design" (4th Ed.). Fort Worth, TX: HBJ College Publishers, 1992

[20] "The Code of Ethics for Scientists of Ukraine," Homepage [Online]. Available: https://zakon.rada.gov.ua/rada/show/v0002550-09#Text, last accessed 2024/04/11.

[21] V. Rudenko, "Mathematical statistics," Kyiv : Centre for Educational Literature, 2021, 304 p.

Optimizing UAE Food Supply Chain Management: Leveraging Fuzzy AHP for Strategic Selection of Optimal Blockchain Platforms

Hour Alburaimi and In-Ju Kim
Department of Industrial Engineering and Engineering Management, College of Engineering, University of Sharjah, UAE
u22104343@sharjah.ac.ae, ikim@sharjah.ac.ae

Keywords: Blockchain Technology, Blockchain Platform, Food Supply Chain Management, Fuzzy AHP.

Abstract: The United Arab Emirates faces significant challenges in food security due to heavy reliance on imported food and vulnerabilities in global supply chains. Blockchain offers decentralized, immutable solutions to improve food supply chain transparency, traceability, and quality assurance. This research focuses on developing a decision-making framework using the Fuzzy Analytical Hierarchy Process to select the optimal blockchain platform tailored to the UAE context. The study reviews blockchain's fundamental components and applications, highlighting its transformative potential in supply chain management. Methodologically, Fuzzy AHP evaluates criteria like technical feasibility, security, regulatory compliance, and cost-effectiveness, prioritizing factors crucial for UAE's food supply chain. Interviews were conducted with 15 experts and stakeholders in the UAE's food supply chain. Results emphasize technical factors such as platform usability, interoperability, and consensus mechanisms as pivotal in platform selection. This study underscores blockchain's potential to enhance transparency, reduce inefficiencies, and build trust in UAE's food supply chain, offering a structured approach for decision-makers to navigate adoption challenges effectively. The theoretical contribution lies in providing a structured approach for decision-makers to navigate the complexities of blockchain adoption in food supply chain management, addressing both technological feasibility and regulatory challenges.

1 INTRODUCTION

The United Arab Emirates (UAE) has a unique challenge in food security, a combination of several factors predisposing the UAE to future food insecurity risks. According to Ammar [1], the UAE imports 80-90% of all the food consumed in the country. Depending on food sourced from the global market poses distinct challenges for the UAE due to the risks of price hikes and disruptions in the global food supply chain. Moreover, the dangers of climate change have increased the risk of harsh weather in the UAE, which diminishes the prospects of producing food in the future; climate change has led to changes in how food is produced and transported to consumers, while advocacy for climate-neutral food production has led to changes in conventional food systems [2].

The blockchain has emerged as an attractive technology for managing food supply chains by encompassing a set of solutions that embrace decentralized systems for immutable data storage and access [3]. According to Bashir [4], blockchain technology is an innovative database system facilitating transparent data sharing among business network members. A blockchain database organizes data into interconnected blocks, forming a sequential chain. Its chronological consistency is because the chain cannot be changed or removed without consensus from the network; by utilizing blockchain technology to create an unchangeable or immutable ledger, orders, payments, accounts, and other transactions can all be tracked. Unauthorized transaction inputs are prevented by built-in system features, ensuring consistency in the shared view of these transactions.

In the food supply chain, blockchain provides 100% traceability of food-related data and multi-party transactions, enables backtracking food provenance in seconds rather than days, makes verifying food safety and quality compliance more manageable, and improves data protection. The basis of the blockchain technology utilized in the food supply chain is a distributed ledger that holds information on every transaction and event that occurs in the chain. Data blocks that are encrypted,

timestamped, and connected chronologically make up the ledger. A batch of transactions verified using the pre-defined consensus protocol is stored in each block.

The overarching purpose of this paper is to develop a decision-making framework for blockchain technology evaluation and selection of an optimal blockchain platform tailored to the unique needs of the UAE food supply chain by conducting a comprehensive analysis of the unique requirements, regulatory framework, and technological considerations specific to the UAE.

2 LITERATURE REVIEW

This chapter examines various aspects of blockchain technology and its potential applications, particularly in the food supply chain industry.

2.1 Components of the Blockchain Technology

In terms of technical structure, a blockchain contains a distributed ledger with a series of data blocks linked together by cryptographic algorithms. Each block contains information about all transactions for a particular batch. A block comprises a header and body, as shown in Figure 1 below. The header contains information that connects to other blocks. Moreover, information relating to the verification of transactions, the block's timestamp, and the previous block's hash value are contained in the block header.

On the other hand, the block body carries all transaction information in the blockchain [5]. A blockchain operation is initiated when a node records new information and broadcasts it to the entire network. Consequently, the nodes that receive this information conduct verification processes and store them in a block. If all blocks reach a consensus on the transaction's validity, a new block is added to the blockchain, and all nodes are updated with the information [6]. Transaction information is added to the blockchain when all blocks verify the information is correct. In addition to the consensus mechanism, blockchain technology is enabled by cryptography, which is comprised of a series of hash functions and public keys that ensure the security, authenticity, integrity, and immutability of the data contained in the distributed ledger. Moreover, blockchains are facilitated by smart contracts, which are digital promises that are automatically execute when certain conditions are met, thereby enhancing the efficiency and reliability of blockchain operations [5].

2.2 Applications of Blockchain Technology

Blockchain technology has been applied in various fields due to its unique decentralization, immutability, and security advantages. One of the earliest and most successful applications of blockchain technology was in developing and operating cryptocurrencies, which are forms of digital or virtual currencies that use cryptography to achieve security functions. These currencies are operated decentralized without requiring a central bank's mediating role [6]. The success of blockchain technology has motivated its application in other fields.

In healthcare, blockchain technologies are used to preserve, exchange, and manage health records for efficient decision-making [7]. In the education sector, blockchain technology offers immense benefits and potential applications in performing tasks such as resource sharing, administration and management, facilitating online l earning and testing, college

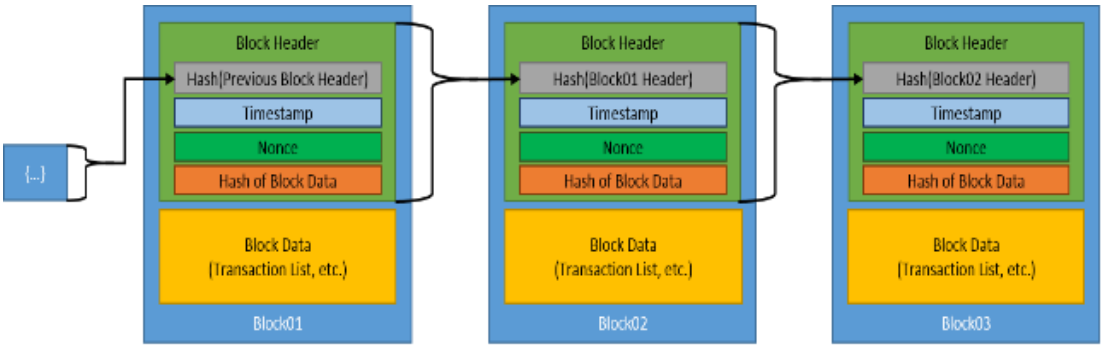


Figure 1: Structure of the blockchain network [5].

crowdfunding, verification and authentication of education records, and data storage and management [8]. In agriculture and food production, blockchain technology may be applied to monitor the origin of food products to build consumer confidence [9]. The technology also offers opportunities for managing food supply chains, management transactions among stakeholders, the transmission of agricultural data, management of agricultural insurance, land registration, and smart farming, among others [9]. Blockchain technology has also been found to be significant in the military and modern warfare. According to Jadav [10], blockchain technology may be adopted in modern warfare to acquire data in critical missions, reconnaissance, and intelligent management of battlefield operations. The immutable nature of blockchain technology allows the military to share susceptible data securely with minimal risk of disclosure to intruders and enemies [10]. The United Arab Emirates has also demonstrated the potential application of blockchain technology in providing government services [11]. Thus, blockchain technology offers immense opportunities for application in diverse fields.

In the UAE, Eletter [12] have reported the utilization of blockchain technology in supply chain management by the retailer Carrefour. Carrefour UAE employed the IBM Food Trust to manage its food traceability by collaborating with food industry players along the value chain. The initiative not only helped to boost trust in the food supply chain but also led to improved efficiency, reduced supply chain risks, and low implementation costs. Eletter [12] demonstrated that the implementation of a full blockchain platform for Carrefour would help the UAE retailer to achieve greater collaboration, enhanced control of food quality, improved transparency, and efficient flow of goods and services. The blockchain platform could also help in real-time platform for negotiation and agreements execution, showcasing the practical benefits of blockchain technology in supply chain management.

2.3 Types of Blockchain Platforms

Several types of blockchain platforms exist, which include public, private, and consortium blockchains (Meng et al., 2021). Of these, public blockchains are are open, decentralized networks with participants that do not require a central entity authority to set it up. They differ from private blockchains, which are centralized networks in terms of authority access, requiring operational authorization, whereby prequalified parties are the only ones authorized to operate the networks [13].

In turn, the participants on a public blockchain can be anonymous and invite other members into the network, while those in a private one must be clearly stated, and only internal participants can manage the network. In turn, the advantage of public blockchains over private ones is the increased visibility of transactions because of the accessible data access to all network users. However, private blockchains have the advantage of increased security because the financially essential data is only sharable and visible within the internal system. In addition, private blockchains are faster than public ones because of the single controlled consensus, increasing transaction rates [13]. The consortium blockchain combines a public and private blockchain. Individuals or groups manage the system's accessibility in a consortium-comprising user and scheduling nodes. This arrangement leads to a fast transactional speed and a high transmission operating efficiency. Figure 2 provides an overview of different types of blockchain platforms.

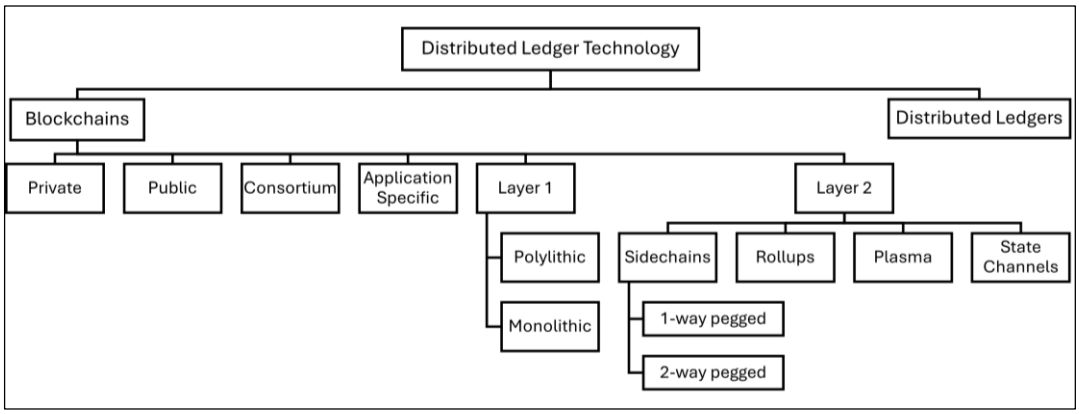


Figure 2: Blockchain platforms types [4].

2.4 Blockchain Technology in the Food Supply Chain

Blockchain technology has been adopted and implemented by several industry players in the food supply chain. In Australia, for instance, a commodity management platform service provider AgriDigital has implemented a blockchain service that verifies and assures agricultural products for consumer benefit [14]. In China, Techrock provides a platform that allows consumers of infant formula to authenticate the product quality and traceability, while the Fiji technology company TraSeable Solutions has developed a blockchain platform for tracing tuna processing (Rodgerson & Parry, 2020). In the United States, Walmart has partnered with IBM to develop a platform incorporating suppliers and distributors to trace the origin of various food products to ensure that they meet quality standards [15]. The fast-moving consumer goods provider Unilever has also implemented a blockchain technology solution that allows the company to manage its tea business. Through the blockchain framework, Unilever monitors all transactions along the supply chain to maintain quality standards [15]. These case studies demonstrate that the UAE could benefit from implementing similar technological solutions.

3 METHODOLOGY

This section outlines the methodology employed for the research, detailing each stage of applying the fuzzy analytical hierarchy process.

3.1 Fuzzy Analytical Hierarchy Process

Multi-criteria decision-making frameworks are useful in decision-making in environments that are characterized by complexities and uncertainties. These frameworks allow the analysis of critical factors based on multiple dimensions. In the present study, a fuzzy analytic hierarchy process (AHP) was used to analyze the evaluation and selection criteria for blockchain platforms in the UAE's food supply chain industry. The process adopted in this study involved the following key steps:

3.1.1 Identification of the Selection Criteria

The first step in the fuzzy AHP process will be identifying the relevant criteria for evaluating blockchain platforms. This study's criteria identified from the literature include technological factors, platform security, organizational factors, regulatory issues, and developer support. Data was collected from the panel of experts to determine the factors that they deemed important in the case study.

3.1.2 Criteria Weighting

The second step involved assigning weights to the criteria identified in the previous step. The weights were assigned based on the relative importance that the experts assigned to each criterion. The experts were assigned to make pairwise evaluations of the factors and their impacts on each other. Consequently, the criteria were ranked hierarchically based on their assigned weights.

3.1.3 Calculating the Fuzzy Weights

The third step involved computing the relative significance of the weights of the criterion established in the previous steps. During this phase, experts were asked to assign scores to the criteria using linguistic terms, as shown in the Table 1.

Table 1: Comparative scale used to assess the importance of criteria and their corresponding fuzzy numbers.

Linguistic terms	Fuzzy scale
Equally important (EI)	[1.00, 1.00, 1.00]
Weakly important (WI)	[2.00, 3.00, 4.00]
Fairly important (FI)	[4.00, 5.00, 6.00]
Strongly important (SI)	[6.00, 7.00, 8.00]
Absolutely important (AI)	[9.00, 9.00, 9.00]

3.1.4 Normalizing the Scores

Next, the scores of criterion performance were normalized through mathematical transformation. This involved replacing linguistic scales with corresponding numerical values, ensuring data consistency, and establishing a common scale for subsequent analysis. The final list of categories and sub-categories used in this evaluation is presented in Table 2

3.1.5 Aggregating the Normalized Scores

In the fifth step, the normalized scores were normalized by multiplying the scores with the corresponding weights. This process culminated in deriving aggregate scores that were used to compare

the criterion from the most to least important. For a criterion j whose fuzzy performance score is $\tilde{a}_{ij}$ and its weight is $\tilde{\omega}_i$, the fuzzy weight $\tilde{u}_i$ is obtained using the following equation:

$$\tilde{u}_i = \sum_{j-1}^{n} \tilde{\omega}_i \tilde{a}_{ij} .$$

Table 2: Final list of categories and sub-categories of blockchain evaluation criteria.

Category	Sub-category	Sub-category code
Technical factors (C1)	Interoperability	C11
	Scalability	C12
	Usability	C13
	Consensus mechanism	C14
	Network performance	C15
Cost (C2)	Initial cost of equipment	C21
	Transaction fees	C22
	Maintenance costs	C23
	Personnel costs	C24
Security factors (C3)	Data security concerns	C31
	User Privacy	C32
	Cybersecurity	C33
Legal and regulatory issues (C4)	Local regulatory framework	C41
	International regulatory framework	C42
	Jurisdictional issues	C43
	Smart contracts	C44
	Antitrust law	C45

3.1.6 Validating the Results

The final step in the fuzzy AHP involves the validation of the results. In this phase, the results of the analysis are validated through comparison with expert opinions. The results were also inspected to ensure their feasibility and consistency with those established in literature.

4 RESULT AND DISCUSION

This chapter discusses the analysis's findings, including the identification of key selection criteria. Interviews were conducted with 15 experts in the UAE's food supply chain.

4.1 Fuzzy AHP Analysis

An initial analysis of the four categories was conducted to determine the hierarchy of the evaluation criteria based on the experts' opinions. As shown in Table 3, technical factors received the most significant weight, with an average score of 0.438. The second-ranked criteria were legal and regulatory issues, with a mean score of 0.205. The third most important selection criterion was the security factors, with an average score of 0.181. Lastly, cost factors were ranked fourth with a score of 0.177. The experts stated that technical factors were the most important in selecting blockchain platforms. A more detailed breakdown of the factors is presented in Tables 4 - Table 7.

Table 3: Main criteria fuzzy evaluation and decision matrix.

	C1	C2	C3	C4	W_{eight}	R_{ank}
C1	(1,1,1)	(3,4,5)	(1,2,3)	(4,5,6)	0.438	1
C2	(0.2, 0.25, 0.33)	(1,1,1)	(0.25, 0.33, 0.5)	(0.33, 0.5,1)	0.177	4
C3	(2,3,4)	(0.25, 0.33, 0.5)	(1,1,1)	(0.33, 0.5,1)	0.181	3
C4	(1,2,3)	(0.33, 0.5,1)	(1,2,3)	(1,1,1)	0.205	2

Table 4: Pairwise comparison and decision matrix for the technological factors sub-criterion.

	C11	C12	C13	C14	C15	W	R
C11	(1,1,1)	(2,3,4)	(0.25, 0.33, 0.5)	(0.33, 0.5, 1)	(1,2,3)	0.268 904	2
C12	(0.25, 0.33, 0.5)	(1,1,1)	(0.25, 0.5,1)	(1,2,3)	(0.2, 0.25, 0.33)	0.162 166	5
C13	(2,3,4)	(0.25, 0.33, 0.5)	(1,1,1)	(0.5, 1,2)	(2,3,4)	0.328 961	1
C14	(1,2,3)	(0.33, 0.5,1)	(1,2,3)	(1,1,1)	(0.33, 0.5,1)	0.239 969	4
C15	(0.33, 0.5,1)	(0.2, 0.25, 0.33)	(3,4,5)	(0.33, 0.5, 1)	(1,1,1)	0.25	3

Table 5: Pairwise comparison and decision matrix for the cost factors sub-criterion.

	C21	C22	C23	C24	W	R
C21	(1,1,1)	(0.33, 0.5,1)	(0.2, 0.25, 0.33)	(0.33, 0.5,1)	0.164967	3
C22	(1,2,3)	(1,1,1)	(0.33, 0.5,1)	(0.2, 0.25, 0.33)	0.257428	2
C23	(0.12, 0.2, 0.25)	(0.2, 0.25, 0.33)	(1,1,1)	(0.33, 0.5,1)	0.137916	4
C24	(0.25, 0.5,1)	(0.33, 0.5,1)	(1,2,3)	(1,1,1)	0.43969	1

Table 6: Pairwise comparison and decision matrix for the security factors sub-criterion.

	C31	C32	C33	W	R
C31	(1,1,1)	(0.25,0.33,0.5)	(0.2,0.25,0.33)	0.109558	3
C32	(2,3,4)	(1,1,1)	(1,2,3)	0.405771	2
C33	(4,5,6)	(0.5,1,2)	(1,1,1)	0.484671	1

Table 7: Pairwise comparison and decision matrix for the regulatory issues sub-criterion.

	C41	C42	C43	C44	C45	W	R
C41	(1,1,1)	(2,3,4)	(0.25, 0.33, 0.5)	(0.33, 0.5,1)	(0.33, 0.5,1)	0.241489	3
C42	(2,3,4)	(1,1,1)	(0.2, 0.25, 0.33)	(1,2,3)	(0.2, 0.25, 0.33)	0.28217	2
C43	(1,2,3)	(0.25, 0.33, 0.5)	(1,1,1)	(0.5,1,2)	(1,2,3)	0.282458	1
C44	(0.2, 0.25, 0.33)	(0.33, 0.5,1)	(1,2,3)	(1,1,1)	(0.33, 0.5,1)	0.193883	4
C45	(0.33, 0.5,1)	(0.2, 0.25, 0.33)	(1,2,3)	(0.25, 0.33, 0.5)	(1,1,1)	0.183064	5

4.2 Ranking of Evaluation Criteria

The final weights criteria were normalized using the scores W = (0.438, 0.177, 0.181, 0.205). As a result, a final weight ranking was obtained for all the 17 sub-criteria. Technical/technological factors emerged as the most critical evaluation criteria for adopting blockchain technology in the UAE's food supply chain industry, as it had the first four top-ranked sub-criteria. Platform usability emerged as the top-ranked factor for blockchain evaluation with a weighted score of 0.125, as shown in Table 8 below. The second and third most relevant factors were interoperability and network performance, and the fourth most important sub-criteria is consensus mechanism. Overall, technical factors occupied the first four positions in the evaluation criteria. The fifth and sixth-ranked positions were the security factors of cybersecurity and user privacy respectively. Furthermore, to round up the top ten positions were personal costs, transaction, scalability, and jurisdictional issues. The Table 8 below shows the ranking of all the 17 criteria used to evaluate blockchain platforms in the UAE's food supply chain.

Table 8: Final weight ranking of blockchain selection criteria.

Category	Weight	Sub-criteria	Weight	Finalized weight	Global ranking
Technical factors (C1)	0.438	C11	0.268904	0.102184	2
		C12	0.162166	0.061623	9
		C13	0.328961	0.125005	1
		C14	0.239969	0.091188	4
		C15	0.25	0.095000	3
Cost (C2)	0.177	C21	0.164967	0.044902	13
		C22	0.257428	0.070068	8
		C23	0.137916	0.037539	16
		C24	0.43969	0.070491	7
Security factors (C3)	0.181	C31	0.109558	0.019940	17
		C32	0.405771	0.073850	6
		C33	0.484671	0.088210	5
Legal and regulatory issues (C4)	0.205	C41	0.241489	0.051920	12
		C42	0.28217	0.060667	11
		C43	0.282458	0.060728	10
		C44	0.193883	0.041685	14
		C45	0.183064	0.039359	15

4.3 Discussion

This study has formulated a strategic decision-making framework for evaluating and selecting blockchain platforms in the UAE's food supply chain. The findings show that technical and technological factors are the most essential evaluation criteria. System usability, interoperability, network performance, consensus mechanism, and cybersecurity were identified as the most critical evaluation factors for blockchain adoption in the UAE's food supply chain. These findings are

consistent with earlier studies that show that technical issues are key factors in blockchain adoption in the food supply chain. For example, a study by Fernandez-Vazquez [16] revealed that technical factors such as decentralization, security, and system resiliency were the priority factors for blockchain adoption in the supply chain management sector. In India [17], have reported technical factors such as traceability, availability of real-time information, and presence of immutable databases as the most important factors in blockchain adoption. On their part, Okorie [18] have reported technical issues such as consensus mechanisms, processing power, scalability, and data infrastructure as critical barriers to blockchain use in the food supply chain. Therefore, developing an effective framework for blockchain technology in the UAE must start with establishing the relevant technological infrastructure.

Cost factors also emerged as key selection factors for blockchain evaluation in the UAE's food supply chain industry. The major cost drivers were identified as personnel costs and transaction fees. Implementing blockchain technology involves significant investment in technology adoption, knowledge management, and technology deployment [19]. Therefore, availability of funds is a critical determinant of the attractiveness of the technology for individual firms. Investigations by Okorie [18] have revealed that the cost of building a blockchain platform is an important priority factor because of the need for firms to acquire new infrastructure, incur development costs, and maintain the systems' operations. Therefore, funds availability could be an important factor in the platform selection in the UAE.

Legal and regulatory factors were also identified as essential criteria for blockchain selection. However, these factors were ranked behind technological, security, and cost factors. This observation could be because the experts interviewed in this study viewed the platform from a technical and economic perspective. Nevertheless, legal and regulatory issues are important in blockchain technology. Public blockchains, for instance, present confidentiality challenges that may have far-reaching legal consequences [20]. Further, the blockchain platforms operate in a nascent and still-evolving legal and regulatory framework. Therefore, their operational environment is unpredictable as more laws and regulations are enacted in different jurisdictions. According to Katopodi [21], blockchains also present antitrust law challenges, which are yet to be addressed by existing legal frameworks.

5 CONCLUSIONS

In conclusion, this paper explores blockchain technology's potential to transform UAE's food supply chain management, addressing food security challenges exacerbated by global dependencies and climate change. By leveraging the Fuzzy Analytical Hierarchy Process (AHP), the study developed a robust decision-making framework that prioritizes technical feasibility, security, cost-effectiveness, and regulatory compliance in selecting optimal blockchain platforms. Key findings emphasize the critical role of technical factors like platform usability, interoperability, and consensus mechanisms in enhancing supply chain transparency and efficiency. Cost considerations are also pivotal, underscoring the importance of financial viability in blockchain adoption. While legal and regulatory aspects are considered, their lower prioritization reflects the evolving nature of blockchain governance frameworks in food supply contexts. Practical implications include blockchain's potential to mitigate supply chain risks, enhance traceability, and strengthen consumer confidence in food safety and quality. Theoretical contributions extend to advancing blockchain's applicability in complex supply chain environments and offering a structured approach for decision-makers navigating technological adoption. Future research could explore implementation challenges, scalability issues, and validate the framework's applicability across diverse industries and global contexts, promoting resilience and sustainability in food systems worldwide.

REFERENCES

[1] K. A. Ammar, A. M. Kheir, B. M. Ali, B. Sundarakani, and I. Manikas, “Developing an analytical framework for estimating food security indicators in the United Arab Emirates: a review,” Environ. Dev. Sustain., vol. 26, pp. 5689–5708, 2024. [Online]. Available: https://doi.org/10.1007/s10668-023-03032-3.

[2] C. M. Galanakis, “The ‘vertigo’ of the food sector within the triangle of climate change, the post-pandemic world, and the Russian-Ukrainian war,” Foods, vol. 12, no. 4, p. 721, 2023. [Online]. Available: https://doi.org/10.3390/foods12040721.

[3] R. Singh, S. Khan, J. Dsilva, and P. Centobelli, “Blockchain integrated IoT for Food Supply Chain: A grey-based Delphi-DEMATEL approach,” Appl. Sci., vol. 13, no. 2, p. 1079, 2023. [Online]. Available: https://doi.org/10.3390/app13021079.

[4] I. Bashir, Mastering Blockchain: Inner workings of blockchain, from cryptography and decentralized identities, to DeFi, NFTs and Web3. Packt Publishing, 2023.

[5] M. K. Lim, Y. Li, C. Wang, and M. L. Tseng, "A literature review of blockchain technology applications in supply chains: A comprehensive analysis of themes, methodologies and industries," Comput. Ind. Eng., vol. 154, p. 107133, 2021. [Online]. Available: https://doi.org/10.1016/j.cie.2021.107133.

[6] M. T. Al Ahmed, F. Hashim, S. J. Hashim, and A. Abdullah, "Hierarchical blockchain structure for node authentication in IoT networks," Egypt. Inform. J., vol. 23, no. 2, pp. 345-361, 2022. [Online]. Available: https://doi.org/10.1016/j.eij.2022.02.005.

[7] A. Haleem, M. Javaid, R. P. Singh, R. Suman, and S. Rab, "Blockchain technology applications in healthcare: An overview," Int. J. Intell. Netw., vol. 2, pp. 130-139, 2021. [Online]. Available: https://doi.org/10.1016/j.ijin.2021.09.005.

[8] X. Chen, D. Zou, G. Cheng, H. Xie, and M. Jong, "Blockchain in smart education: Contributors, collaborations, applications and research topics," Educ. Inf. Technol., vol. 28, no. 4, pp. 4597-4627, 2023. [Online]. Available: https://doi.org/10.1007/s10639-022-11399-5.

[9] G. S. Sajja et al., "Towards applicability of blockchain in agriculture sector," Mater. Today: Proc., vol. 80, pp. 3705-3708, 2023. [Online]. Available: https://doi.org/10.1016/j.matpr.2021.07.366.

[10] N. K. Jadav et al., "Blockchain-based secure and intelligent data dissemination framework for UAVs in battlefield applications," IEEE Commun. Stand. Mag., vol. 7, no. 3, pp. 16-23, 2023. [Online]. Available: https://doi.org/10.1109/MCOMSTD.0005.2200052.

[11] A. Alketbi, Q. Nasir, and M. Abu Talib, "Novel blockchain reference model for government services: Dubai government case study," Int. J. Syst. Assur. Eng. Manage., vol. 11, no. 6, pp. 1170-1191, 2020. [Online]. Available: https://doi.org/10.1007/s13198-020-00971-2.

[12] S. F. Eletter et al., "Leveraging Blockchain-Based Smart Contracts in the Management of Supply Chain: Evidence from Carrefour UAE," in Proc. 2022 Int. Arab Conf. Inf. Technol. (ACIT), IEEE, 2022, pp. 1-5. [Online]. Available: https://doi.org/10.1109/ACIT57182.2022.9994083.

[13] R. Astuti and L. Hidayati, "How might blockchain technology be used in the food supply chain? A systematic literature review," Cogent Bus. Manage., Aug. 8, 2023. [Online]. Available: https://doi.org/10.1080/23311975.2023.224673.

[14] M. Rogerson and G. C. Parry, "Blockchain: Case studies in food supply chain visibility," Supply Chain Manage. Int. J., vol. 25, no. 5, pp. 601-614, 2020. http://dx.doi.org/10.1108/SCM-08-2019-0300.

[15] M. Sharma and P. Kumar, "Adoption of blockchain technology: A case study of Walmart," in Blockchain Technology and Applications for Digital Marketing, IGI Global, 2021, pp. 210-225.

[16] S. Fernandez-Vazquez, R. Rosillo, D. De la Fuente, and J. Puente, "Blockchain in sustainable supply chain management: An application of the analytical hierarchical process (AHP) methodology," Bus. Process Manage. J., vol. 28, no. 5/6, pp. 1277-1300, 2022. [Online]. Available: https://doi.org/10.1108/BPMJ-11-2021-0750.

[17] V. S. Yadav, A. R. Singh, R. D. Raut, and N. Cheikhrouhou, "Blockchain drivers to achieve sustainable food security in the Indian context," Ann. Oper. Res., pp. 1-39, 2021. [Online]. Available: https://doi.org/10.1007/s10479-021-04308-5.

[18] O. Okorie et al., "Removing barriers to Blockchain use in circular food supply chains: Practitioner views on achieving operational effectiveness," Cleaner Logist. Supply Chain, vol. 5, p. 100087, 2022. [Online]. Available: https://doi.org/10.1016/j.clscn.2022.100087.

[19] E. Katopodi, "Blockchain market: Regulatory concerns arising from the 'Diem' example in the field of free competition," EU Comp. Law Issues Challenges Ser., 2021. [Online]. Available: https://doi.org/10.25234/eclic/18821.

[20] A. Konkin and S. Zapechnikov, "Privacy methods and zero-knowledge proof for corporate blockchain," Procedia Comput. Sci., pp. 471-478, 2021. [Online]. Available: https://doi.org/10.1016/j.procs.2021.06.055.

Novel Computer Software for Interpolation and Approximation of Ravine and Stiff Digital Dependencies Using Root-Polynomoal and Root-Fractional-Rational Functions

Igor Melnyk[1], Mykhailo Skrypka[1], Alina Pochynok[2] and Olga Demyanchenko[3]

[1]*Department of Electronic Devices and Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteiska Avenue 37, 03056 Kyiv, Ukraine*

[2]*Research Institute of Electronics and Microsystems Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteiska Avenue 37, 03056 Kyiv, Ukraine*

[3]*Institute of Applied Mathematics and Fundamental Sciences of the National University "Lviv Polytechnic", Mytropolyta Andreia Str. 5, Building 4, 79016 Lviv, Ukraine*

imelnik@phbme.kpi.ua, scientetik@gmail.com, alina_pochynok@yahoo.com, olgademyanchenko@gmail.com

Keywords: Interpolation, Approximation, Ravine Dependence, Stiff Dependence, Root-Polynomial Function, Root-Fractional-Rational Function.

Abstract: The article considers the possibilities of solving interpolation and approximation problems using special types of functions, such as root polynomials and root fractional rational, and provides relevant examples. It is prove, that the use of root polynomial functions is especially effective for interpolation and approximation of numerical dependencies with a ravine data set, and the use of root fractional rational functions gives the best results for various data sets with a more rigid functional dependence. To solve the approximation problem, a new approximation by reference points is proposed and tested. With a small number of points in the data set for the approximation problem, equal to twenty or less, the convergence of the proposed method is usually guaranteed. In general, the proposed algorithms are very universal and can be easily adapted to any complex problems. All the proposed methods are implemented and tested in the newly developed computer software created in the Python programming language.

1 INTRODUCTION

In general, the problem of interpolation and approximation of digital data is well known and described in various manual books on numerical methods and computational technologies [1 – 7]. But, although classical approaches in the theory of interpolation and approximation are still effectively used today, their application is often difficult or even impossible to solve complex modeling problems with minimization of data exchange. Furthermore, elaboration of novel approaches for solving interpolation and approximation tasks is very important today for organizing effective data exchange in modern network technologies, including cloud and fog computing [8].

Polynomial functions, especially when increasing the order of interpolation and approximation, always give unwanted outliers that are not typical for the behaviour of the function, and interpolation using splines requires a large number of reference points, which increases the set of necessary numerical data. There is often no compromise solution to achieve high interpolation accuracy with a small number of points, especially when using high-order functions [1 – 7]. For this reason, since 2019, a new approach to interpolation and approximation of ravine, vertex, and other types of rigid functions with local and extrema, based on the use of root-polynomial and root-fractional-rational functions, has been proposed and implemented. Basic principles of this approach have been formulated in the papers [9 – 11].

In the papers [9, 10], it has been shown and proven that for symmetric ravine data sets, the level of relative error of interpolation using root-polynomial functions from second to fifth order is in the range of a fraction of a percent.

Further research on root-polynomial functions has been provided in the papers [12 – 14]. Namely, in

papers [12, 14], the algorithm of approximation of ravine data sets by the tangent to the linear branch of the ravine set has been proposed and described. Generally, such an approach is based on the well-known standard least-squares method of approximation [15, 16], but the necessity of obtaining the derivatives for root-polynomial functions leads to sophisticated calculations, especially for root-polynomial functions of higher order [12]. Using the advantages of the symbolic processor of the MATLAB system for scientific and technical calculation [3] has been proposed in the paper [13] for solving this sophisticated problem. A general description of this method of approximation has also been provided in the work [18].

In [13], a point-based approximation method was also proposed. It is simpler from a computational point of view than tangent approximation, therefore, if the convergence of this method is ensured, its use is preferable in terms of the use of computer resources. The conducted studies have shown that for 15–20 points of the approximated data set, the approximation error by points usually does not exceed several percent and approaches the error of the least squares method. Corresponding examples are given in [13].

And, finally, in the paper [14], generalized formulas for root-polynomial functions from second to sixth order are given, and conditions of convergence for the proposed methods for solving the tasks of interpolation and approximation are considered. In [14] are generally systemized all provided research, connected with root-polynomial functions and study its properties.

On the basis of the provided research, advanced computer software has been created using the programming language Python [17 – 20]. Firstly, this software has been created for solving the tasks of electron optics [21 – 27], but it became clear that the possibilities of using root-polynomial functions are significantly greater. Therefore, the aim of this paper is to describe elaborated computer software from the point of view of interpolation and approximation of ravine data sets by root-polynomial functions of different orders. Corresponding examples are given. Applying root-fraction-rational functions for solving interpolation tasks for stiff functions is also proposed.

2 BASIC CONCEPTION

As noted in the previous section, root polynomial functions were first proposed to be used to solve electron optics problems [21 – 27]. Namely, it was proven that the boundary trajectory of short-focus electron beams propagating in ionized gases corresponds to a ravine function with one global minimum and quasi-linear dependencies outside the minimum region [9, 11, 14]. It was also proven in [14] that the root polynomial functions correspond to the differential equation for the boundary trajectory of an electron beam propagating in an ionized gas. In general, this equation is based on the laws of electron optics [21 – 27] and the basic concepts of plasma physics [14, 29 – 31]. In general, these set of algebra-differential equations is written as follows [14]:

$$f=\frac{n_e}{n_{i0}-n_e};C=\frac{I_b\left(1-f-\beta^2\right)}{4\pi\varepsilon_0\sqrt{\frac{2e}{m_e}}U_{ac}^{1.5}};\frac{d^2r_b}{dz^2}=\frac{C}{r_b};\theta=\frac{dr_b}{dz}+\theta_s;$$

$$n_e=\frac{I_b}{\pi r_b^2};\ v_e=\sqrt{\frac{2eU_{ac}}{m_e}};$$

$$n_{i0}=r_b^2B_i pn_e\sqrt{\frac{\pi M\varepsilon_0 n_e}{m_eU_{ac}}}\exp\left(-\frac{U_{ac}}{\varepsilon_0 n_e r_b^2}\right); \qquad (1)$$

$$\gamma=\sqrt{1-\beta^2};\ \tan\left(\frac{\theta_{\min}}{2}\right)=\frac{10^{-4}Z_a^{4/3}}{2\gamma\beta^2};\ \tan\left(\frac{\theta_{\max}}{2}\right)=\frac{Z_a^{3/2}}{2\gamma\beta^2};$$

$$\overline{\theta^2}=\frac{8\pi\left(r_bZ_a\right)^2dz}{n_e}\ln\left(\frac{\theta_{\min}}{\theta_{\max}}\right),\ \beta=\frac{v_e}{c},$$

where z is the longitudinal coordinate, r_b is the radius of the boundary trajectory of the electron beam, I_b is the electron beam current, U_{ac} is the accelerating voltage, p is the residual gas pressure, n_{i0} is the concentration of residual gas ions on the beam symmetry axis, n_e is the beam electrons' concentration, f is the residual gas ionization level, B_i is the gas ionization level, $\theta_{\min}$ and $\theta_{\max}$ are the minimum and maximum scattering angles of the beam electrons, corresponding to Rutherford model, $\overline{\theta}$ is the average scattering angle of the beam electrons, ε_0 is the dielectric constant, v_e is the average velocity of the beam electrons, m_e is the electron mass, c is the light velocity, γ is the relativistic factor, Z_a is the nuclear charge of the residual gas atoms, dz is the length of the electron path in the longitudinal direction at the current iteration.

The numerical solving of the set of equations (1) can be described with high accuracy by root-polynomial functions, which in generalized form is written as follows [9, 10]:

$$r_b(z)=\sqrt[n]{C_nz^n+C_{n-1}z^{n-1}+\ldots+C_1z+C_0}, \qquad (2)$$

where n is the degree of the polynomial and the order of the root-polynomial function, and $C_0 - C_n$ are the polynomial coefficients.

Finding of coefficient $C_0 - C_n$ is the simple task,

and with the known set of basic points $\{P_1(z_1, r_1), P_2(z_2, r_2), \ldots, P_{n+1}(z_{n+1}, r_{n+1})\}$ it led to analytical solving the set of linear equations:

$$r_j - C_0 = \sum_{i=1}^{n} C_i z_j^i, \forall j\left(j = 1\ldots n, i \in N\right), \quad (3)$$

where N is natural number.

Interpolation error $\varepsilon(z)$ is estimated relative to numerical data $r_{num}(z)$, which has been obtained by solving the set of algebra-differential equations (1). This numerical data is considered as etalon values. Corresponded analytical relation for error estimation is follows [9, 10]:

$$\varepsilon(z) = \frac{\left|r_{num}(z) - r_{INT}(z)\right|}{r_{num}(z)} \cdot 100\%, \quad (4)$$

where r_{INT} is the result of interpolation using relation (2) and set of equations (3).

The analytical method of solving the set of linear equations (3) as well as the obtained sets of analytical relations for defining polynomial coefficients for root-polynomial functions (2) from second to sixth order are complexly presented in works [18, 20]. The general theory of root-polynomial functions is also given in [14]. Examples of solving such interpolation tasks are presented in papers [13].

It should be pointed out that the simulation of electron beam propagation in a soft vacuum of ionized gas is also an important industrial task. Solving this task is especially important for the elaboration and application in industry of high-voltage glow discharge electron guns [14]. Such guns are considered today as an advanced tools for providing in the soft vacuum such technological operations, as welding the contacts of electron devices [32, 33], deposition of ceramic coatings [33, 34], as well as refining refractory materials [35, 36].

Another concept for simulation relativistic electron beams propagating in plasma for microwave devices is based on the accurate processing and systematization of experimental data. This approach was proposed in [37, 38].

3 EXAMPLES OF SOLVING INTERPOLATION TASK

Example 1. Using root-polynomial function (2) of forth order interpolate the boundary trajectory of electron beam for the following parameters of simulation task: U_{ac} = 20 kV, I_b = 5 A, p = 3 Pa, operation gas nitrogen, start beam radius r_{b0} = 6 mm, start angle of beam convergence $\theta_0 = 12^0$, start point on z coordinate z_0 = 0.1 m.

Corresponded results of interpolation, obtained using elaborated software, is presented at Fig. 1.

For this task maximal error of interpolation is 2.9 % and average error in the range of z coordinate [0.1 m; 0.16 m] is 1,28 %. Should been pointed out, that for sixth-order root-polynomial functions values of errors are smaller: maximal value is 1.73 % and average value is 0.635 %. As proven in the [18, 20], generally even high-order functions give the smallest error of interpolation. Odd-order functions usually give the larger value of error. It has also been found that at large values of the accelerating voltage and small values of the current, the interpolation algorithm often diverges due to the small values of the minimum radius of the electron beam [18, 20].

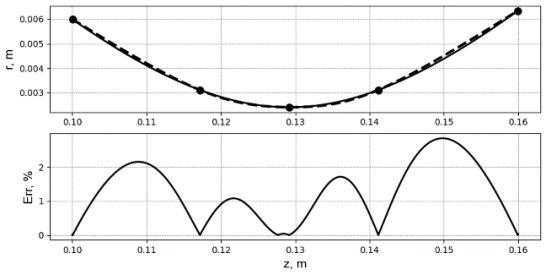


Figure 1: Solving of interpolation task for Example 1.

In the developed software, the set of equation (1) is first numerically solved using the fourth-order Runge-Kutt method [1, 2], and at the second stage, the obtained numerical data is transferred to the interpolation problem. To provide such access to a large amount of numerical data, the well-known tool for describing global variables in the Python programming language was used. Data transfer is implemented by analyzing the button click event using the functional programming paradigm [17 – 20]. Corresponding forth-order root-polynomial function for this example is as follows:

$$r(z) = \sqrt[4]{1.534 \cdot 10^{-3} z^4 - 7.9 \cdot 10^{-3} z^3 + 1.5343 \cdot 10^{-4} z^2 - \\ -1.32326 \cdot 10^{-4} z + 4.2854 \cdot 10^{-7}}.$$

Example 2. Interpolate manually using root-polynomial function ravine data set, presented in Table 1. It is clear that since the presented data set contains 7 values, only a sixth-order root-polynomial function is suitable for solving this task.

Table 1: Data set for ravine function of Example 2.

z, m	0.1	0.2	0.3	0.4
r, mm	7.0	5.0	4.0	3.5
z, m	0.5	0.6	0.7	–
r, mm	4.0	5.0	7.0	–

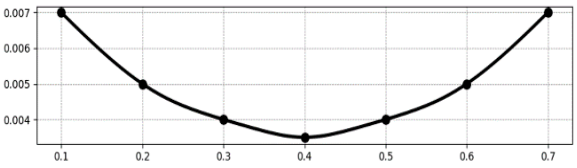


Figure 2: Solving of interpolation task for Example 2.

The root-polynomial function, obtained in this example, is written as follows:

$$r(z) = \sqrt[6]{\begin{matrix} 1.86\cdot10^{-10}z^6 - 4.464\cdot10^{-10}z^5 + \\ +4.4104\cdot10^{-10}z^4 - 2.2953\cdot10^{-10}z^3 + \\ +6.6558\cdot10^{-11}z^2 - 1.027\cdot10^{-11}z + 6.69\cdot10^{-13}. \end{matrix}} \quad (5)$$

The coefficient $C_0 - C_6$ in relation (5) has been calculated by solving the set of equations (3) analytically. Corresponding analytical relations for defining the values of these coefficients by the coordinates of the points $P_1 - P_7$ are given in [13, 14].

4 METHOD OF SOLVING THE APPROXIMATION TASK BY THE REFERENCE POINTS

Firstly, it should be pointed out that the task of approximation, in contrast to the interpolation task, is usually solved for a much larger number of numerical data points, which describes the ravine function, than the value of polynomial order *n*. This presumption is very important because only if this rule is followed, the error of approximation is relatively small.

Taking into account this presumption, the task of approximation is simplified by its transforming to interpolation task. Interpolation is provided between *n* +1 selected reference points, where *n* is the order of root-polynomial function. Corresponded method of interpolation has been described in the pervious section of this paper. On the next iteration interpolation is provided again between following *n* +1 reference points, which give the maximal error in the pervious step. And finally, the two best curves with minimal integrated error are selected and the optimal solution is located between them with using well-known dichotomy method [1 – 5] for the values of polynomial coefficients [13].

Given this assumption, the approximation problem is simplified by converting it into an interpolation problem. Interpolation is carried out between *n* + 1 selected reference points, where *n* is the order of the root polynomial function. The corresponding interpolation method was described in the previous section of the article. At the next iteration, interpolation is again carried out between the next *n* + 1 reference points for which the maximum approximation error was obtained in the previous step. And finally, using the well-known dichotomy method [1 – 5], the two best curves with the minimum integral error are selected for the values of the polynomial coefficients, and the optimal solution is found between them [13]. The last step of solving the approximation problem can also be carried out repeatedly with a decrease in the error value. As test experiments have shown, with the correct implementation of this method for 10-20 reference points, the approximation error is small and approaches the corresponding value for the well-known least squares method [15, 16].

But there are many particularities in realizing this approach that are connected with the mathematical properties of root-polynomial functions. Complexly, this approach is based on dividing the root-polynomial function into linear branches and the region of the local minimum. A corresponding flowchart of this algorithm is presented in the paper [12].

Since the number of reference points n_a in the approximation task is significant value, its coordinate z and *r* are formed in one vector and given in square bracket, for example: $\mathbf{Z} = [z_1, z_2, \ldots, z_{na}]$. Such approach of presentation structured data is fully corresponding to basic conception of MATLAB system for scientific and technical calculations.

Considering corresponding example.

Example 3. Solving manually the task of approximation using root-polynomial functions from fourth to sixth order ravine data set, presented in Table 2.

Solving of this task using proposed algorithm of approximation is presented in Figure 3.

Table 2: Data set for ravine function of Example 3.

z, m	0.1	0.2	0.3	0.4	0.5
r, mm	4.394	4.02	3.594	3.201	2.792
z, m	0.6	0.7	0.8	0.9	1.0
r, mm	2.501	2.294	2.203	2.305	2.495
z, m	1.1	1.2	1.3	1.4	1.5
r, mm	2.801	3.193	3.607	3.96	4.405

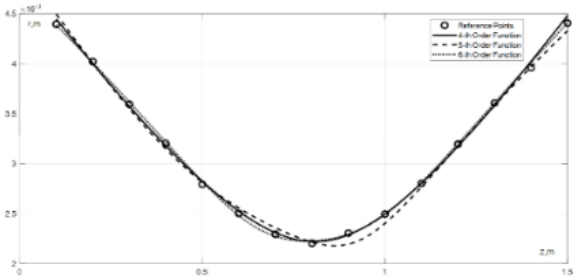


Figure 3: Solving of approximation task for Example 3.

It is clear from the obtained results of manual approximation that only the fifth-order root-polynomial function, as a function of odd order, gives a relatively large value of error. The total error of the approximation δ has been estimated as follows:

$$\delta = \sum_{i=1}^{n_a} \frac{\left| r_{ref}(z_i) - r_{APPR}(z_i) \right|}{r_{ref}(z_i)} \cdot 100\%, \quad (6)$$

where z_i are the z coordinate of the reference points, $r_{ref}(z_i)$ – r coordinate of the reference points, $r_{APPR}(z_i)$ – the corresponding values of the root polynomial-function at the reference points, $i = 1 \ldots n_a$. For solving approximation task value of δ was $\delta_4 = 1.693$ % for fourth-order root-polynomial function and $\delta_6 = 3.938$ % for sixth-order function. The fourth-order root-polynomial function has been obtained by solving this task is written as follows:

$$r(z) = \sqrt[4]{8.369 \cdot 10^{-10} z^4 - 2.65 \cdot 10^{-9} z^3 + 3.495 \cdot 10^{-9} z^2 - \\ -2.224 \cdot 10^{-9} z + 5.8068 \cdot 10^{-10}}.$$

Also, provided research has shown that the proposed approximation algorithm always gives a smaller error if the reference points are not located exactly uniformly but with a slight deviation along the radius coordinate [12 – 14].

5 ADVANCED POSSIBILITIES OF USING ROOT-FRACTIONAL RATIONAL FUHCTIONS

Unfortunately, the convergence of interpolation methods based on the use of relations (2, 4) significantly depends on the maximum value of the derivative for the set of interpolated data and is not always guaranteed [10, 14]. Therefore, the use of root-factional-rational functions has been proposed. In a generalized, form such a function is written as follows:

$$r(z) = \sqrt[n]{\frac{z^n + C_1 z^{n-1} + \ldots + C_{n-1} z + C_n}{C_{n+1} z^{n-1} + \ldots + C_{2n} z + C_{2n+1}}}. \quad (7)$$

Generally, the task of interpolation by using function (7) leads to the analytical solution of such a system of linear equations:

$$\sum_{i=1}^{n} C_i z_j^i - \left(r_j(z_j) \right)^n \sum_{i=n+1}^{2n+1} C_i z_j^i = -z_j^n; \; j \in N, \; j \le 2n+1. \quad (8)$$

The developed computer software makes it possible to study the properties of interpolation of various analytical dependencies using fractional rational functions. The mathematical function interpolated over the specified interval is entered into the appropriate text field. To ensure that the corresponding calculations are carried out at the programming level, the well-known concept of lambda-functions in the Python programming language is used [17 – 20]. Let's consider a relevant example.

Example 4. Describe the analytical function $r(z) = \exp\left(\frac{-z^4}{2}\right) + 1.5$ by a fourth-order root-fractional-rational function. Providing interpolation using the following reference points: $z_1 = 0.045$; $z_2 = 0.225$; $z_3 = 0.405$; $z_4 = 0.8775$; $z_5 = 1.3545$, $z_6 = 1.8$, $z_7 = 2.295$, $z_8 = 2.79$, $z_9 = 3.195$.

The result of interpolation for a given set of numerical data is presented in Fig. 4. Clear, that the interpolation error for this example does not exceed $5 \cdot 10^{-3}$ %.

Generally, the provided research shows, that root-fractional-rational functions written in the form (7) can be used to solve a wide range of practical scientific and engineering problems. Some of these problems are as follows.

1) To interpolate the magnetic field of focusing lenses in tasks of electron optics [28 – 33].
2) Interpolation of the values of probability distribution and probability density functions in problems of probability theory and mathematical statistics. The main feature here is the high degree of stiffness of such functions and the impossibility of obtaining an analytical expression for the inverse function in a simple algebraic form [15, 16].
3) Interpolation of membership function values in fuzzy logic problems.

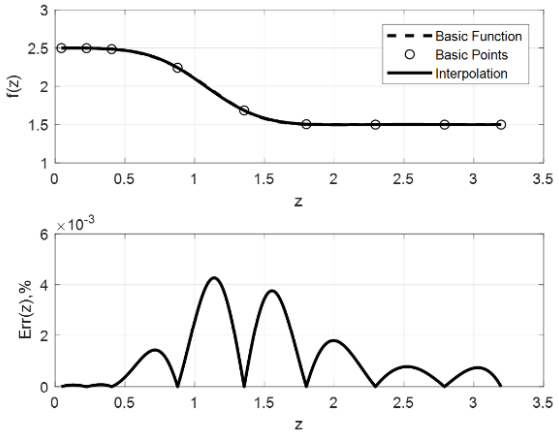


Figure 4: The result of interpolation by a fourth-order root-fractional-rational function for example 4

Generally, the main aspects of using root-fraction-rational functions (7) as an advanced method of interpolation and approximation for the stiff function are under study today.

6 OBTAINED RESULTS AND RECOMMENDATIONS

As it is clear from the examples presented in previous sections, elaborated computer software is a very effective tool for solving the tasks of interpolation and approximation. Root-polynomial functions are effective for solving such problems for ravine data sets, which can be obtained both by numerical calculations and experimentally. Obtained experience with the simulation of the boundary trajectory of an electron beam confirms that changing an enormous amount of numerical data to root-polynomial interpolation is the best way to solve the sophisticated tasks of electron-beam technologies. Such data transformation leads to a significant reduction in processor time and memory resources while maintaining the same accuracy in calculations. Therefore, such an approach is very effective in solving practical tasks, which lead to analyzing the ravine functions, including different physical tasks, tasks of probability theory, of economics, social science, etc. The advanced possibilities of using proposed approaches to solving different interpolation and approximation tasks are the subjects of future research.

7 CONCLUSIONS

The created computer software and the results of the provided research show that the use of root-polynomial functions is a very effective tool for interpolation and approximation of the data sets of ravine functions. Such an approach leads to a significant reduction of computer resources in the case of solving sophisticated mathematical tasks. The use of root-fractional-rational functions can also be considered an advanced possibility for solving the tasks of computation physics, probability theory, and fuzzy logic. The use of root-fractional-rational functions is especially effective for the interpolation of data sets with stiff functional dependences. Generally, the obtained results may be interesting to a wide range of specialists in the field of computational mathematics.

REFERENCES

[1] S.C. Chapra and R.P. Canale, Numerical Methods for Engineers, 7th ed. New York: McGraw Hill, 2014, 992 p.

[2] M.K. Jain, S.R.K. Iengar, and R.K. Jain, Numerical Methods for Scientific and Engineering Computation. New Delhi: New Age International Pvt. Ltd., 2010, 733 p.

[3] J.H. Mathews and K.D. Fink, Numerical Methods: Using Matlab. Amazon, 1998, 680 p.

[4] I.N. Bronshtein, K.A. Semendyayev, G. Musiol, and H. Mühlig, Handbook of Mathematics, 5th ed. Berlin: Springer, 2007, 1164 p.

[5] M. Abramowitz and I.A. Stegun, Eds., Handbook on Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Washington: National Bureau of Standards, 1964, 1046 p.

[6] G.M. Phillips, Interpolation and Approximation by Polynomials. Cham: Springer, 2023, 312 p. [Online]. Available: http://bayanbox.ir/view/2518803974255898294/George-M.-Phillips-Interpolation-and-Approximation-by-Polynomials-Springer-2003.pdf.

[7] N. Draper and H. Smith, Applied Regression Analysis, 3rd ed. New York: Wiley Series, 1998, 706 p. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9628935.

[8] I.V. Melnyk and A.O. Luntovskyy, "Estimation of Energy Efficiency and Quality of Service in Cloud Realizations of Parallel Computing Algorithms for IBN," in Future Intent-Based Networking: On the QoS Robust and Energy Efficient Heterogeneous Software Defined Networks, Lecture Notes in Electrical Engineering, vol. 831, M. Klymash, M. Beshley, and A. Luntovskyy, Eds. Cham: Springer, 2022, pp. 339–379. ISBN 978-3-030-92433-1. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-92435-5_20.

[9] I. Melnyk, S. Tuhai, and A. Pochynok, "Interpolation of the Boundary Trajectories of Electron Beams by the Roots from Polynomic Functions of Corresponded Order," in Proc. 2020 IEEE 40th Int. Conf. Electronics and Nanotechnology (ELNANO), 2020, pp. 28–33. [Online]. Available: https://ieeexplore.ieee.org/servlet/opac?punumber=9085228.

[10] I.V. Melnyk and A.V. Pochynok, "Study of a Class of Algebraic Functions for Interpolation of Boundary Trajectories of Short-Focus Electron Beams," System Researches and Information Technologies, no. 3, pp. 23–39, 2020. [Online]. Available: http://journal.iasa.kpi.ua/issue/view/13416/7188.

[11] I.V. Melnyk, S.B. Tuhai, and A.V. Pochynok, "Interpolation Functions for Describing the Boundary Trajectories of Electron Beams Propagated in Ionised Gas," in Proc. 15th Int. Conf. Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET–2020), 2020, pp. 79–83. [Online]. Available: https://ieeexplore.ieee.org/document/9088609.

[12] I. Melnyk and A. Pochynok, "Basic Algorithm for Approximation of the Boundary Trajectory of Short-Focus Electron Beam Using The Root-Polynomial Functions of the Fourth and Fifth Order," System Research and Information Technologies, no. 3, pp.

127–148, 2023. [Online]. Available: http://journal.iasa.kpi.ua/article/view/290474/284062.

[13] I. Melnyk, S. Tuhai, M. Skrypka, T. Khyzhniak, and A. Pochynok, "A New Approach to Interpolation and Approximation of Boundary Trajectories of Electron Beams for Realizing Cloud Computing Using Root-Polynomial Functions," in Information and Communication Technologies and Sustainable Development (ICT&SD 2022), Lecture Notes in Networks and Systems, vol. 809, S. Dovgyi, O. Trofymchuk, V. Ustimenko, and L. Globa, Eds. Cham: Springer, 2023, pp. 395–427. ISBN 978-3-031-46879-7. [Online]. Available: https://doi.org/10.1007/978-3-031-46880-3_24.

[14] I. Melnyk and A. Pochynok, "Theoretical Justification of Application Possibility of Different Order Root-Polynomial Functions for Interpolation and Approximation of Boundary Trajectory of Electron Beam," Radioelectronics and Communications Systems, vol. 66, no. 1, pp. 53–73, 2023. [Online]. Available: https://radioelektronika.org/issue/view/2023-02.

[15] E. Wentzel and L. Ovcharov, Applied Problems of Probability Theory. Mir, 2002, p. 432. [Online]. Available: https://mirtitles.org/2022/06/03/applied-problems-in-probability-theory-wentzel-ovcharov/.

[16] J.A. Gubner, Probability and Random Processes for Electrical and Computer Engineers. Cambridge, UK: Cambridge University Press, 2006. [Online]. Available: http://www.cambridge.org/gb/academic/subjects/engineering/communications-and-signal-processing/probability-and-random-processes-electrical-and-computer-engineers.

[17] P. Robbins, Python Programming for Beginners. Independently published, 2022, p. 114. [Online]. Available: https://www.amazon.com/Hweryho-Python-Programming-Beginners-Hands/dp/B0BTRRLCYZ/ref=zg_bs_285856_sccl_6/144-8112897-9282521?psc=1.

[18] W. McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter, 3rd ed. O'Reilly Media, 2023, p. 579.

[19] F. Chollet, Deep Learning with Python, 2nd ed. Manning, 2022, p. 509. [Online]. Available: https://www.amazon.com/Learning-Python-Second-François-Chollet/dp/1617296864/ref=zg_bs_285856_sccl_17/144-8112897-9282521?psc=1.

[20] M. Lutz, Learning Python, 5th ed. O'Reilly, 2013, p. 1643..

[21] S. Schiller, U. Heisig, and S. Panzer, Electron Beam Technology. New York: John Wiley & Sons, 1982, p. 509.

[22] M. Szilagyi, Electron and Ion Optics. Springer Science & Business Media, 2012, p. 539. [Online]. Available: https://www.amazon.com/Electron-Optics-Microdevices-Miklos-Szilagyi/dp/1461282470.

[23] S.J.R. Humphries, Charged Particle Beams. Courier Corporation, 2013, p. 834. [Online]. Available: https://library.uoh.edu.iq/admin/ebooks/76728-charged-particle-beams---s.-humphries.pdf.

[24] E. Kasper and P. Hawkes, Principles of Electron Optics: Applied Geometrical Optics. Elsevier Science, 1989, p. 1188. [Online]. Available: https://books.google.com.ua/books?id=XHEDruIilgEC&hl=uk&source=gbs_book_similarbooks.

[25] P. Grivet, P.W. Hawkes, and A. Septier, Electron Optics. Elsevier, 2013, p. 412. [Online]. Available: https://books.google.com.ua/books?id=jDgvBQAAQBAJ&hl=uk&source=gbs_book_similarbooks.

[26] J. Orloff, Handbook of Charged Particle Optics. CRC Press, 1997, p. 528. [Online]. Available: https://books.google.com.ua/books?id=_rYbAyf-iAkC&hl=uk&source=gbs_similarbooks.

[27] A.B. El-Kareh, Ed., Electron Beams, Lenses, and Optics. Academic Press, 2012, p. 428.

[28] B.M. Smirnov, Theory of Gas Discharge Plasma. Springer, 2015, p. 433. [Online]. Available: https://www.amazon.com/Theory-Discharge-Springer-Optical-Physics/dp/3319110640.

[29] M.A. Lieberman and A.J. Lichtenberg, Principles of Plasma Discharges for Materials Processing. New York: Wiley Interscience, 1994, p. 572. [Online]. Available: https://people.physics.anu.edu.au/~jnh112/AIIM/c17/Plasma_discharge_fundamentals.pdf.

[30] Yu.P. Raizer, Gas Discharge Physics. New York: Springer, 1991, p. 449. [Online]. Available: https://d-nb.info/910692815/04.

[31] A.A. Druzhinin, I.P. Ostrovskii, Y.N. Khoverko, N.S. Liakh-Kaguy, and A.M. Vuytsyk, “Low temperature characteristics of germanium whiskers,” Functional Materials, vol. 21, no. 2, pp. 130–136, 2014.

[32] A.A. Druzhinin, I.A. Bolshakova, I.P. Ostrovskii, Y.N. Khoverko, and N.S. Liakh-Kaguy, “Low temperature magnetoresistance of InSb whiskers,” Materials Science in Semiconductor Processing, vol. 40, pp. 550–555, 2015..

[33] T.O. Prikhna et al., “Electron-Beam and Plasma Oxidation-Resistant and Thermal-Barrier Coatings Deposited on Turbine Blades Using Cast and Powder Ni(Co)CrALY(Si) Alloys I. Fundamentals of the Production Technology, Structure, and Phase Composition of Cast NiCrAlY Alloys,” Powder Metallurgy and Metal Ceramics, vol. 61, no. 1–2, pp. 70–76, 2022. [Online]. Available: https://www.springer.com/journal/11106.

[34] T.O. Prikhna et al., “Electron-Beam and Plasma Oxidation-Resistant and Thermal-Barrier Coatings Deposited on Turbine Blades Using Cast and Powder Ni(Co)CrAlY(Si) Alloys Produced by Electron-Beam Melting II. Structure and Chemical and Phase Composition of Cast CoCrAlY Alloys,” Powder Metallurgy and Metal Ceramics, vol. 61, no. 3–4, pp. 230–237, 2022.

[35] S.V. Akhonin, B.A. Berezos et al., “Production of high-strength titanium alloy VT22 by electron beam melting method,” Advanced in Electrometallurgy, no. 3(131), pp. 8–14, 2018.

[36] T. Kemmotsu, T. Nagai, and M. Maeda, “Removal Rate of Phosphorous from Melting Silicon,” High Temperature Materials and Processes, vol. 30, no. 1–2, pp. 17–22, 2011.

[37] V.M. Lazurik, V.T. Lazurik, G. Popov, and Z. Zimek, “Two-parametric model of electron beam in computational dosimetry for radiation processing,” Radiation Physics and Chemistry, vol. 124, pp. 230–234, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0969806X1530133X.

[38] V. Lazurik and V. Moskvin, “Monte Carlo calculation of charge-deposition depth profile in slabs irradiated by electrons,” Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms, vol. 108, no. 3, pp. 276–281, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/0168583X95010521.

SECTION 3

Control and Automation

Proceedings of the 12th International Conference on Applied Innovations in IT (ICAIIT)

NOVEMBER 2024

Impact of Cyber Risks and Threats on the Critical Infrastructure Development: Visualization of Scientific Research

Aleksy Kwilinski[1] and Nataliia Trushkina[2]
[1]*The London Academy of Science and Business,*
Unit 3, Office A, 1st Floor, 6-7 St Mary At Hill, London, EC3R 8EE, England
[2]*Research Center for Industrial Problems of Development of the National Academy of Sciences of Ukraine,*
1a Inzhenernyi Lane, 61165 Kharkiv, Ukraine
a.kwilinski@london-asb.co.uk, nata_tru@ukr.net

Keywords: National Economy, Critical Infrastructure, Cyber Risk, Cyber Threat, Cyberattack, Cyberwars, Cybersecurity, Cyber defense, Cyberspace, Digital Environment, Development Strategy, Bibliometric Analysis, Trend Analysis.

Abstract: In the global world, there is a constant dynamic process that leads to structural changes (economic, social, organizational, environmental, etc.) of objects and infrastructure networks in a multi-component spatial system. At this stage, the issue of transformation of economic systems through the transition of key elements of critical infrastructures to a qualitatively new level of development due to adaptation to risks and threats of the external environment is becoming particularly relevant. One of these risks is cyberattacks on important infrastructure facilities in the energy, IT, financial, transport sectors. According to the European Agency for Network and Information Security (ENISA), the number of cyberattacks on critical infrastructure in the EU increased in 2019-2020 by 2 times, or from 150 to 300. According to expert estimates, global losses from cyberattacks in 2020 amounted to approximately 945 billion dollars. In this regard, this study is devoted to the analysis of the impact of cyber risks and threats on the development of critical infrastructure using bibliometric and trend approaches. The purpose of the article is to identify the main directions of scientific research, assess the dynamics of publishing activity, as well as identify key trends and gaps in this field. As part of the study, a bibliometric analysis of the database of scientific publications was carried out to identify the most cited works and authors, as well as a trend analysis to identify changes in research topics. The results of the study will allow obtaining a holistic view of the current state and prospects for the development of research in the field of critical infrastructure development, taking into account cyber risks and threats, which can contribute to the formation of more effective strategies and policies of cybersecurity and the protection of critical infrastructure objects in the countries of the world.

1 INTRODUCTION

In modern society, cyberattacks on critical infrastructure are becoming an increasingly serious threat and risk to the development of national economies of the world. Therefore, the fact that reliable protection against cyberattacks actively affects the economic, political, defense and other components of the national security of states is undeniable. It is obvious that the disruption of the functioning of critical infrastructure facilities can lead to the appearance of risks, emergency situations and crisis phenomena of economic systems of various levels.

Therefore, cyber risks are considered the main global risk for strategic sectors of critical infrastructure. The type of information and communication technology risks that infrastructure objects are exposed to has not changed in recent years, but the frequency of cyber incidents and the scale of their impact on enterprise activity have increased. According to Cybersecurity Ventures [1], global annual costs of cybercrime will reach 9.5 trillion dollars in 2024. Global losses from

cybercrime are predicted to grow 3.5 times over 2015-2025, or from 3 to 10.5 trillion dollars. Global spending on cybersecurity will increase to 1.75 trillion dollars by 2025 (for comparison: the volume of the global cybersecurity market was only 3.5 billion dollars in 2004 [2]).

Constant geopolitical tensions are one of the key factors focusing business leaders' attention on creating an effective cyber risk management strategy. According to the World Economic Forum's Global Security Outlook 2023 report [3], 74% of organizations indicated that global geopolitical instability has impacted their cyber strategy.

It should be noted that critical infrastructure businesses face an increased risk of disruption due to sophisticated cyber threats, from state-sponsored ransomware groups to supply chain vulnerabilities and new threats arising from ongoing geopolitical tensions.

Microsoft Digital Defense Report [4] showed that the number of cyberattacks targeting critical infrastructure had grown significantly and now accounts for 40% of all government attacks (20% in 2021). According to Verizon's 2023 Data Breach Report [5], the majority of attacks targeted government administrators, as well as organizations in the IT, finance, manufacturing, and professional services sectors.

In view of this, in the conditions that have developed today, it is important to ensure the appropriate level of security, including cybersecurity of critical infrastructure facilities. For this, it is necessary to substantiate conceptual provisions and develop practical recommendations for the formation of an appropriate secure information environment and the application of a risk-oriented approach to managing the development of critical infrastructure.

2 LITERATURE REVIEW

The analysis shows that most of the world's leading countries pay special attention to the formation and development of national cyber security systems and the protection of critical infrastructure facilities. Thus, in the European Union, the European Union Agency for Cybersecurity ENISA is the main body engaged in achieving a high common level of cybersecurity. ENISA developed a single pan-European concept of protection “Cyber Europe”, which was adopted in 2009 and is updated every two years. The basis of this concept is the safety and stability of objects of critical information infrastructure. The EU also ensured tactical actions and operational cooperation of countries at the pan-European level. In addition, the requirements for the protection of such important objects are determined by the national legislation of individual EU member states. Directive (EU) 2016/1148 of the European Parliament and the Council of July 6, 2016 “On measures to ensure a high general level of security of network and information systems on the territory of the Union” is of great importance in the European Union for protection against cyberattacks. NIS (The Security of Network and Information Systems) primarily concerns critical infrastructure companies and digital service providers (online marketplaces, online search engines, and cloud computing services).

On January 16, 2023, the New Cybersecurity Directive NIS 2 came into force, introducing mandatory information security measures and information security incident reporting requirements. Many companies in certain sectors will be subject to significant fines for failure to comply with these requirements. This Directive applies to organizations from the following sectors of critical importance for the economy: energy, transport, banking, financial market infrastructure, health care, drinking water supply, sewage systems, digital infrastructure, B2B management, IT services, public administration, and space research. Therefore, NIS 2 aims to improve the current state of cybersecurity in the EU by creating the necessary cyber crisis management framework, increasing the level of harmonization of security requirements and reporting obligations, as well as establishing a baseline level of cybersecurity risk management measures and reporting obligations in all critical sectors covered by the directive.

In the US, the organization that develops requirements in the field of cybersecurity is NIST – National Institute of Standards and Technology. For certain organizations in the USA, when building information systems, compliance with the requirements of the NIST Cybersecurity Framework is mandatory, in particular for objects of critical information infrastructure. This document appeared in 2014 and has been updated several times since then. In the United States of America, the Cybersecurity and Infrastructure Security Agency (CISA) deal with issues of cybersecurity in general and the protection of critical infrastructure. Facilitation of the broad exchange of critical

infrastructure information between owners and operators of critical infrastructures and government agencies responsible for their protection is carried out in accordance with the Critical Infrastructure Information Act of November 25, 2002. In this regard, the country's vulnerability to terrorism is reduced.

In Singapore, attention is drawn to the Cybersecurity Act, which is the framework for protecting critical information infrastructure from cybersecurity threats, taking measures to prevent, manage and respond to cybersecurity threats and incidents of critical infrastructures. The Cybersecurity Agency of Singapore is also interested.

In the Republic of China, security measures for critical information infrastructure are entrusted to the state. The country adopted the Law on Cybersecurity, in which Critical Information Infrastructure is interpreted as public communications and information services, public administration, water supply, finance, public services, electronic management and other critical information infrastructure, which in case of its destruction, violation functionality or data loss may actually threaten national security, national welfare, people's livelihoods, or the public interest.

In India, there is an Information Technology Act (2008), according to which critical information infrastructure (Critical Information Infrastructure) computer resources, the failure or destruction of which will affect the national security, economy and social welfare of the nation (Article 70). The legislative document delineates the sector of telecommunications and information technologies. That is, information technologies are considered as an independent, critically important sector of the national infrastructure. According to the Information Technology Act, the National Critical Information Infrastructure Protection Center (NCIIPC) of India was established in 2014.

It is worth noting that not only at the government level, various aspects of increasing the level of cyber protection of critical infrastructure objects are being discussed, and appropriate methodological recommendations and practical techniques for combating cyberattacks on critical infrastructures are being developed.

In recent years, in the scientific and educational environment, they are also actively engaged in research and development on the chosen research topic. The study of various aspects of the development of infrastructure as a multifunctional system that ensures the functioning of economic systems is given considerable attention in the works of leading scientists (M. Blaiklock [6]; B. Frischmann [7]; G. Hedtkamp [8]; R. Jochimsen [9]; W. Rostow [10]; U. Simonis [11]; H. Singer [12]; A. Youngson [13] and others). Based on the generalization of the existing scientific approaches to the formulation of the term “infrastructure”, they are conditionally systematized according to the following groups: system; resource; mechanism; systemic economic category; a component of the economic system; complex of types of economic activity; part of the economy; appropriate conditions (institutional, economic, social, investment, financial, environmental [14; 15; 16]); a component of the environment; component of the spatial system.

In the scientific literature (R. Wróbel [17]; B. Rathnayaka et al. [18]; D. Rehak et al. [19]; L. Shen et al. [20]; C. Scholz et al. [21] and others), many interpretations of the concept of “critical infrastructure” are used from different positions, including cyber security in the national security system. Summarizing the existing scientific developments regarding the conceptual apparatus, it was established that scientists usually understand critical infrastructure as: a complex system; its key components or components; critical infrastructure facilities; network structure; physical structure; organizational structures; institutes; institutions; institutions; set of assets; object of administrative and legal protection; object of cyber protection; security direction; one of the security tasks of the state; a component of the national infrastructure; a set of objects, technologies, state and scientific structures; object of state administration; component of information security; an element of the national security system of the state or region.

The theoretical analysis shows that scientists (A. Coning, F. Mouton [22]; D. Decker, K. Rauhut [23]; A. Elmarady, K. Rahouma [24]; M. Komarov et al. [25]; M. Gazzan, F. Sheldon [26]; S.Venkatachary et al. [27]; A. Golgota, U. Cerma [28] – in chronological order) are focused on conducting thorough scientific research on the development of risk management tools for implementation in the operations of critical infrastructure facilities.

Researchers take a detailed look at the challenges associated with cybersecurity and cyberterrorism for critical infrastructures. The papers highlight the

complexity of monitoring, managing, and measuring cybersecurity threats and discuss the critical need for analysis in this area, especially in the energy sector where command and control operations are performed in a networked environment. Despite effective risk management practices in the energy industry, it remains vulnerable to cyberterrorism, as evidenced by the Stuxnet attack. In addition, the economic consequences of cyberattacks on critical infrastructure are discussed, including the potential for significant financial losses and reputational damage. The authors provide practical advice on safeguards and defense mechanisms such as network segmentation, access control and encryption to help prevent cyberattacks. Scientists emphasize the need to continue developing effective risk management strategies and implementing appropriate measures to protect critical infrastructure objects from cyber threats. At the same time, scientists emphasize the creation of a digital security model in strategic sectors of critical infrastructure in various countries of the world.

Despite the wide range of scientific research on the chosen topic, the multifacetedness and debatable nature of certain issues require further development. Considering the above, it is relevant and necessary to analyse the impact of cyber risks and threats on the development of critical infrastructure using bibliometric and trend approaches.

3 METHODOLOGY

The theoretical and methodological basis of the study is the provisions of economic theory, institutional theory, theories of systems, network economy, digital economy, infrastructure, globalization, national interests of H. Morgenthau, possible conflicts of interest in the field of ensuring national stability according to J. Anderis, P. Martin-Breen, D. Chandler; concepts of information society, cyber and information security, sustainable development, strategic and energy management; models of national stability and development of the security environment.

The following general scientific methods were used in the research process: dialectical, historical, formal-logical, axiomatic, theory of logic and hypothetical-deductive, analysis and synthesis, induction and deduction, component analysis, trend analysis, bibliometric analysis, comparative analysis, analogy, classification, structural-logical generalization.

The information base of the research is statistical and analytical materials of Cybersecurity Ventures, Cybersecurity and Infrastructure Security Agency (CISA), The Cybersecurity Agency of Singapore, The European Agency for Network and Information Security (ENISA), Microsoft, the National Critical Information Infrastructure Protection Center (NCIIPC), Verizon, World Economic Forum (WEF), as well as legislative and regulatory documents the Critical Infrastructure Information Act of November 25, 2002; the Cybersecurity Directive NIS 2; the Cybersecurity Act of Singapore; Directive (EU) 2016/1148 of the European Parliament and the Council of July 6, 2016 “On measures to ensure a high general level of security of network and information systems on the territory of the Union”; Information Technology Act of India; the Law on Cybersecurity of China; the NIST Cybersecurity Framework; Pan-European concept of protection “Cyber Europe”.

4 RESEARCH RESULTS

Based on the bibliometric analysis, it was established that various aspects of ensuring the development of critical infrastructure, taking into account the impact of cyber risks and threats, are part of the long-term scientific interests of most leading foreign scientists. According to the title of articles, abstracts and keywords “Cyber risk”, “Critical infrastructure” or “Critical infrastructure facilities” in the international scientometric database Scopus, 1510 documents were found for the years 2002-2024.

As the analysis shows, these issues became especially relevant in the period from 2010. For 2010-2024, the number of scientific works increased from 24 to 134 or 5.6 times. During this period, the average growth rate was 13.1%. The following keywords are mostly used in publications: Cyber Security (521 documents), Risk Assessment (460), Network Security (446), Critical Infrastructures (424), Computer Crime (230), Risk Management (219), Cyber Attacks (213), Cyber Physical System (172), Security of Data (138), Industrial Control Systems (120), Risk Analysis (119), Internet of Things (111), Control Systems (97), SCADA Systems (93), Electric Power Transmission Networks (92), Risk Perception (91), Security

Systems (87), Cyber Threats (78), National Security (70), Resilience (69), Critical Infrastructure Protection (65) etc.

Among the most cited scientific works on the chosen subject, the following can be mentioned:

1) A. A. Cárdenas et al., "Attacks against process control systems: Risk assessment, detection, and response" [29] – the article examines how, using knowledge about the physical system under control, it is possible to detect computer attacks that change the behaviour of the target control system. By using knowledge about the physical system, it is possible to focus on the ultimate goal of the attack, rather than on specific mechanisms for exploiting vulnerabilities and hiding the attack. The authors analysed the protection and security of mechanisms, investigating the consequences of hidden attacks and ensuring that automatic mechanisms for responding to attacks do not lead the system to an unsafe state;

2) S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security" [30] – the article argues that industrial systems address security only partially, relying mostly on "isolated" networks and access-controlled environments. Monitoring and control systems, such as SCADA/DCS, are responsible for managing critical infrastructure that operates in environments where a false sense of security is common. The article explores the highly complex aspects of Stuxnet, the impact it may have on existing security considerations, and offers some thoughts on next-generation SCADA/DCS systems from a security perspective;

3) Y. Ashibani, and Q. H. Mahmoud, "Cyber physical systems security: Analysis, challenges and solutions" [31] – an analysis of security problems at different levels of the architecture of cyber-physical systems (CPS), an assessment of risks and CPS protection methods is given;

4) H. Sandberg, S. Amin, and K. H. Johansson, "Cyber physical security in networked control systems: An introduction to the issue" [32] – hypothesized that cyber-physical security applications of networked control systems (NCS) range from large-scale industrial to critical infrastructures such as water supply, transportation, and power grids. NCS security naturally depends on the integration of cyber and physical dynamics, and the different ways in which they are affected by the actions of decision makers. Emphasis is placed on developing a principled approach to NCS cyber-physical security;

5) I. Stellios et al., "A survey of IoT-enabled cyberattacks: Assessing attack paths to critical infrastructures and services" [33] – the article states that for some sectors, such as industry, intelligent networks, transportation and healthcare, the importance of cyber-attacks using the Internet of Things is obvious, since IoT technologies are part of mission-critical server systems. Therefore, the purpose of this study is threefold: to assess IoT-enabled cyberattacks using a risk-like approach to demonstrate their current threat landscape; identification of hidden and subliminal ways of attacks on critical infrastructures and services supported by the Internet of Things; study of mitigation strategies for all areas of application;

6) A. Nicholson et al., "SCADA security in the light of cyber-warfare" [34] – the article reviews current research and provides a consistent overview of SCADA security threats, risks, and mitigation strategies;

7) P. A. S. Ralston, J. H. Graham, J. L. Hieb, "Cyber security risk assessment for SCADA and DCS networks" [35] – In the article, the authors emphasize that the growing dependence of critical infrastructure and industrial automation on interconnected physical and cyber control systems has led to a previously unforeseen cybersecurity threat to supervisory control and data acquisition (SCADA) and distributed control systems (DCS). This article provides a broad overview of cybersecurity and risk assessment for SCADA and DCS, introduces the major industry organizations and government groups working in the field, and provides a comprehensive review of the literature to date.

Key publications that publish works on the subject of cyber risks and threats to the functioning of critical infrastructure facilities include: Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics (70 documents); IFIP Advances In Information And Communication Technology (40); ACM International Conference Proceeding Series (30); International Journal Of Critical Infrastructure Protection (23); Computers And Security (19 documents).

The main organizations dealing with the development of critical infrastructure in cyberspace are: Norges Teknisk-Naturvitenskapelige Universitet (28 documents); University of Piraeus (21); Austrian Institute of Technology (19); Pacific Northwest National Laboratory (18); University of Illinois

Urbana-Champaign (16); University of Jyväskylä (16); Sandia National Laboratories, New Mexico (14); Idaho National Laboratory (13); The George Washington University, Queensland University of Technology, University of Oxford (11 documents each); Virginia Polytechnic Institute and State University, The Grainger College of Engineering (10 documents each).

The results of the analysis show that most of the works on the studied issue are published by scientists from the United States (448 documents), United Kingdom (173), Italy (96), India (91), Germany (76), Greece (60), Norway (60), Australia (57), France (44), Spain (42), Canada (38), China (36), Sweden (35), Austria (33), Finland (28), Netherlands (28 documents), etc. In Ukraine, 26 documents were found based on the established search details.

By types of documents, works can be ranked as follows: 1st place is occupied by Conference Paper (790 documents or 52.3% of the total number of publications on the selected research topic); 2nd – Article (434 or 28.7%); 3rd – Book Chapter (145 or 9.6%); 4th – Conference Review (57 or 3.8%); 5th – Review (49 or 3.2%); 6th place – Book (29 documents or 1.9% of the total number of publications).

For the most part, scientific works on cyber risks and threats to critical infrastructure are published in the following fields of knowledge: Computer Science (953 documents or 30.6% of the total number of publications on this issue); Engineering (767 or 24.6%); Social Sciences (256 or 8.2%); Decision Sciences (231 or 7.4%); Energy (154 or 4.9%); Business, Management and Accounting (87 or 2.8%); Environmental Science (81 or 2.6%); Economics, Econometrics and Finance (42 documents). This shows that the researched topic is multidisciplinary and multifaceted.

The main sponsors that finance scientific publications on selected issues include the following: Horizon 2020 Framework Programme (75 documents); European Commission (64); National Science Foundation (52); U.S. Department of Energy (28); Engineering and Physical Sciences Research Council (22); U.S. Department of Homeland Security (18); Horizon 2020, National Natural Science Foundation of China (17 documents each); Norges Forskningsråd (15); Seventh Framework Programme (13); European Regional Development Fund (10 documents) etc.

As bibliometric analysis shows, in the international scientometric database Scopus there are 2,297 publications that contain the keywords “Cyber threat” and “Critical infrastructure”. The first publication on this topic appeared in 1997. Until 2005, publication activity was insignificant. And since 2005, scientists began to actively pay attention to this topic. For 2005-2024, the number of publications increased 26.5 times (from 10 to 265), and for 2015-2024 – 2.5 times (from 106 to 265).

Based on the analysis, it was established that the Scopus international scientometric database contains 3,537 publications that deal with various types of cyberattacks on critical infrastructure facilities. During 2005-2024, the number of such scientific works increased 17.3 times (from 18 to 312), during 2005-2015 – 8.6 times (from 18 to 155), during 2015-2024 – almost 2 times (from 155 to 312 documents).

Therefore, the analysis of publication activity confirmed that since 2010 there has been an increase in scientific interest in the study of the development of critical infrastructures in the context of global cyber risks and threats.

Further processing and analysis of bibliographic data was carried out using the VOSviewer software, which is a software tool for constructing and visualizing maps of bibliometric networks [36]. VOSviewer software was used to construct network maps of relationships between keywords based on bibliographic records from Scopus databases. The visual results of the obtained map of the bibliometric network are shown in *Figure 1*.

The map of the bibliometric network displays the frequency of use of terms by the size of the circle and the intensity of communication, and allows you to track variants of combinations of terms both within clusters and between them. The colour of the circle indicates that the keyword belongs to a certain cluster. The larger the diameter of the circle, the more often this term appears in scientific publications. Links on the map show the frequency of repetition of keywords in publications, while the smaller the distance between keywords, the stronger the connection between them [36].

According to *Figure 1* using the VOSviewer program, 785 keywords are systematized into 10 clusters, each of which symbolizes a separate direction of scientific research on the development of critical infrastructure, taking into account possible cyber risks and threats: the first cluster (red) contains 178 words, its share is 22.7% of the total

number of key concepts; the second (green) – 149 words (or 19%); the third (blue) – 110 words (14%); fourth (yellow) – 94 words (11.9%); fifth (purple) – 84 words (10.7%); sixth (turquoise) – 54 words (6.9%); seventh (orange) – 37 words (4.7%); eighth (brown) – 33 words (4.2%); ninth (dark pink) – 28 words (3.6%); the tenth cluster (pale pink) includes 18 words, which is 2.3% of the total number of terms from the selected topic.

Let's consider the 4 main clusters in more detail. The grouped keywords in the first cluster indicate that scientists consider the security aspects of the development of critical infrastructure, that is, from the standpoint of different types of security. This cluster contains keywords such as critical infrastructure, critical infrastructure resilience, cyber conflict, cyber event, cyber insurance, cyber protection, cybercrime, cyber challenge, cybersecurity strategy, cyberspace, cyber threat, data protection, defense, economic security, national security, national infrastructure, national cybersecurity strategy, potential attack, potential cyber threat and others.

The second cluster is related to the search and definition of tools for minimizing risks and threats to the functioning of critical infrastructure facilities. This cluster includes the following concepts: behavioral research, budget control, compliance control, computer control system, control system analysis, control system security, critical infrastructure protection, critical infrastructure security, cyber physical security, cyber physical system, cyber physical threat, multi agent system, network security, networked control system, risk assessment, risk management, SCADA system, and security system.

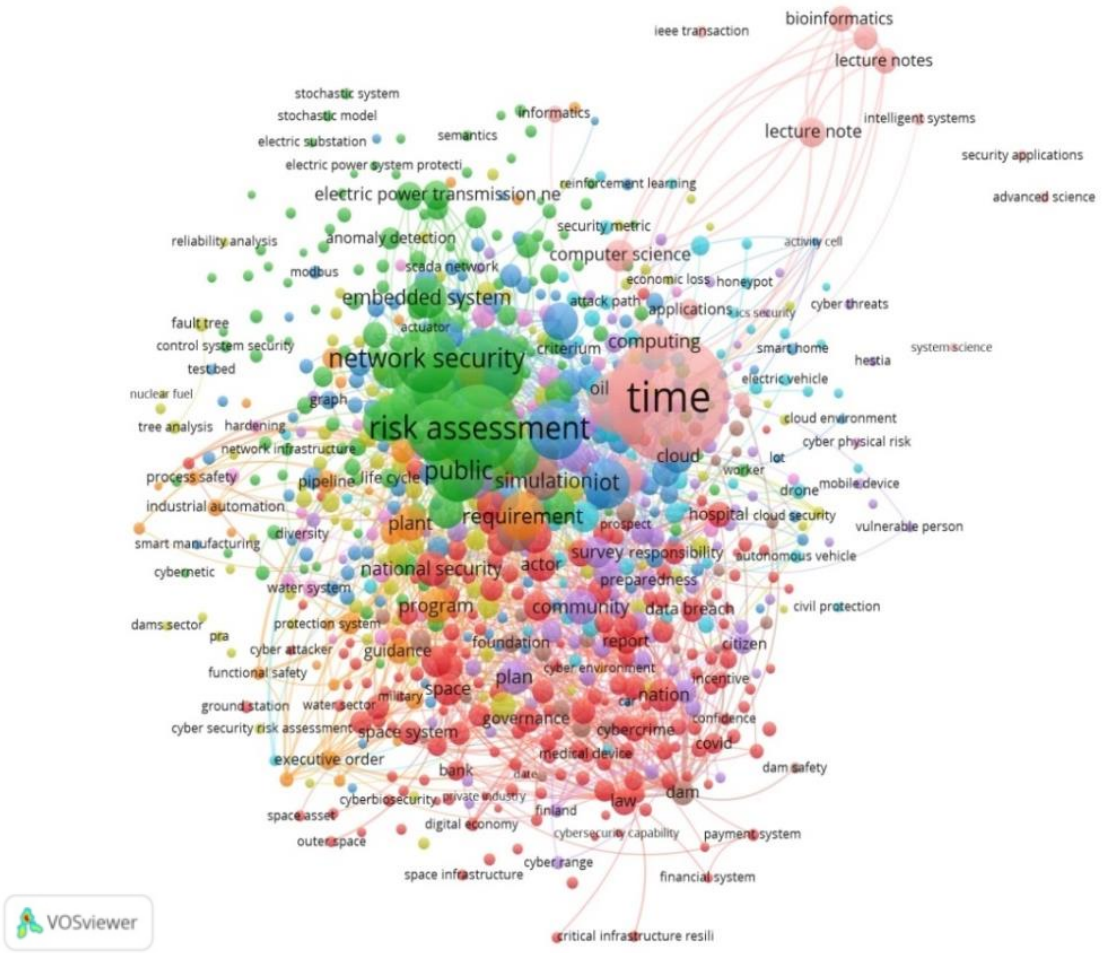


Figure 1: Network visualization of citations of articles on the impact of cyber risks on the critical infrastructure development, implemented using the VOSviewer tool[1].

[1] *Source:* built on the basis of data from the Scopus scientometrics database using the VOSviewer program.

The third cluster is related to the problems of forming a critical information infrastructure using modern information technologies and systems. The cluster includes Big Data, blockchain technology, cloud, cyberinfrastructure, diagnosis, IoT, industrial Internet, intelligent electronic device, machine learning, SCADA network.

The fourth cluster takes into account the processes of digital transformation of critical infrastructure. Emphasis is mostly placed on energy infrastructure. The cluster includes key terms such as communication protocol, critical energy infrastructure, defense strategy, digital instrumentation, digital technology, digital transformation, modernization, mathematical model, navigation, network environment, nuclear energy, safety critical system, security compliance, security risk assessment and others.

Thus, based on the results of the study, the following conclusions can be drawn:

1) The number of publications indexed in Scopus, whose titles, abstracts and keywords contain the terms “Cyber risk”, “Cyber threat”, “Cyberattack”, “Critical infrastructure”, “Critical infrastructure facilities”, “Critical infrastructure development” grows at an accelerated pace every year. Research on digital transformations of critical infrastructure has become increasingly popular since the 2000s. The key reason for the growing popularity of these scientific studies is the intensification of digitization processes and the introduction of digital technologies [37].
2) The term “critical infrastructure” has an interdisciplinary nature; it is used in studies of various branches of science, namely: it is found in publications on engineering, computer science, energy, ecology, social sciences, management, economics, decision science, etc.
3) Visualization of the network map of keywords based on bibliographic data made it possible to single out 10 clusters that characterize the priority areas of research: formation of a security environment, identification, adaptation, digitalization, development, cyber risk management, measures to reduce vulnerabilities and cyber threats, security, protection of critical infrastructure, development of a comprehensive national cyber security strategy.
4) The leaders in terms of the number of publications indexed in the international scientometric database Scopus are the USA, Great Britain, Italy, India, Germany, Canada, and China.

It should be noted that issues related to the definition of contextual and temporal patterns of the development of the views of scientists who investigate the impact of cyber risks and threats on the development of critical infrastructure in the countries of the world are gaining special relevance. For this, the toolkit of trend analysis is used – Google Trends.

Based on the trend analysis for the years 2004-2024, a high level of interest in the topics of “Critical Infrastructure” (on average 17 points) and “Cyber risk” (on average 13 points) was revealed worldwide (*Figures 2, 3*).

Looking back over time, we can see that in 2004, the popularity of the topic related to cyber risks was 0 points, and the development of critical infrastructure was 35 points. Since 2008, topics related to cybersecurity problems began to become popular, while the development of critical infrastructure began to decline. In 2012, the level of interest in both topics was 12 points. Since 2017, the level of interest in cyber risk management has been growing annually, while critical infrastructure has been shrinking. So, in 2019, the value of this indicator was 25 and 12 points, respectively; in 2023 – 47 and 24 points.

Queries are given points from 0 to 100, where 100 points means the location with the highest share of query popularity, 50 points - the location where the query popularity level is half as low as the first one. A score of 0 indicates a location for which there is insufficient data for the query in question. It is worth noting that the more points, the higher the proportion of relevant requests from all requests, and not their absolute number. Therefore, a small country, where queries with the words “Cyber risk” or “Critical infrastructure” make up 80% of all queries, will be assigned twice as many points as a large one, where only 40% of all queries contain this word.

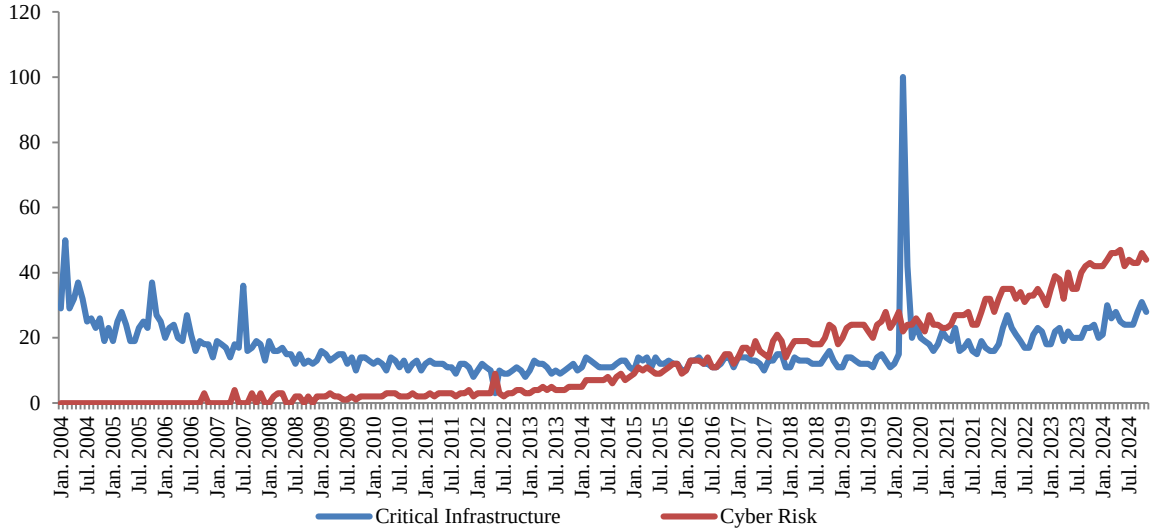
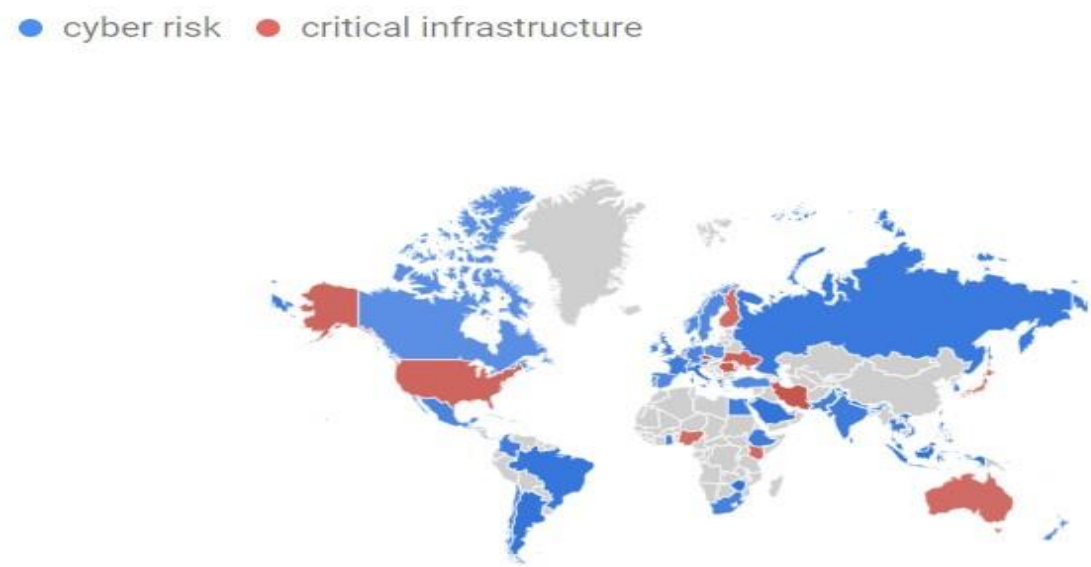


Figure 2: The dynamics of changes in search frequency in terms of the definitions of "Cyber risk" (blue colour) and "Critical Infrastructure" (red colour) in the world [3].



Note: the intensity of the colour depends on the percentage of requests.

Figure 3: Popularity of user searches for the topics "Cyber risk" and "Critical Infrastructure" in the world for the years 2004-2024 [4].

[3,4] *Source:* built using the Google Trends toolkit.

The topic, which is dedicated to the solution of urgent issues of the development of critical infrastructure, taking into account cyber threats and risks, has been updated again since 2022. This is due to the full-scale invasion of Russia on the territory of Ukraine and constant cyberattacks on critical infrastructure objects around the world.

Currently, in most countries of the world, in the last 20 years, the topics of cyber risks, threats, cyberattacks, cyberwars, cybersecurity, etc. are popular. In a number of countries, the share of requests for cyber risks exceeds 50% of the total number of requests in the respective country. For example, in Italy the value of this indicator is 81%, Switzerland – 77, Vietnam – 72, Korea and Great Britain – 71, France – 68, the Netherlands – 64, Portugal – 61, Germany – 59, Israel – 58, Spain – 56, Turkey – 55, Norway – 54, Canada and Poland – 50%. And in some countries, various aspects of development and implementation of critical infrastructure development strategies are gaining popularity and prevalence among sourcing. Thus, the level of popularity of the topic "Critical infrastructure development" in Finland and Japan is 51%, in Australia – 53, in the USA – 57, in Ukraine – 59, in the Czech Republic – 62, in Romania – 64% of the total number of requests in the respective country.

The following topics can be named among the leaders in terms of popularity: Cyber risk security (100 points); Cyber security (93); Cyber risk management (50); Risk management (49); Risk in Cyber security (35); Risk assessment (28); Cyber risk assessment (27); Risk management cyber security (26); Cyber risk insurance (17); Cyber insurance (15); Information security (14); Cyberattack (14); Cyber threat (13); Cybersecurity and Risk management (10 points).

The most common user search queries in countries around the world include: Critical infrastructure security (100 points); Critical infrastructure protection (85); Critical national infrastructure (64); Security of Critical infrastructure (51); Cybersecurity (36); Cybersecurity critical infrastructure (35); Critical infrastructure act (31); Critical infrastructure sectors (26); Critical infrastructure systems (25); Critical infrastructure definition (20); Critical information infrastructure protection (10 points).

The analysis shows that the subject of "Cyber threat" began to become more active since 2016. If in 2004 the level of popularity for this issue was 0 points, then in 2016 – 12, in 2017 – 15, in 2019 – 26, in 2022 – 39, in 2023 – 46 points. Users from Indonesia (84% of the total number of queries in the country), Malaysia (75%), Brazil (75%), Turkey (73%), Pakistan (72%), Korea (69%), Vietnam (69%), France (69%), Israel (69%), UK (67%), Portugal (60%), Italy (60%), Japan (57%), Poland (56%), Hungary (54%), Germany (54%), Canada (47%), USA (46%), Ukraine (45%), China (34%).

Since 2014, Cyberattack has become a popular search topic for users all over the world. The dynamics of popularity changed every year: in 2014, the level of interest was 14 points, in 2017 – 100, in 2020 – 40, in 2021 – 59, in 2023 – 61 points.

The leading countries in which users actively search for information on this issue include the following: The United Kingdom (94% of the total number of requests in the country), Pakistan and France (92%), Turkey (91%), Israel (90%), Canada, Switzerland, Sweden (87%), Germany (84%), USA and Japan (83%), Iran and Poland (80%), Republic of Korea (74%).

Among the leaders when searching for users using the keyword "Cyber threat" can be named Cyber security threat (100 points), Security threat (97), Threat intelligence (66), Cyber intelligence (63), Cyber threat intelligence (63 points). Popular queries on the subject of "Cyberattack" are Cybersecurity attack (100 points), Cyberattacks (54), Cyberattack types (17 points).

Thus, the study of trend patterns of publishing activity from the analysis of the relationship between the concepts of "Cyber risk" and "Critical infrastructure" proved the significant popularity of this issue in scientific circles, as well as its permanent growth.

At the same time, according to the results of the conducted trend analysis (based on the analysis of the dynamics of the number of publications on the researched topic, indexed by the Scopus scientometric database, for 2002–2024, the analysis of trends in user interest in this issue based on the Google Trends toolkit for the period 2004–2024 y.), as well as the generalization of existing conceptual developments in the scientific literature regarding the justification of national strategies for cybersecurity of critical infrastructure [22-28], it can be concluded that this problem is complex and multifaceted. It causes a synergistic effect on the national economy and is inextricably linked to

ensuring information security in the national security system.

5 CONCLUSIONS

In today's conditions, the world is on the threshold of new challenges, as there is a growing trend of increasingly complex and large-scale attacks on critical infrastructure facilities. Therefore, cybersecurity of critical infrastructure is one of the priorities of national security in the countries of the world in the conditions of a changing information space.

At the same time, it is extremely important to diagnose the state of cybersecurity as an effective tool, the procedure of which should include the following stages: assessment of the state of cybersecurity according to the NIST Cybersecurity Framework; determination of the target state of cybersecurity; development of recommendations (high-level design, specifications of technical architecture, operational models of cybersecurity); development of a road map for the implementation of recommendations; cybersecurity reassessment.

Based on the results of cyber diagnostics, critical infrastructure objects must receive a comprehensive assessment of digital security; recommendations for increasing the level of cyber maturity and preparedness for cyber incidents.

Since the nature of cyberspace is changing rapidly, the countries of the world need to improve the mechanism of regulatory and legal provision of cyber protection of critical infrastructure and information systems of objects, as well as the structure and content of national cyber security strategies. It is necessary to implement a comprehensive and comprehensive approach to the development of critical infrastructure in cyberspace, which should take into account constant changes in the security segment.

In addition, the governments of most countries of the world should pay attention to the development or improvement of national cyber security strategies. This was confirmed by the results of the bibliometric analysis. This strategic document should be understood as defining the concept, common goals, principles and priorities that should guide the country in solving cybersecurity problems; a description of the steps, programs and initiatives (i.e. the “Roadmap”) that the country intends to take to protect its critical infrastructure (including information) and to improve security, protection and resilience.

Early definition of the concept, goals and priorities allows governments to comprehensively consider cybersecurity within the framework of their national digital ecosystem, rather than at the level of a separate economic sector, a single goal or a response to a specific risk – it allows them to act strategically. National cybersecurity strategy priorities vary from country to country, so one country's focus may be on addressing risks to critical infrastructure, while another may focus on protecting intellectual property, building trust in the online environment, or raising public awareness of issues of cybersecurity or a combination of these tasks. The strategy should emphasize the importance of protecting critical infrastructure from cyber risks and recommend a comprehensive approach to risk management in the risk management system.

Finally, in the process of developing a national cybersecurity strategy, the government's vision must be translated into a coherent and feasible policy that helps it achieve its goal. This includes not only the activities, programs, and initiatives that must be accomplished, but also the resources allocated to those efforts and how those resources are used. Also, during this process, the indicators that will be used to achieve the desired results within the established budgets and deadlines should be determined.

Prospects for further research are the substantiation of the National Cybersecurity Strategy of critical infrastructure in Ukraine and the need to apply a risk-oriented approach to managing the development of critical infrastructure, taking into account the best European practices and developing practical recommendations for their implementation.

REFERENCES

[1] 2023 Official Cybercrime Report. Cybersecurity Ventures. https://www.esentire.com/resources/library/2023-official-cybercrime-report.

[2] S. Morgan, “Top 10 Cybersecurity Predictions and Statistics For 2023”, Cybercrime Magazine, Dec. 10, 2022. https://cybersecurityventures.com/stats/.

[3] Global Cybersecurity Outlook 2023. World Economic Forum. Jan. 18, 2023. https://www.weforum.org/publications/global-cybersecurity-outlook-2023/

[4] Microsoft Digital Defense Report 2022. Microsoft. https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/microsoft-digital-defense-report-2022.

[5] 2023 Data Breach Investigations Report (DBIR). Verizon. June 6, 2023. https://inquest.net/wp-content/uploads/2023-data-breach-investigations-report-dbir.pdf.

[6] M. Blaiklock, The infrastructure finance handbook: principles, practice and experience. London: Euromoney Books, 2014.

[7] B. M. Frischmann, Infrastructure: the social value of shared resources. New York: Oxford University Press, 2013.

[8] G. Hedtkamp, Die Bedeutungder Infrastruktur in makrookonomischer Sicht. Munchen: Osteuropa-Inst., 1996.

[9] R. Jochimsen, Theorie der Infrastruktur: Grundlagen der marktwirtschaftlichen Entwicklung. Tübingen: J.C.B. Mohr., 1966.

[10] W. W. Rostow, The Stages of Economic Growth. London: Cambridge University Press, 1962.

[11] U. E. Simonis, "Ecological Modernization of Industrial Society – Three Strategic Elements", In: F. Archibugi, P. Nijkamp (eds.), Economy and Ecology: Towards Sustainable Development. Volume 1: Economy & Environment. Dordrecht: Springer, pp. 119-137, 1989. https://doi.org/10.1007/978-94-015-7831-8_7.

[12] H. W. Singer, International Development: Growth and Change. New York: McGraw-Hill, 1964.

[13] A. Youngson, Overhead Capital: Study Development Economics. 1st ed. Edinburgh: Edinburgh University Press, 1967.

[14] H. Dzwigol, N. Trushkina, and A. Kwilinski, "The Organizational and Economic Mechanism of Implementing the Concept of Green Logistics", Virtual Econ., vol. 4(2), pp. 74-108, 2021. https://doi.org/10.34021/ve.2021.04.02(3).

[15] O. Hutsaliuk et al. "Factor-criteria assessment of greening prerequisites for transport infrastructure development in Ukraine", IOP Conf. Ser. Earth Environ. Sci., 1126(1), 012009, 2023. https://doi.org/ 10.1088/1755-1315/1126/1/012009.

[16] A. Kwilinski, and N. Trushkina, "Green Investments as Tools for Stimulating the Sustainable Financing of Logistics Systems Development", E3S Web of Conf., vol. 456, 01003, 2023. https://doi.org/10.1051/e3sconf/202345601003.

[17] R. Wróbel, "Dependencies of elements recognized as critical infrastructure of the state", Transp. Res. Proc., vol. 40, pp. 1625-1632, 2019. https://doi.org/10.1016/ j.trpro.2019.07.225.

[18] B. Rathnayaka et al., "Improving the resilience of critical infrastructures: Evidence-based insights from a systematic literature review", Int. J. Disast. Risk Re., vol. 78, 103123, 2022. https://doi.org/10.1016/ j.ijdrr.2022.103123.

[19] D. Rehak et al., " Dynamic robustness modelling of electricity critical infrastructure elements as a part of energy security", Int. J. Elec. Power, vol. 136, 107700, 2022. https://doi.org/10.1016/ j.ijepes.2021.107700.

[20] L. Shen, J. Li and W. Suo, "Risk response for critical infrastructures with multiple interdependent risks: A scenario-based extended CBR approach", Comput. Ind. Eng., vol. 174, 108766, 2022. https://doi.org/10.1016/ j.cie.2022.108766.

[21] C. Scholz, S. Schauer, and M. Latzenhofer, "The emergence of new critical infrastructures. Is the COVID-19 pandemic shifting our perspective on what critical infrastructures are?", Int. J. Disast. Risk Re., vol. 83, 103419, 2022. https://doi.org/10.1016/j.ijdrr.2022.103419.

[22] A. D. Coning and F. Mouton, "Bulk infrastructure management for facilities management", 2020 Resilience Week, RWS 2020, pp. 181-187, 92412622020, 2021. https://doi.org/ 10.1109/RWS50334.2020.9241262.

[23] D. K. Decker, and K. Rauhut, "Incentivizing Good Governance Beyond Regulatory Minimums: The Civil Nuclear Sector", J. Crit. Infrastruct. Policy, vol. 2, no. 2, pp. 19-43, 2021. https://doi.otrg/10.18278/jcip.2.2.3.

[24] A. A. Elmarady and K. Rahouma, "Studying Cybersecurity in Civil Aviation, Including Developing and Applying Aviation Cybersecurity Risk Assessment", IEEE Access, vol. 9, pp. 143997-144016, 2021. https://doi.org/ 10.1109/ACCESS.2021.3121230.

[25] M. Komarov et al., "Requirements for a taxonomy of cyber threats of critical infrastructure facilities and an analysis of existing approaches", In: Studies in Systems, Decision and Control, vol. 346. Berlin: Springer Science and Business Media Deutschland GmbH, pp. 189-205, 2021. https://doi.org/ 10.1007/978-3-030-69189-9_11.

[26] M. Gazzan and F. T. Sheldon, "Opportunities for Early Detection and Prediction of Ransomware Attacks against Industrial Control Systems", Future Internet, vol. 15, iss. 4, 144, 2023. https://doi.org/ 10.3390/fi15040144.

[27] S. K. Venkatachary et al., "Cybersecurity and cyber-terrorism challenges to energy-related infrastructures – Cybersecurity frameworks and economics – Comprehensive review", Int. J. Crit. Infrastruct. Prot., vol. 45, 100677, 2024. https://doi.org/10.1016/j.ijcip.2024.100677.

[28] A. Golgota and U. Cerma, "Securing Durres Port's Digital Transformation: Cybersecurity Strategy for Maritime Industry", 13th Mediterranean Conference on Embedded Computing, MECO 2024, 771, 2024. https://doi.org/10.1109/MECO62516.2024.10577771.

[29] A. A. Cárdenas et al., "Attacks against process control systems: Risk assessment, detection, and response", Proc. of the 6th International Symposium on Information, Computer and Communications Security, ASIACCS 2011, pp. 355-366, 2011. https://doi.org/10.1145/ 1966913.1966959.

[30] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security", 7th Annual Conf. of the IEEE Industrial Electronics Society, IECON 2011, pp. 4490-4494, 61200483, 2011. https://doi.org/10.1109/IECON.2011.6120048.

[31] Y. Ashibani and Q. H. Mahmoud, "Cyber physical systems security: Analysis, challenges and solutions", Comput. Secur., vol. 68, pp. 81-97, 2017. https://doi.org/10.1016/j.cose.2017.04.005.

[32] H. Sandberg, S. Amin, and K. H. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue", IEEE Control Systems, vol. 35, iss. 1, pp. 20-23, 7011179, 2015. https://doi.org/10.1109/MCS.2014.2364708.

[33] I. Stellios et al., “A survey of iot-enabled cyberattacks: Assessing attack paths to critical infrastructures and services”, IEEE Communications Surveys and Tutorials, vol. 20, iss. 4, pp. 3453-3495, 8410404, 2018. https://doi.org/10.1109/COMST.2018.2855563.

[34] A. Nicholson et al., “SCADA security in the light of cyber-warfare”, Comput. Secur., vol. 31, iss. 4, pp. 418-436, 2012. https://doi.org/10.1016/j.cose.2012.02.009.

[35] P. A. S. Ralston, J. H. Graham, and J. L. Hieb, “Cyber security risk assessment for SCADA and DCS networks”, ISA Transactions, vol. 46, iss. 4, pp. 583-594, 2007. https://doi.org/10.1016/j.isatra.2007.04.003.

[36] VOSviewer – Visualizing scientific landscapes. https://www.vosviewer.com.

[37] V. Khaustova, M. Kyzym, N. Trushkina, and M. Khaustov, “Digital transformation of energy infrastructure in the conditions of global changes: bibliometric analysis”, Proc. of the 12th Int. Conf. on Applied Innovations in IT (Koethen, Germany, Match 7, 2024). Koethen: Anhalt University of Applied Sciences, vol. 12, iss. 1, pp. 135-142, 2024. http://dx.doi.org/ 10.25673/115664.

Towards Automated Quality Control in Industrial Systems: Developing Markov Decision Process Model for Optimized Decision-Making

Katerina Mitkovska-Trendova[1], Robert Minovski[2], Verica Bakeva[3], Simeon Trendov[4] and Dimitar Bogatinov[5]

[1]*Faculty of Engineering Technology and Science, Higher Colleges of Technology, Al Mizn - Baniyas North, 20020 Abu Dhabi, UAE*

[2]*Faculty of Mechanical Engineering, Ss. Cyril and Methodius University, Rugjer Boshkovikj 18, 1000 Skopje, N. Macedonia*

[3]*Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Rudzer Boshkovikj 16, 1000 Skopje, N. Macedonia*

[4]*Anhalt University of Applied Sciences, Bernburger Str. 57, 06366 Köthen, Germany*

[5]*Military Academy Gen. Mihailo Apostolski, st. Vasko Karangelevski n.n., 1000 Skopje, N. Macedonia*

ktrendova@hct.ac.ae, robert.minovski@mf.edu.mk, verica.bakeva@finki.ukim.mk, simeon.trendov@hs-anhalt.de, dimitar.bogatinov@morm.gov.mk

Keywords: Markov Decision Processes, Performance Measurement Systems, Quality Control, Percent of Scrap, Costs Minimization, Automatization.

Abstract: In the context of rapidly evolving industrial environments, optimizing decision-making for quality control is crucial. This paper develops a Markov Decision Process (MDP) model aimed at enhancing automated quality control and reducing scrap in manufacturing systems, addressing challenges posed by complex and uncertain decision scenarios. The study focuses on improving the sub-key element of quality-accuracy within a Performance Measurement System (PMS) framework, specifically targeting scrap minimization and cost reduction. The research employs a mathematical model that integrates vector random processes, each representing critical factors such as machine condition, operator behaviour, tools, and materials. These factors are modeled as individual one-dimensional MDPs, which are combined to create a multi-dimensional MDP capable of monitoring and offering optimal policy for minimizing scrap rates and costs. The research methodology leverages advanced data analytics, statistical modeling, and real-time monitoring to accurately estimate transition probabilities and optimize policies. Different MDP models and methods are explored to enhance adaptability and iterative learning, allowing for optimal policy refinement over time. The proposed model is validated through its application to a real-world printing enterprise identified critical element, demonstrating a reduction in scrap and costs. This improvement underscores the model's effectiveness in practical settings, offering structured, subsystem-specific interventions that enhance manufacturing quality control. The results hold both theoretical and practical significance. Theoretically, the study contributes to the body of knowledge on MDP modeling for industrial quality control, providing a scalable approach that addresses complex interdependencies and decision-making under uncertainty. Practically, the model offers a robust tool for optimizing manufacturing processes, supported by modern IT systems, integration of advanced technologies, predictive maintenance, and data-driven decision-making. This integrated approach enables manufacturers to proactively identify and mitigate quality issues, enhancing operational efficiency, reducing waste, and driving continuous improvement in industrial systems.

1 INTRODUCTION

Enterprises today are compelled to evolve due to a convergence of technological advancements, globalization, shifting market demands, economic fluctuations, regulatory pressures, and internal transformations such as workforce and cultural changes. These factors collectively influence enterprise performance and require agility and innovation to maintain competitiveness in a dynamic environment. Various innovative approaches have

been introduced to address these changes, particularly in improving enterprise functionality and adaptability.

The belief that "You cannot improve what you cannot measure" reflects the prevailing mindset among today's business leaders. Performance Measurement Systems (PMS) emerged as a central focus in the early 1990s, becoming a key concept for guiding enterprises in a dynamic environment towards improvement and competitive advantage. Over time, numerous PMS concepts have been developed, all centered around a common goal: to segment the enterprise's operations into distinct areas and monitor performance by evaluating these areas.

Initially in our study, we aimed to improve a PMS by introducing and applying mathematical models to support decision-making in enterprise restructuring while coping with rapid changes. The PMS under consideration is referred to as COMPASS (COmpany's Management Purpose ASSistance) [1]. COMPASS focuses on key areas of success, including time, quality, costs, flexibility, and productivity. Due to their broad nature, these areas are further refined into sub-key elements of success (subKEs), resulting in a total of 18 subKEs. The extensive range of elements within PMSs poses challenges when it comes to generating actionable improvements, known as Success Factors. Each subKE may be influenced by multiple factors, and each factor can be enhanced through various actions. Given the complexity and interdependencies, action generation is typically approached heuristically, but modern systems increasingly leverage data analytics, real-time monitoring, and automation to support the decision making.

Therefore, in this study, when we refer to new approaches, we primarily focus on mathematical modeling. However, it is essential to consider related IT advancements supported by mathematical modeling which are further observed. Regarding mathematical modeling specifically, a diverse range of models and tools have been implemented in recent years to enhance organizational performance.

Markov Decision Processes (MDPs) are utilized to model specific aspects of enterprise operations due to their proven capabilities and advantages. MDPs are supported by a well-developed theory and have become a mature modeling tool. Their successful application is largely due to the availability of efficient algorithms for finding optimal solutions. Furthermore, MDPs provide a flexible framework for solving optimization problems across a wide range of fields and are particularly valuable in sequential planning applications where accounting for process uncertainty is critical [2].

Applying MDP to solve a restructuring, decision-making, or planning problem within an enterprise, while managing the dimensionality of the problem, yields an optimal policy based on a specified optimality criterion. If this criterion aligns with a key performance metric for a specific subKE (e.g., percent of scrap for the subKE quality-accuracy), it implies that the subKE is being optimized. This approach enables the generation of success factors or improvement actions for various elements of the PMS using quantitative methods supported by software, rather than relying on heuristic approaches.

To accurately assess the performance of subKEs, a specific measure must be assigned to each subKE. The primary objective of monitoring and recording these measures is to identify issues and generate actions for continuous improvement. Additionally, it is essential to pinpoint the influential factors contributing to a given situation.

Taking all of this into consideration, the primary research challenge of this study was to develop a mathematical model based on MDP to enhance one aspect of the PMS, quality-accuracy, by managing its key measure: the percent of scrap.

Further we look at the MDPs extensive utilization in real-world scenarios that require decision-making under uncertainty, highlighting the pivotal role of modern IT systems. Advanced IT tools facilitate data collection, processing, and real-time analysis, enabling more precise modeling and optimization of complex decision processes, thereby enhancing the practical implementation of MDPs. They are particularly valuable in enhancing the automation of industrial processes by improving efficiency, adaptability, and reliability. MDPs can optimize various aspects, including process optimization and control, quality control and inspection, predictive maintenance, inventory, supply chain management, logistics, scheduling and resource allocation, adaptive process control in dynamic environments, etc. By leveraging MDPs, automated industrial systems can effectively manage complex, interdependent decisions, striking a balance between short-term costs and long-term efficiency and sustainability. This results in highly efficient, resilient, and sustainable enterprise environments.

2 LITERATURE REVIEW

The evolution of PMS has been driven by the need to better understand, measure, and improve organizational performance beyond traditional financial metrics. Early systems primarily focused on

financial indicators, which often failed to capture the complexity and multidimensional nature of modern business operations [3]. This limitation paved the way for the development of more holistic frameworks, such as the Balanced Scorecard by Kaplan and Norton [4], which integrated financial measures with non-financial perspectives, including customer satisfaction, internal processes, and organizational learning and growth. Concurrent with the development of the Balanced Scorecard, Total Quality Management (TQM) approaches, championed by Deming [5] and Juran [6], emphasized continuous improvement, customer-centric performance metrics, and process optimization. These methodologies highlighted the need for aligning strategic goals with operational measures, laying a foundation for comprehensive PMS approaches. As businesses faced rapid globalization, technological advances, and increased complexity, PMS evolved further to incorporate dynamic and flexible measurement tools such as the implementing MDPs. Modern systems increasingly leverage data analytics, real-time monitoring, and automation to provide more granular insights and facilitate agile decision-making [7]. With advances in data collection, artificial intelligence, and business intelligence systems, PMS has become more sophisticated, enabling organizations to optimize their performance and respond effectively to market changes [8]. Today, PMS is recognized as a critical tool for guiding strategic initiatives, measuring success, and fostering continuous improvement and competitiveness.

On the other hand, advancements in MDPs have allowed their adoption across a wide range of applications, greatly supported by the latest developments in Information Technology (IT). MDPs are a mathematical framework for modeling decision-making in environments where outcomes are partly random and partly under the control of a decision-maker. The foundation of MDPs dates back to the work of Andrey Markov in the early 20th century, who introduced the concept of Markov chains to model stochastic processes [9]. MDPs extend these concepts by incorporating actions and rewards, allowing for optimization of long-term outcomes through sequential decision-making. The formalization of MDPs for decision problems was developed in the 1950s by Richard Bellman, who introduced the principle of dynamic programming as a method to solve MDPs and coined the term "Bellman equation" to describe the recursive decomposition of value functions [10]. This framework has since become foundational in operations research, artificial intelligence, and control theory, providing a structured way to handle uncertainty in decision-making. Throughout the late 20th century, MDPs were further refined with the development of exact algorithms such as value iteration and policy iteration [2]. However, many real-world applications present challenges such as large or continuous state spaces, which led to advancements in approximation methods and reinforcement learning algorithms, including Q-learning and policy gradient methods [11]. For this study it was important to examine the development of application of MDPs in quality control. For instance, Markov Chains to model and simulate transitions between different product quality states in manufacturing processes were used in [12]. Their approach identifies influential factors and proposes measures for continuous quality improvement, highlighting the value of stochastic modeling in industrial quality management. Today, MDPs are widely used in fields such as robotics, automated control systems, finance, healthcare decision-making, and artificial intelligence for complex planning and optimization problems under uncertainty [13]. A methodology for determining the transition probabilities of MDP for quality accuracy improvement inside a PMS framework was proposed in [14]. Modern applications continue to push the boundaries of scalability and efficiency in different MDP models and solutions through techniques like Monte Carlo methods, approximate dynamic programming, deep reinforcement learning, Q-learning, etc. Some of the reviewed references like [15], [16], [17], [18], [19], [20] were found insightful for this work.

3 METHODOLOGIES

As mentioned before, the PMS methodology COMPASS served as the foundational framework within which the research was conducted. The approach leverages mathematical modeling to address real-world challenges, utilizing operations research models and methods, with a particular focus on MDP models and the policy iteration optimization technique. Extensive literature review was conducted in order to select the right model for the modeled quality control problem. During the modeling and the application of the mathematical model to a specific enterprise problem, various management techniques were employed to analyze the organization, complemented by statistical methods for data collection and processing. To identify and address the causes of scrap, quality management tools such as

data collection lists, process flowcharts, Ishikawa diagrams, histograms, scatter plots, and control charts were utilized. This research developed methodologies to calculate transition probability matrices and revenue (cost) matrices for each factor individually, as well as collectively, to monitor the percentage and cost of scrap for each influential factor and overall [14]. Custom software is developed to perform the calculations with the provided data.

All stages of the research were synthesized into a model for generating optimal decision policies and success factors, specifically targeting the management of the critical quality-accuracy element within the selected PMS. This model is designed to facilitate practical implementation within a real enterprise, enhancing the functionality and effectiveness of its PMS.

The methodology introduced in [14] has influenced subsequent research by providing a robust framework for applying MDP models in quality management, facilitating better understanding and control of key performance factors through precise probabilistic modeling. It has contributed the way for more accurate and data-driven decision-making strategies across various applications in quality and operational management [15].

4 MATHEMATICAL MODELING OF A REAL SYSTEM USING FOUR-DIMENSIONAL MDP

The model for managing quality-accuracy, specifically focusing on scrap and scrap cost management, serves as a decision-making support tool for stochastic, multi-stage planning processes. This model represents a system comprising a single job station, consisting of one machine and one operator. Due to prolonged usage, both the machine and its tools experience deterioration, resulting in decreased quality and increased scrap production. However, additional factors can contribute to scrap generation. SubKE quality-accuracy can be influenced by various elements, such as the machine, operator, tools, materials, environment, and methods. To strike a balance between model complexity and realism, this study focuses on the most significant factors—machine, operator, tools, and materials. However, with advancements in modern IT, the model can be further expanded, addressing previous limitations and enhancing its capabilities. The model aims to provide a detailed breakdown of scrap production by cause, as well as overall scrap levels. To achieve this, stochastic processes represented by random variables are defined to describe the conditions of the machine, operator, tools, and materials, specifically in terms of their contributions to scrap production. Subsequently, a vector random process is constructed, comprising these four individual stochastic processes to capture both the individual and cumulative scrap production for the system. At the conclusion of each production cycle, the conditions of the four factors, in terms of the percentage of scrap they generate, are recorded and classified into a finite number of states, which represent the values of the random processes. Historical data was used to determine the transition probabilities for each possible state change between production cycles, for each influential factor. One of the major challenges today lies in the need for extensive data to implement these models, which is difficult to obtain using traditional methods. However, advancements in modern IT provide powerful tools for capturing, recording, processing, and utilizing data effectively in determining the decision-making policy. Since the transition probabilities are independent of the states from previous cycles, these stochastic processes can be modeled as discrete-time, finite, homogeneous Markov chains. Finite action spaces are defined for each Markov chain, representing available decision alternatives. The revenue structure associated with each process yields matrices corresponding to all possible transitions, with the revenue function reflecting gains or losses through scrap percentages and costs for each transition step. As a result, four one-dimensional MDPs are derived, which are then combined into a four-dimensional MDP represented as a vector random process, with a specifically designed action space and revenue structure based on the one-dimensional MDPs. The first one-dimensional MDP is described by the random process “the condition of the machine after every run”, as one of the most important influence factors or cause for scrap identified in quality-accuracy management, and for quality measure the percent of scrap is chosen. For the random variable X_n^1 which is the condition of the machine in a discrete moment n, it is assumed that the stochastic process $\{X_n^1 | n \in \mathbb{N}\}$ is homogeneous Markov chain. At any point of time n, the condition of the machine can be classified in one of several possible states and the random variable X_n^1 in a given moment n, takes values from the defined state space for the condition of the machine. It is assumed that in every discrete moment of time the random variable X_n^1, takes values from the same state space, and further for simpler notation, this random process will be denoted only by X_1 for every transition

moment and this is valid for the other stochastic processes. It is assumed that all random variables describing the given Markov chains are mutually independent. The second random process X_2, i.e. the second Markov chain, which is "the condition of the operator after every run of the work place", expressed by the caused percent of scrap from the operator. The third one-dimensional MDP is the random process X_3, which is "the condition of the machine tools after every run of the work place", expressed by the caused percent of scrap from the tools. The fourth one-dimensional MDP is the random process X_4, which is "the condition of the materials after each run of the work place", expressed by the caused percent of scrap from the materials. Let the random process X_l, $l \in \{1, 2, 3,4\}$, takes $|R_{X_l}| = n_l$ values. For example, let the sets of values for these stochastic processes are $R_{X_l} = \{x_1^l, x_2^l, x_3^l\}$, $n_l = 3$. To simplify, the same notations for the states of the one-dimensional MDPs and the values of associated random variables are used. A_l, denote the sets of primary actions (decisions) for the one-dimensional MDPs and $|A_l| = m_l$ are their numbers, for $l \in \{1, 2, 3,4\}$. For example, let $A_l = \{a_1^l, a_2^l, a_3^l\}$, $|A_l| = m_l = 3$.

One significant limitation of the MDP model is the challenge in accurately determining transition probabilities. Typically, historical data are used to estimate these probabilities, but such data are often unavailable, hard to collect, or outdated. Additionally, transition probabilities can be influenced by a variety of changing factors in the environment, making them susceptible to fluctuation over time. As a result, the values of these probabilities may shift, leading to potential inaccuracies in the model and impacting the reliability of the decision-making process. In an automated environment, systematically tracking causes, and recording and categorizing failure data can help address this challenge. By continuously collecting accurate data on failure patterns and updating the MDP model accordingly, the system can adapt more effectively to changes in the operational environment, improving the accuracy of transition probabilities over time. This model assumes that transition probabilities do not change over time, i.e. Markov chains are homogenous. The transition probabilities are denoted with $p_{ij}^k(l)$, and they are the conditional probabilities that the random variable X_l takes value x_j^l if its previous value was x_i^l, under the influence of the action a_k^l,$i \in \{1, 2, 3\}, j \in \{1, 2, 3\}, k \in \{1, 2, 3\}, l \in \{1, 2, 3,4\}$. For the given example, the matrices of the transition probabilities are:

a_k^l	x_1^l	x_2^l	x_3^l
x_1^l	$p_{11}^k(l)$	$p_{12}^k(l)$	$p_{13}^k(l)$
x_2^l	$p_{21}^k(l)$	$p_{22}^k(l)$	$p_{23}^k(l)$
x_3^l	$p_{31}^k(l)$	$p_{32}^k(l)$	$p_{33}^k(l)$

Each transition probability matrix is followed by revenue or cost matrix.

a_k^l	x_1^l	x_2^l	x_3^l
x_1^l	$c_{11}^k(l)$	$c_{12}^k(l)$	$c_{13}^k(l)$
x_2^l	$c_{21}^k(l)$	$c_{22}^k(l)$	$c_{23}^k(l)$
x_3^l	$c_{31}^k(l)$	$c_{32}^k(l)$	$c_{33}^k(l)$

The state space $\mathcal{S}$ for the four-dimensional MDP is defined as the set of all ordered quadruplets formed by the elements of the value sets of X_1, X_2, X_3, and X_4, and that is $\mathcal{S} = \left\{\left(x_i^1, x_j^2, x_k^3, x_l^4\right), i \in \{1, 2, \dots, |R_{X_1}|\}, j \in \{1, 2, \dots, |R_{X_2}|\}, k \in \{1, 2, \dots, |R_{X_3}|\}, l \in \{1, 2, \dots, |R_{X_4}|\}\right\}$ and it consists of $|\mathcal{S}| = |R_{X_1}| \cdot |R_{X_2}| \cdot |R_{X_3}| \cdot |R_{X_4}| = n_1 \cdot n_2 \cdot n_3 \cdot n_4$ states. The number of all possible transitions between the states of the system is calculated by $|\mathcal{S}|^2$. The action space is defined similarly as the state space $\mathcal{A} = \left\{\left(a_i^1, a_j^2, a_k^3, a_l^4\right), i \in \{1,2, \dots, |A_1|\}, j \in \{1,2, \dots, |A_2|\}, k \in \{1,2, \dots, |A_3|\}, l \in \{1,2, \dots, |A_4|\}\right\}$ and it consists of $|\mathcal{A}| = |A_1| \cdot |A_2| \cdot |A_3| \cdot |A_4| = m_1 \cdot m_2 \cdot m_3 \cdot m_4$ actions. Using the fact that the random processes are independent, a method for calculating the joint transition probabilities is proposed, knowing those transition probabilities in the one-dimensional MDPs. The number of transition probabilities is $|\mathcal{S}|^2 \cdot |\mathcal{A}|$. For the transition $\left(x_{i_1}^1, x_{i_2}^2, x_{i_3}^3, x_{i_4}^4\right) \xrightarrow{\left(a_{k_1}^1, a_{k_2}^2, a_{k_3}^3, a_{k_4}^4\right)} \left(x_{j_1}^1, x_{j_2}^2, x_{j_3}^3, x_{j_4}^4\right)$, where $i_1, j_1 \in \{1, 2, \dots, |R_{X_1}|\}$, $i_2, j_2 \in \{1, 2, \dots, |R_{X_2}|\}$, $i_3, j_3 \in \{1, 2, \dots, |R_{X_3}|\}$, $i_4, j_4 \in \{1, 2, \dots, |R_{X_4}|\}$, $k_1 \in \{1,2, \dots, |A_1|\}$, $k_2 \in \{1,2, \dots, |A_2|\}$, $k_3 \in \{1,2, \dots, |A_3|\}$, $k_4 \in \{1,2, \dots, |A_4|\}$, the transition probability is calculated by $p_{i_1 j_1}^{k_1}(1) \cdot p_{i_2 j_2}^{k_2}(2) \cdot p_{i_3 j_3}^{k_3}(3) \cdot p_{i_4 j_4}^{k_4}(4)$, and the corresponding revenue is calculated by $c_{i_1 j_1}^{k_1}(1) + c_{i_2 j_2}^{k_2}(2) + c_{i_3 j_3}^{k_3}(3) + c_{i_4 j_4}^{k_4}(4)$.

The primary transition matrices are stochastic, $\sum_{j_l} p_{i_l j_l}^{k_l}(l) = 1$, so $\sum_{j_1} \sum_{j_2} \sum_{j_3} \sum_{j_4} p_{i_1 j_1}^{k_1}(1)\, p_{i_2 j_2}^{k_2}(2) p_{i_3 j_3}^{k_3}(3) p_{i_4 j_4}^{k_4}(4) = \sum_{j_1} p_{i_1 j_1}^{k_1}(1) \sum_{j_2} p_{i_2 j_2}^{k_2}(2) \sum_{j_3} p_{i_3 j_3}^{k_3}(3) \sum_{j_4} p_{i_4 j_4}^{k_4}(4) = 1$, i.e. the new matrices are also stochastic. For example, if $n_1 = n_2 = n_3 = n_4 = 3$, $m_1 = m_2 = m_3 = m_4 = 3$, the system has 12 primary transition

matrices 3x3, and that is $9 \cdot 12 = 108$ primary transition probabilities. The number of primary revenues is the same. The state space has $3^4 = 81$ states, and the number of actions in the action space is the same. The number of all possible transitions between the states of the system is $3^8 = 6561$, and $|\mathcal{S}|^2 \cdot |\mathcal{A}| = 3^8 \cdot 3^4 = 3^{12} = 531441$ is the number of the joint transition probabilities. The number of the joint revenues is the same as the number of the joint transition probabilities. Exhaustive enumeration of all stationary policies is only practical in problems with small dimensions. In this model the number of all stationary policies is $|\mathcal{A}|^{|\mathcal{S}|}$ and it is a very big number. For the example, this number is $81^{81} \approx 3{,}87 \cdot 10^{154}$. So, the ranking of all stationary policies is not practical and the focus is on finding the optimal decision i.e. planning policy using some optimality method that can handle these dimensions. Later the policy iteration method is chosen (with and without discount rate). Convergence and optimal solution existence are considered in the research for the real data, collected in concrete enterprise. Table 1 gives the summary of the state space and the action space for the created four-dimensional MDP, for the MDP real example.

Table 1: State and Action spaces for the MDP example.

State Number	State	Action Number	Action
1	$(x_1^1, x_1^2, x_1^3, x_1^4)$	1	$(a_1^1, a_1^2, a_1^3, a_1^4)$
2	$(x_1^1, x_1^2, x_1^3, x_2^4)$	2	$(a_1^1, a_1^2, a_1^3, a_2^4)$
…		…	…
22	$(x_1^1, x_3^2, x_2^3, x_1^4)$	22	$(a_1^1, a_3^2, a_2^3, a_1^4)$
…		…	…
27	$(x_1^1, x_3^2, x_3^3, x_3^4)$	27	$(a_1^1, a_3^2, a_3^3, a_3^4)$
28	$(x_2^1, x_1^2, x_1^3, x_1^4)$	28	$(a_2^1, a_1^2, a_1^3, a_1^4)$
29	$(x_2^1, x_1^2, x_1^3, x_2^4)$	29	$(a_2^1, a_1^2, a_1^3, a_2^4)$
…		…	…
49	$(x_2^1, x_3^2, x_2^3, x_1^4)$	49	$(a_2^1, a_3^2, a_2^3, a_1^4)$
…		…	…
54	$(x_2^1, x_3^2, x_3^3, x_3^4)$	54	$(a_2^1, a_3^2, a_3^3, a_3^4)$
55	$(x_3^1, x_1^2, x_1^3, x_1^4)$	55	$(a_3^1, a_1^2, a_1^3, a_1^4)$
56	$(x_3^1, x_1^2, x_1^3, x_2^4)$	56	$(a_3^1, a_1^2, a_1^3, a_2^4)$
…		…	…
76	$(x_3^1, x_3^2, x_2^3, x_1^4)$	76	$(a_3^1, a_3^2, a_2^3, a_1^4)$
…		…	…
81	$(x_3^1, x_3^2, x_3^3, x_3^4)$		$(a_3^1, a_3^2, a_3^3, a_3^4)$

The calculated values of the primary transition probabilities and the primary revenues are input for the software designed to calculate the joint transition probabilities and the associated revenues for the four-dimensional MDP. Because of the relatively small number of states and actions in the example and the relatively short time of finding the optimal solution, the discounted policy iteration method to solve the MDP is chosen. It gives the optimal decision policy and the respective state-value functions for every state for the optimal policy, i.e. the average expected returns for every state. The optimal policy, determines the associate matrix of transition probabilities P and the matrix of revenues R. The vector $X = (x_1, x_2, \dots, x_{81})^T$ of the long-run stationary probabilities for the optimal policy is determined by solving the system of linear equations obtained from the matrix equation $P \cdot X = X$ and the equation $x_1 + x_2 + \dots + x_{81} = 1$. Further vectors $v = diag(P \cdot R^T)$ and $E = X^T \cdot v$ are calculated. The value of E is the expected revenue of the optimal policy per transition step and later it reflects the improvement of the condition of the system [14].

5 REAL APPLICATION AND RESULTS

The mathematical model, developed as part of the methodology for generating optimal decision policies and success factors, was applied to a specific enterprise within the printing industry to address a real-world quality-accuracy management issue focused on minimizing scrap and reducing associated costs. According to their PMS, the sub-key element of success quality-accuracy is located as critical. The opinion of the experts from the company was that the importance of the sub-key element quality-accuracy in this company is 0.7. The performance of this element was measured by the indicator percent of scrap. The average value of percent of scrap for the selected sample was 13.59%, which was grade 4 as per the scale for the performance-axis in the I/P (Importance/Performance) matrix. The developed mathematical model proposed optimal or suboptimal decision policy to improve the performance of the located critical sub-key element.

The printing machine is linked to the design studio through specialized software and includes functionality to record total scrap production. However, it does not categorize scrap by individual causes. Instead, operators manually document these causes in detailed daily reports, which are subsequently processed into specific forms to track scrap sources. This system provides a clear method for identifying the root causes of scrap, but as discussed earlier, this can be enabled with the new IT development. A portion of these records was made available for this research. The collected data for scrap production referred to 396 consecutive orders

(printing machine runs), possible causes for scrap, appropriate corrective actions and associate costs. This information was instrumental in modeling the real problem. Considering the expert's opinion and the analysis of two measures, percent of scrap and number of scrap sheets, it is concluded that it is more appropriate to use the measure (the indicator) percent of scrap and the average of percents of scrap in defining the states of MDPs. Real data of available corrective actions were used to define action spaces for MDPs. Clearly, the available sets of states and corrective actions are too big and complex and the model needed to simplify the state and action sets in order to avoid the problem of dimensionality.

According to the opinion of the experts from the company and the collected real data for 396 consequent series with different number of sheets printed on the monitored machine, it is decided that it's appropriate to consider three intervals of percents of scrap for the scrap cause – the machine. This defined the states for the random process "condition of machine" based on percent of scrap it caused for the considered sample. The limits are determined according to the average of percents of scrap for all series. The percents of scrap caused by the machine for the monitored sample were provided from real records. The operator as factor of influence is most difficult to evaluate, because of its complex and unpredictable behavior in different situations. Therefore, the systems in which the influence of the operator is greater are more difficult for modeling and analyzing. The modeling and the analysis in this paper mainly rely on collected real data. The condition of the operator is defined in terms of the induced percent of scrap, for which the cause is the operator. The condition of the tools is defined in terms of the induced percent of scrap, for which the cause are the tools. The condition of the materials is defined in terms of the induced percent of scrap, for which the cause are the materials. Intervals of percents of scrap to define the states of all four causes are determined based on real data, the average of percents of scrap of the considered sample for this cause and experts' opinion.

Based on corrective actions taken in reality, three types of actions were selected for each cause. To make the model more realistic in the future, the spaces of states and actions should be more detailed and more comprehensive, but large size problems require more complex algorithms and software for solving. However, the mathematical model is open in this sense and has opportunity to explore larger issues.

The collected real data were used to calculate the transition probabilities with the formula $p_{ij}^{k}(l) = \frac{N_{ij}^{k}(l)}{N_{i}^{k}(l)}$, where $N_i^k(l)$ denotes the number of times the one-dimensional MDP described by the random process l was in state with index i, under the influence of primary action with index k, and $N_{ij}^k(l)$ denotes the number of times it made transition from state with index i to state with index j, under the influence of primary action with index k. The notations of the primary states and actions are simplified identifying them with their indexes. For the given example, $i \in \{1,2,3\}$, $j \in \{1,2,3\}$, $k \in \{1,2,3\}$, $l \in \{1,2,3,4\}$.

Real data give information of all transitions from one state to another under the influence of certain action. Thus, the transition probabilities were determined [14].

The enterprise uses special price lists and software to calculate the price for an order. These information were used to calculate the elements of costs matrices. The revenue matrices that follow the percent of scrap from the collected real data of percent of scrap divided by cause were obtained. They were calculated separately for each cause. First the differences between the percents of scrap for all consecutive transitions from one state to another under the influence of some action were calculated. Then they were divided in 18 columns representing the feasible transitions. Those representing the same transition under the same action were put in the same column. After that, averages for each column were calculated, for the considered sample, and for every cause the average percent of scrap's increase or decrease for every transition under the influence of every action was obtained. These values were denoted by $r_{ij}^{k}(l)$, $i,j,k \in \{1,2,3\}$, $l = 1,2,3,4$. Clearly, for a larger sample, more accurate results should be obtained. Thus, the percent of scrap is followed individually for each cause and totally for the whole system. Also, the optimal decision policy was obtained, minimizing the percent of scrap. For all unfeasible transitions, for $l = 1,2,3,4$, $r_{21}^{1}(l) = r_{31}^{1}(l) = r_{32}^{1}(l) = r_{12}^{2}(l) = r_{13}^{2}(l) = r_{23}^{2}(l) = r_{12}^{3}(l) = r_{13}^{3}(l) = r_{23}^{3}(l) = 0$.

For the considered model the output from the designed software was 81 transition probabilities matrices with dimensions 81x81 and 81 revenue matrices with the same dimensions. Applying policy iteration method, for different discount rates, different optimal policies were obtained represented in the form of two vector columns with dimensions 81x1, which means that for every state of the system (numerated from 1 to 81 in Table 3), optimal policy suggests the corrective action numerated in Table 4. The other output vector column consists of state-

value functions i.e. the average expected return for every state. Optimal policy determines transition probability matrix P and revenue matrix R. Vector $X = (x_1, x_2, \dots, x_{81})^T$ of the stationary probabilities is solution to the system of linear equations obtained from $P \cdot X = X$ and $x_1 + x_2 + \dots + x_{81} = 1$. Further, vectors $v = (v_1, v_2, \dots, v_{81})^T$, $v = diag(P \cdot R^T)$ are determined and $E = X^T \cdot v$ is obtained. The obtained value is $E = 5.2107\%$ which is the expected percent of scrap for the optimal policy per transition step and grade 8. The results for discount rate 0.99 are given in Table 2. The optimal decision policy can be compared to any other decision policy, or any two stationary policies can be compared. Figure 1 shows the revised I/P matrix.

Table 2: Results from the optimization of the percent of scrap for discount factor 0.99.

State i	Optimal policy	V *1.0e+03	x_i	v_i
1	81	1.2577	0.0164	-36.5635
2	81	1.2580	0.0116	-30.7814
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
14	78	1.2601	0.0009	7.4408
15	78	1.2624	0.0019	8.2197
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
80	81	1.2831	0.0705	30.6179
81	81	1.2831	0.0705	30.6179

The average value of percent of scrap for the selected sample is 13.5928%, and with the optimal policy it decreased to 5.2107%. This improved value for the performance of the sub-key element quality-accuracy is entered in the revised I/P matrix, which shows the improvement of the condition of the system.

In order to examine the sensitivity of the model to changes in transition matrices, the performed sensitivity analysis led to the conclusion that the model is not very sensitive to such changes.

Determination of cost matrices was also needed to find the optimal decision policy minimizing costs. In these matrices costs are with negative sign, and the revenue with positive sign. For all unfeasible transitions, for $l = 1,2,3,4$, $c_{21}^1(l) = c_{31}^1(l) = c_{32}^1(l) = c_{12}^2(l) = c_{13}^2(l) = c_{23}^2(l) = c_{12}^3(l) = c_{13}^3(l) = c_{23}^3(l) = 0$. Other costs are computed with the $c_{ij}^k(l) = -t_l^k + \frac{r_{ij}^k(l) \cdot 1470 \cdot 12.7}{100}$, $i, j \in \{1,2,3\}$, $l = 1,2,3,4$, $k \in \{2,3\}$, and for k=1, $c_{ij}^1(l) = -\frac{r_{ij}^1(l) \cdot 1470 \cdot 12.7}{100}$, where 1470 is the average of the sum of the order size and the number of scrap sheets for the considered sample. 12.7 is the average price of one finished sheet calculated according the methodology of the enterprise. The research included analysis of costs t_l^k for the corrective actions for each cause, which gives Table 3, illustrated with matrix.

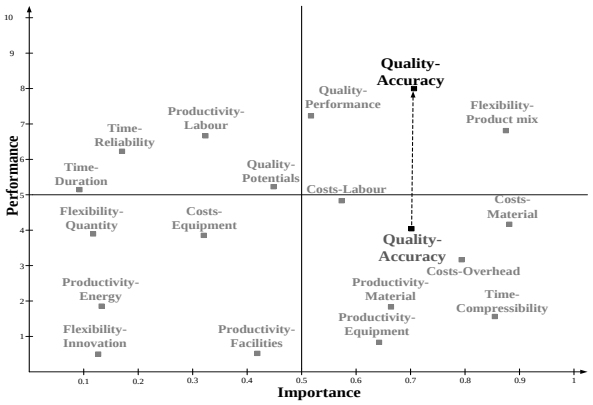


Figure 1: Revised I/P matrix.

Table 3: Costs for the corrective actions for each cause.

t_l^k	Action 1	Action 2	Action 3
Machine	0	250	1000
Operator	0	150	230
Tool	0	200	1473
Material	0	300	1395

The calculated average costs induced with the transitions between the states, under the influence of the primary actions for all the MDPs, were calculated and presented with cost matrices and joined table, similarly as for the transition probabilities in [14].

Table 4: Results from the optimization of the scrap costs for discount factor 0.99.

State i	Optimal policy	V *1.0e+05	x_i	v_i *1.0e+03
1	49	1.3159	0.0527	-4.7413
2	50	1.3151	0.0080	-4.7913
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
14	50	1.3200	0.0017	0.8922
15	50	1.3233	0.0010	-0.1339
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
80	50	1.3558	0.0031	-0.3511
81	50	1.3558	0.0031	-0.3511

Same methodology as for determining the optimal decision policy for minimization of percent of scrap is used to get the optimal decision policy for minimization of costs of scrap and the computed value was $E = -546.0880$ which is the expected cost for scrap for the optimal policy, per transition step. Table 4 shows the results from the calculations with the respective currency.

To compare, in the same way the expected costs per transition step are calculated, for the optimal policy obtained previously minimizing the percent of scrap, and the value is $E = -1189$.

6 EMERGING CHALLENGES AND FUTURE PROSPECTS

Markov Decision Processes (MDPs) are highly suitable for automating manufacturing processes due to their ability to model decision-making in complex and uncertain environments, where current actions influence future states and outcomes. Automated quality control systems benefit from MDPs as they provide a framework for optimizing inspection frequencies and methods, thereby ensuring consistent product quality. Key advantages of incorporating MDPs in manufacturing automation include data-driven decision-making, adaptability to changing conditions, operational efficiency and cost savings, and enhanced predictive capabilities. Within automated quality control systems, various MDP models are particularly effective in managing uncertainty and adapting policies based on quality metrics and outcomes.

However, the successful application of MDPs, especially in automated quality control systems, is subject to several challenges and risks that can limit their effectiveness in real-world industrial settings. Key challenges include inaccurate transition probabilities, oversimplified system assumptions, the need for a stationary environment, data quality and availability issues, the exploration versus exploitation trade-off, model overfitting, computational complexity, and human factors. To address these challenges, industries must prioritize robust model validation and verification, continuously update transition probabilities with real-time data, and incorporate adaptive mechanisms to respond to evolving conditions. Regular testing and refinement of models based on actual operational data are essential for maintaining the accuracy and efficiency of automated quality control systems and reducing scrap in manufacturing.

To further enhance flexibility and adaptability, MDPs can be extended through vector random processes, enabling the simultaneous management of multiple quality and process metrics within a single system. In some models, independent state variables—representing different subsystems such as machines, operators, tools, and materials—simplify the MDP structure by allowing separate consideration of each subsystem's dynamics and transitions. This approach enables a multi-dimensional MDP framework in which each subsystem functions as a one-dimensional MDP, focusing on minimizing scrap by addressing significant contributors individually. Monitoring each subsystem's state facilitates targeted, subsystem-specific interventions, optimizing quality control in manufacturing environments.

Reinforcement Learning (RL)-enhanced MDPs and hierarchical MDPs offer additional adaptability in multi-dimensional contexts, allowing coordinated control across subsystems. These models are particularly promising for reducing scrap in complex manufacturing setups, as they learn and refine optimal policies over time. Nonetheless, determining accurate transition probabilities from historical data remains a challenge and requires ongoing attention to data integrity and validation.

In MDP-based manufacturing quality control models, subsystems are treated as stochastic processes governed by different probability distributions, which capture the uncertainty and variability in state transitions. These distributions form the basis for modeling scenarios such as machine failure, tool wear, material defects, and operator errors. By incorporating these distributions into MDP frameworks, manufacturers can make informed decisions on process adjustments to minimize defects and improve overall accuracy. Specialized software tools like MATLAB/Simulink, AnyLogic, Arena, Python, @Risk, R, and Simio facilitate the simulation and optimization of MDP models, allowing tailored probability distributions to enhance decision-making and production efficiency.

To further enhance quality control, integrated technologies such as sensors, wearable devices, predictive maintenance systems, and data analytics platforms can be used for failure detection and responsibility assignment within subsystems. These tools, when combined within unified systems like Manufacturing Execution Systems (MES), enable precise tracking of defects and their causes, enhancing accountability and proactive mitigation efforts. Emerging technologies, including machine learning (ML), artificial intelligence (AI), predictive analytics, IoT, and cloud-based solutions, are shaping the future of automated quality control by enabling

more accurate, proactive defect detection and process optimization.

7 CONCLUSIONS

This study introduces a comprehensive methodology for integrating MDPs into industrial quality control systems, emphasizing automation and decision-making under uncertainty. By developing a four-dimensional MDP model within a PMS framework, the research effectively minimizes scrap and associated costs. Application in a real-world printing enterprise validated the model's practical value, achieving significant scrap reductions and cost savings while providing a structured approach to addressing inefficiencies.

Key contributions include methodologies for transition probability estimation and revenue matrices, alongside policy iteration optimization to generate actionable decisions. The study highlights the adaptability and scalability of MDPs, accommodating additional factors and larger datasets to tackle complex manufacturing challenges.

Integrating advanced IT tools, such as real-time monitoring and data analytics, enhances model accuracy and responsiveness, improving efficiency, product quality, and sustainability. Future work should expand state-action spaces, leverage machine learning and IoT, and address data challenges to advance automation and industrial competitiveness.

ACKNOWLEDGEMENT

We acknowledge support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Open Access Publishing Fund of Anhalt University of Applied Sciences.

REFERENCES

[1] R. Minovski, D. Jovanoski, and K.-P. Zeh, "Ein universelles Modell zur Unternehmensstrukturierung," Compass, REFA – Nachrichten, Heft 5, pp. 53-58, Oct. 2002.

[2] M. Puterman, "Markov Decision Processes: Discrete Stochastic Dynamic Programming," 1st ed., John Wiley & Sons Inc., 1994.

[3] A. Neely, "The performance measurement revolution: why now and what next?" Int. J. Oper. Prod. Manag., 1999.

[4] R. S. Kaplan and D. P. Norton, "The Balanced Scorecard—Measures That Drive Performance," Harv. Bus. Rev., 1992.

[5] W. E. Deming, "Out of the Crisis," MIT Press, 1986.

[6] J. M. Juran, "Juran on Planning for Quality," Free Press, 1988.

[7] U. S. Bititci, P. Garengo, V. Dörfler, and S. S. Nudurupati, "Performance measurement: Challenges for tomorrow," Int. J. Manag. Rev., 2012.

[8] B. Marr, "Key Performance Indicators: The 75+ Measures Every Manager Needs to Know," Pearson UK, 2015.

[9] A. A. Markov, "Extension of the Limit Theorems of Probability Theory to a Sum of Variables Connected in a Chain," 1906.

[10] R. Bellman, "Dynamic Programming," Princeton Univ. Press, 1957.

[11] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction," MIT Press, 1998.

[12] I. Janicevic, J. Filipovic, and J. Miscevic, "Using a Markov Chain for Product Quality Improvement Simulation," UPB Sci. Bull., Ser. D: Mech. Eng., vol. 76, no. 1, pp. 227-242, 2014.

[13] W. B. Powell, "Approximate Dynamic Programming: Solving the Curses of Dimensionality," Wiley-Interscience, 2007.

[14] K. Mitkovska-Trendova, R. Minovski, and D. Boshkovski, "Methodology for transition probabilities determination in a Markov decision processes model for quality-accuracy management," J. Eng. Manag. Competitiveness (JEMC), vol. 4, no. 2, pp. 59-67, 2014.

[15] E. Tolstaya, A. Koppel, E. Stump, and A. Ribeiro, "Nonparametric Stochastic Compositional Gradient Descent for Q-Learning in Continuous Markov Decision Problems," ACC, pp. 6608-6615, 2018.

[16] R. J. Boucherie and N. M. van Dijk, "Markov Decision Processes in Practice," Springer Int. Publ., 2017.

[17] Y. Yang, J. Wu, X. Song, D. Wu, L. Su, and L. Tang, "A Novel Black Box Process Quality Optimization Approach based on Hit Rate," arXiv preprint arXiv:2305.20003, 2023.

[18] C. Reisinger and J. Tam, "Markov Decision Processes with Observation Costs: Framework and Computation with a Penalty Scheme," arXiv preprint arXiv:2201.07908, 2023.

[19] A. Nasir, S. Mekid, Z. Sawlan, and O. Alsawafy, "Optimized Task Assignment and Predictive Maintenance for Industrial Machines using Markov Decision Process," arXiv preprint arXiv:2402.00042, 2024.

[20] S. Chen, D. Simchi-Levi, and C. Wang, "Experimenting on Markov Decision Processes with Local Treatments," arXiv preprint arXiv:2407.19618, 2024.

SECTION 4

Power Engineering

Proceedings of the 12th International Conference on Applied Innovations in IT (ICAIIT)

NOVEMBER 2024

Mathematical Modeling of Heat Parameters of Photothermal Device

Boysori Yuldoshov, Mamasobir Tursunov, Gayrat Raimov, Shakhvoz Karshiev and Sirojiddin Toshpulatov
Termez State University, Barkamol Avlod Str. 43, 190111 Termez, Uzbekistan
b.yuldoshov10@mail.ru, mtursunov@tersu.uz, raimov.gayrat@mail.ru, shqarshiyev@tersu.uz, toshpulatovs@tersu.uz

Keywords: Solar Cell, Photovoltaic Panel, Photothermal Device, Heat Collector, Polycarbonate, Solar Radiation, Temperature.

Abstract: The article investigates hot water production by cooling a photovoltaic panel in a PV device. For this purpose, a parallel-channel polycarbonate heat collector was installed on the back surface of the PVT. The results obtained based on mathematical modelling of PVT thermal parameters are presented in the article. Mathematical modelling was performed based on Comsol multiphysics 6.1 integrated software. In mathematical modelling, the dependence of thermal energy obtained from PVT on radiation intensity, temperature and speed of cooling water was studied. Mathematical modeling results were compared with experimental results. According to the results of mathematical modelling, the maximum value of PV surface temperature was 47.6ºC in summer when the water speed was 0.001m/s. In the summer season, the results of the experiment conducted on the PVT device and mathematical modelling were compared. The results of experiments and mathematical modelling showed that the temperature of hot water difference by 2ºC. It was determined that 41.4 liters of hot water per hour is obtained from the PVT device.

1 INTRODUCTION

Solar energy is the largest and most environmentally friendly source of renewable energy. The amount of solar energy reaching the Earth is about 85PW. This is 500 times more than the world's power consumption (15TW) [1]. Solar energy can be converted into electrical and thermal energy using PV [2]. During the use of PV, its effectiveness decreases with the increase in temperature [3]. This process is related to the part of the radiation that is not absorbed by the solar elements and turns into heat. Various PVT models have been developed by researchers in recent years [4, 5, 6, 7]. Such devices differ in heat-absorbing (absorber) material, heat-carrying substance, and other physical parameters [8]. In [9], a complete description of the air and water configurations of the PVT system is presented. In [10], the electrical and thermal parameters of the PVT system, which consists of rectangular copper pipes on the back side of the polycrystalline PV, were described. In article [11], a flat water heater combined system with PVT is studied.

In [12], analytical expressions for energy balance equations and thermal parameters for various components of PVT were obtained. In some works, it is possible to see hydrodynamic equations in PVT [13, 14, 15, 16], and analyses obtained by mathematical modeling of electrical and thermal parameters [17, 18].

Since solar energy is a more sustainable source of energy than other renewable sources, it is important to follow research in this direction. In this article, the change in thermal parameters of the PVT system is analyzed using mathematical modelling.

2 METHODS AND MATERIALS

The temperature of the hot water obtained from the PVT mainly depends on the intensity of solar radiation and the surface of the area where the radiation falls. Therefore, reflectors installed in PVT also serve to increase thermal energy.

Figure 1 shows the solar radiation falling on the PV surface directly and through reflectors and the process of water heating in the PV heat collector.

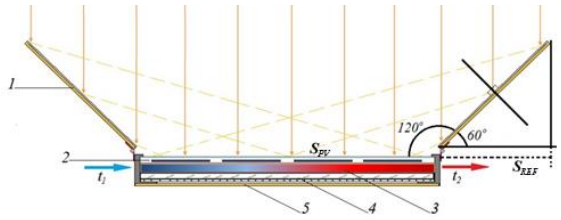


Figure 1: PVT reflectors are installed at an optimal angle. 1 – reflector, 2 – PV, 3 – heat collector, 4 – heat retaining materials, 5 – back cover.

The variation of PV and water temperatures at different velocities of water flowing through the channels of the heat collector of the PVT device was simulated using the integrated software Comsol multiphysics 6.1. In order to save computer RAM and reduce calculation time, mathematical modeling was performed for a part of the width of one solar cell in the direction of water flow channels (Fig. 2).

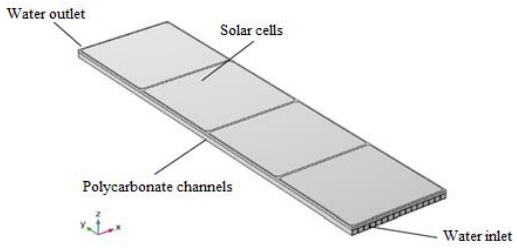


Figure 2: Part of PVT extracted for mathematical modelling.

In the modeling, hydrodynamic equations were solved to determine the dynamic quantities of flowing water in the channels. In the considered device, the speed of water flow varies in the range of 0.001-0.004 m/s. In this case, the following kinematic viscosity (1) is used for the Reynolds number (Re):

$$Re = \frac{V \cdot d}{v} \qquad (1)$$

herein, V – flow rating speed (m^3/s), d – channel size (m), v – kinematic viscosity (m^2/s). At a temperature of 40°C, the kinematic viscosity is equal to $0{,}659 \cdot 10^{-6} m^2/s$, will be less than 2300 at our chosen speed values.

At such small values of the Reynolds number, the water in the channels consists of a laminar flow. For this reason, we choose the "Laminar Flow" interface for the water flow in modeling.

The "Heat Transfer in Solids and Fluids" and "Surface-to-Surface Radiation" physics interfaces were used to simulate the heat exchange between fluid, PV, and the environment. In this case, the "Heat Transfer with Surface-to-Surface Radiation" multiphysics interface was used, which combines the description of "Heat transfer" and "Surface-to-surface radiation".

To describe the heat exchange of a water flow with a solid body, the "Laminar Flow (spf)" and "Heat Transfer in Solids and Fluids (ht)" interfaces through the "Nonisothermal Flow" nonisothermal flow multiphysics interface was connected.

For the convenience of modelling, we divide the input parameters into the following two types (Tables 1-2).

Table1: Geometric and dynamic parameters of the device.

Name of variables	Size (m)	Description
Wg	0.16	Width of PV
Lg	0.16	The length of the solar cell
H	0.00471	The thickness of PV
Ws	0.156	Width and height of the solar cell
Wk	0.0095	Water channel width
Hk	0.008	The height of the water channel
n	4	Solar element number
L	0.64 m	The length of PV
T_0	298.15 K	Initial water temperature
u_in	0.001 m/s	The speed of water entering the channels
Q_{water}	$1.28 \cdot 10^{-6}$ (m^3/s)	The amount of water coming out of the PVT in one second
N	16	Number of channels
M_{water}	4.608 kg	Mass of water heated in 1 hour

Table 2: Environment parameters.

Name of variables	Size	Description
T_{avg}	27 °C	Ambient temperature at start time
dT	3 °C	Ambient temperature gradient
Day	1	The day of the experiment
Month	7	The month of the experiment
Year	2021	The year of the experiment

Data from the ASHRAE Meteorological Database were used to account for changes in ambient temperature and solar activity. On July 1, 2021, the weather conditions in the city of Termez were taken from 10:00 a.m. to 4:00 p.m. The number of iterations for parametric research was 128, and the calculation time was 3 hours 31 minutes 23 seconds.

In the program, the following analytical expression was created to determine the temperature change in the movement of the earth about the sun during the day (2):

$$T_{amb} = T_{avg} + dT \cdot cos\left(\frac{x-14}{12} \cdot \pi\right), \quad (2)$$

here, Tamb is the ambient temperature (°C), Tavg – average temperature (°C), x – the time being viewed. The graph of this expression is presented in Figure 3.

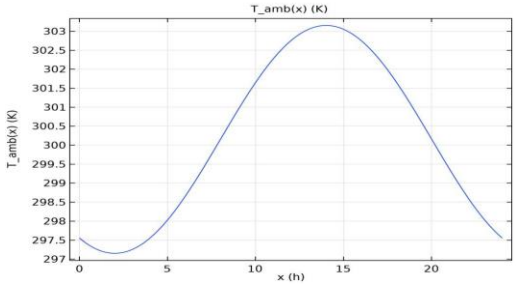


Figure 3: Changes in ambient temperature during the day.

Modeling was carried out in three-dimensional space. The model was implemented in an automated system for planar shear layers and turbulent flows. This model includes the least set of equations in the case where the unknown quantities are minimized for all-natural heat and flow processes [19, 20].

In the modelling, the blog of the nonstationary Navye-Stokes equation through nonlocal couplings was used in the Comsol multiphysics 6.1 platform and solved together with the continuity (3), (4):

$$\rho\frac{\partial u}{\partial t} + \rho(u\cdot\nabla)u = \nabla\cdot\left[-pI + K\right] + F \quad (3)$$

$$\rho\nabla\cdot u = 0 \quad (4)$$

therein, ρ – water density (kg/m^3), u – flow speed (m/s), I – impulse (kg•m/s), K – kinetic energy (J), F – external forces (N).

Boundary conditions were set as follows:

- the liquid does not slip on the walls of the channel, that is u=0, the flow speed is zero;
- the velocity at the entrance to the channels is directed along the normal of the entrance surface;
- the pressure difference at the exit from the channel is zero.

The developed equations for heat transfer in PV layers were used in the "Heat Transfer in Solids and Fluids" interface. The following heat balance equation and heat transfer (6) between layers were solved for solids:

$$\rho c_p\frac{\partial T}{\partial t} + \rho c_p u\nabla T + \nabla q = \sum Q_i \quad (5)$$

$$q = -k\nabla T \quad (6)$$

therein, u – flow speed of water (m/s), c_p – heat capacity (J/(kg•°C)), q – heat flux density (W/m^2), k – heat transfer coefficient (W/(m^2•°C)).

The following heat balance (7), (8) was solved for heat transfer between fluid and solid:

$$\rho c_p\frac{\partial u}{\partial t} + \rho c_p u\nabla T + \nabla q = Q + Q_p + Q_{vd} \quad (7)$$

$$\rho = \frac{p_A}{R_s T} \quad (8)$$

therein, Q – amount of heat (J), Q_p – spot heat quantity (J), Q_{vd} – Viscous dissipation (J), p_A – absolute pressure (Pa), R_s – Universal gas constant (J/kg•K). The effect of temperature on the density and viscosity of water was taken into account. It was considered that there is no heat exchange with the external environment (9):

$$-nu = 0 \quad (9)$$

n – normal to the wall. The following expression was used for the heat transfer to the external environment for the PV surface (10), (11):

$$-n\cdot q = q_0 \quad (10)$$

$$q_0 = k(T_{ext} - T) \quad (11)$$

therein, q – heat flux density (W/m^2), q_0 – internal heat flow (W/m^2), k – heat transfer coefficient W/(m^2•K), T_{ext} – exterior temperature (K).

The following quantities are available in the Comsol multiphysics 6.1 platform for the effects of incident radiation:

- solar radiation intensity;
- energy balance;
- Stefan-Bolsmann equation.

Functions indicating the ambient temperature were calculated as a result of the sum of radiation for all wavelengths, taking into account the geographical width of the selected area. The (12) - (16) were solved using the finite element method.

$$J_i = \varepsilon_i e_b(T) FEP_i(T) + \rho_{dj} G_i \quad (12)$$

$$G_j = G_{m,j} + G_{amb,j} + G_{ext,j} \quad (13)$$

$$G_{ambj} = F_{ambj}\varepsilon_{amb} e_b(T_{amb}) FEP_j(T_{amb}) \quad (14)$$

$$e_b(T) = n^2\sigma T^4 \quad (15)$$

$$FEP_j(T) = \frac{15}{\pi^4}\int_{c_2/(\lambda_{j-1}T)}^{c_2/(\lambda_j T)} \frac{x^3}{1-e^x}dx \quad (16)$$

therein, J_i – Intensity of radiation falling on the PV surface (W/m^2), ε_i – surface irradiation capacity, e_b – grayness of the surface, FEP_i – seasonal radiation conditions of the area, ρ_{dj} – coefficient of reflection, G_i – surface radiation intensity (W/m^2), G_{mj} – mutual surface radiation (W/m^2), $G_{amb,j}$ – environmental radiation (W/m^2), G_{ext} – external radiation (W/m^2), $F_{amb,j}$ – environmental factor, ε_i – environmental emissivity, n – light refraction indicator, c_2 – heat capacity (J/(kg•°C)), λ_j – thermal conductivity (W/m^2).

3 RESULTS AND DISCUSSION

The number of solar elements and the channel length were considered in mathematical modelling. Determined experimentally 0.001m/s value was given for the velocity of water entering the cooling system channels. Primarily, the surface heating of a 640mm long solar panel piece consisting of 4 elements was studied at four different time points (Fig. 4).

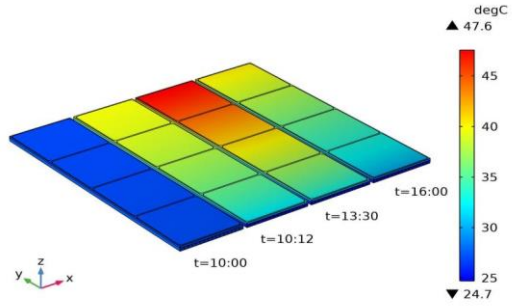


Figure 4: Change of PV surface temperature when water speed u_in(1)=0.001 m/s.

During periods of perpendicular solar radiation, the water outlet temperature of PVT reaches its maximum value of 47ºC. In the rest of the time, the outgoing water temperature is 41-42ºC. The velocity of the water inside the channels is small and the thickness of the polycarbonate wall is 0.5mm, so the temperature of the water is close to the temperature of the PV. It differs by 1-1.5ºC at most.

Figure 5 shows the change in water temperature in the channels of the PVT heat collector during the day. It can be seen from the picture that the difference between the temperature of the PV surface and the temperature of the water shown in Figure 5 was equal to 1.1^{0}C. This situation is very close to the experimental result, with a difference of 1.5-2ºC due to the presence of external temperature and wind speed.

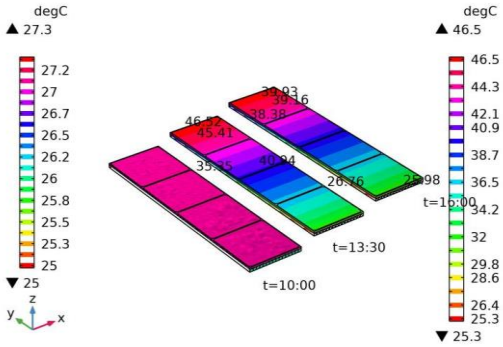


Figure 5: Change of water temperature at a water speed u_in(1)=0.001 m/s.

Figure 6 shows the change in water temperature at different water speeds under the same conditions. It can be seen from the figure that when the water speed is 0.001 m/s, the PVT surface temperature is 47ºC, while the exit water temperature is 45ºC.

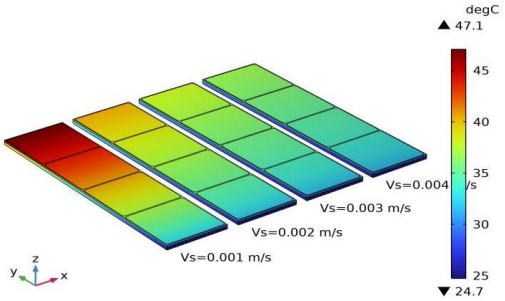


Figure 6: Change in water temperature at different water speeds under the same conditions.

As a result of changes in water density, viscosity and pressure due to the temperature gradient along the length of the channels, an increase in water flow speed is observed. The speed of water in the channels increases along the length of the PV and reaches 0.00199 m/s almost twice due to the loss of the pressure difference at the outlet (Fig. 7-8).

Figure 9 shows the absorption of radiation in the heat collector through the PV surface. It shows the amount of radiation converted into heat during the day in one line of PV when the water flow speed is 0.003 m/s.

Figure 10 shows the absorption of solar radiation at different water flow rates in a PVT heat collector under the same conditions. In the graph, the result of the above opinion was approved.

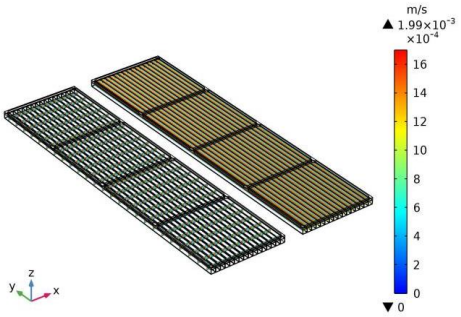


Figure 7: Increase in water speed in channels, t=13:30.

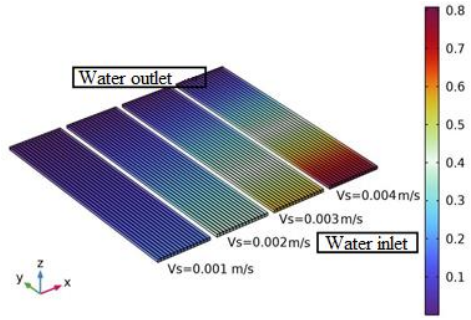


Figure 8: Change of pressure in channels at different speeds, t=13:30.

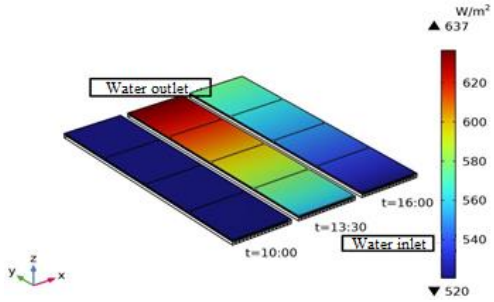


Figure 9: Absorption of solar radiation in PVT, (W/m^2).

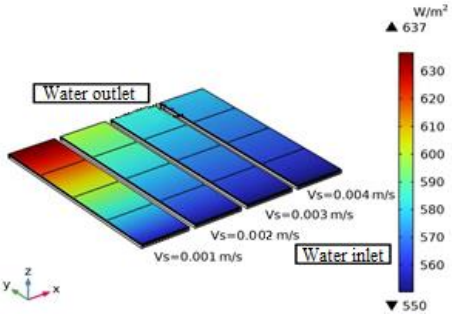


Figure 10: Absorption of solar radiation in FIB at different water speeds, (W/m^2).

The thickness of the normal to the surface of the device consisting of PV and cooling system is 14.71 mm. The walls of the polycarbonate channel are 0.5 mm thick. The width and the height of the channel are 5 mm and 8mm respectively. Figure 11 shows the temperature change in polycarbonate, water, polycarbonate and solar elements as a result of the heat exchange process in the layers due to the absorption of the radiation falling on the PV surface. At a water speed of 0.001 m/s, the temperatures ranged from 31°C to 34.4°C in QE1, 32°C to 35°C in QE2, 33.3°C to 37.4°C in QE3, and 37.2°C to 41.2°C in QE4 can be seen to change.

Figure 12 shows the variation of the PV surface temperature along the length of the channel during the day. Under the same conditions, the temperature of the solar cell will be from 40^{o}C to 47.5^{o}C at values of water flow velocity of 0.001-0.002 m/s. The arrangement of four solar cells along the length of the channel can be called the optimal option.

Figure 13 can be used to determine the optimal value of the water flow rate for the use of the cooling system as a heat collector. It can be seen from the picture that for use as a heat collector, the flow rate in the channels of 0.001m/s gives the possibility of receiving water at a temperature of 39^{o}C-40^{o}C.

4 CONCLUSIONS

The mathematical modeling of thermal parameters in photovoltaic-thermal (PV/T) devices plays a key role in understanding and improving the processes within combined systems. For this purpose, modeling of thermal processes in a PVT device which has a polycarbonate heat collector was carried out. When the speed of water in the PVT channel is 0.001m/s, the temperature of the PV surface is 47.6°C, and the temperature of the outgoing water is 46.5°C. When the reflector is installed at the optimal angle, the temperature of the PV surface is 61.2°C, and the temperature of the outgoing water is 58°C. Also, the temperature change in the first solar cell (QE1) when cold water passes through the heat collector is 31^{o}C - 34.4^{o}C, in QE2 it is 32^{o}C - 35^{o}C, in QE3 it is 33.3^{o}C - 37.4^{o}C and in QE4 it is 37.2^{o}C - 41.2^{o}C, determined. These results differed by 2^{0}C from the experimental result. It was found that 41.4 litres of hot water per hour was taken from the device. By increasing the incoming water consumption, electricity losses are reduced, however, the decreasing outgoing water temperature has been proven based on experimental and modeling results.

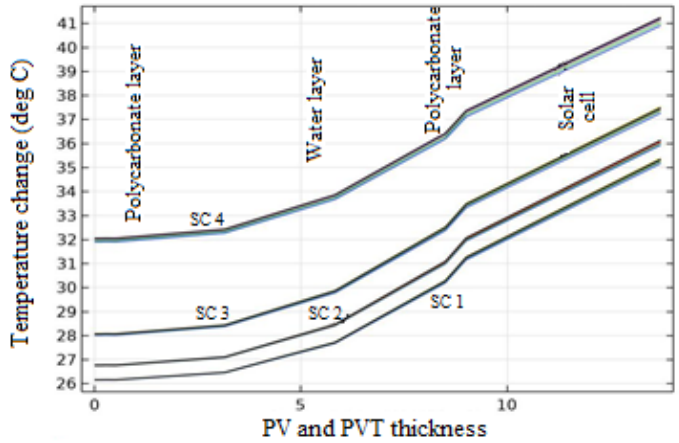


Figure 11: Variation of PV temperature during the day.

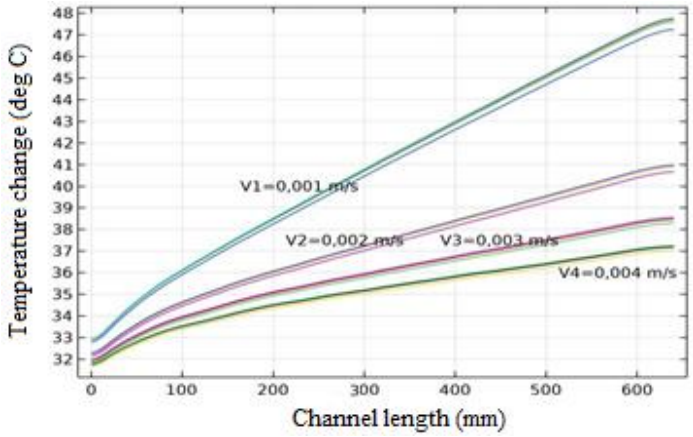


Figure 12: Variation of temperature on the surface of PV during the day.

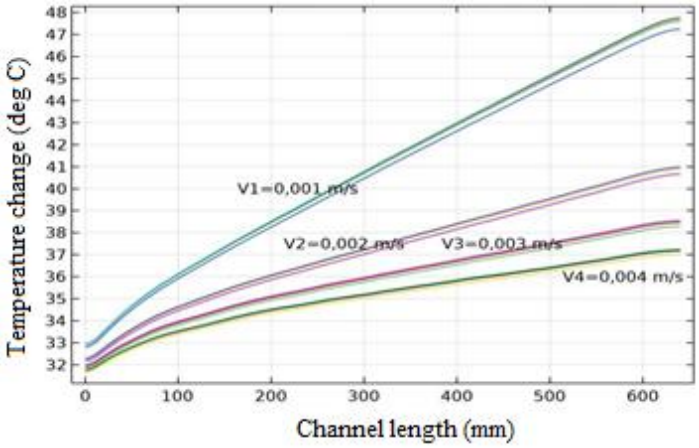


Figure 13: Change in water temperature in the channel.

REFERENCES

[1] M. Khamooshi, H. Salati, F. Egelioglu, A. Hooshyar Faghiri, J. Tarabishi, and S. Babadi, "A review of solar photovoltaic concentrators," Int. J. Photoenergy, vol. 2014, pp. 1-17.

[2] M. Tursunov, K. Sabirov, T. Axtamov, U. Abdiyev, B. Yuldoshov, J. Khaliyarov, S. Bobomuratov, and S. Toshpulatov, "Analysis of electric and thermal efficiency of crystal silicon small power suppliers," in Proc. Int. Conf. Appl. Innov. IT, vol. 11, no. 2, 2023, pp. 167-171.

[3] R. Muminov, M. Tursunov, Kh. Sabirov, Sh. Abilfayziyev, B. Yuldoshov, and S. Toshpulatov, "Testing of crystalline silicon-based photoelectric and photothermal batteries in real climate conditions and comparison of parameter changes," J. Phys.: Conf. Ser., vol. 2388, 2022, 012128. doi: 10.1088/1742-6596/2388/1/012128.

[4] W. He, J. Zhou, C. Chen, and J. Ji, "Experimental study and performance analysis of a thermoelectric cooling and heating system driven by a photovoltaic/thermal system in summer and winter operation modes," Energy Convers. Manag., vol. 84, 2014, pp. 41-49.

[5] A. Monjezi, Y. Chen, R. Vepa, A. Kashyout, G. Hassan, H. Fath, and M. Shaheed, "Development of an off-grid solar energy powered reverse osmosis desalination system for continuous production of freshwater with integrated photovoltaic thermal (PVT) cooling," Desalination, vol. 495, 2020, pp. 1-7.

[6] H. Nasef, S. Nada, and H. Hassan, "Integrative passive and active cooling system using PCM and nanofluid for thermal regulation of concentrated photovoltaic solar cells," Energy Convers. Manag., vol. 199, 2019, pp. 1-15.

[7] U. Sajjad, M. Amer, H. Ali, A. Dahiya, and N. Abbas, "Cost-effective cooling of photovoltaic modules to improve efficiency," Case Stud. Therm. Eng., vol. 14, 2019, pp. 1-7.

[8] S. Bhakre, P. Sawarkar, and V. Kalamkar, "Performance evaluation of PV panel surfaces exposed to hydraulic cooling – A review," Sol. Energy, vol. 224, 2021, pp. 1193-1209.

[9] R. Avezov, J. Akhatov, and N. Avezova, "A review on photovoltaic-thermal (PV-T) air and water collectors," Appl. Sol. Energy, vol. 47, no. 3, 2011, pp. 169-183.

[10] A. Alzaabi, K. Badawiyeh, O. Hantoush, and A. Hamid, "Electrical/thermal performance of hybrid PV/T system in Sharjah, UAE," Int. J. Smart Grid Clean Energy, vol. 3, no. 4, 2014, pp. 385-389.

[11] M. Arefin, "Analysis of an integrated photovoltaic thermal system by top surface natural circulation of water," Front. Energy Res., vol. 7, 2019, pp. 1-10.

[12] F. Sobhnamayan, F. Sarhaddi, M. Alavi, S. Farahat, and J. Yazdanpanahi, "Optimization of a solar photovoltaic thermal (PV/T) water collector based on exergy concept," Renew. Energy, vol. 68, 2014, pp. 356-365.

[13] F. Yazdanifard, E. Ebrahimnia-Bajestan, and M. Ameri, "Investigating the performance of a water-based photovoltaic/thermal (PV/T) collector in laminar and turbulent flow regime," Renew. Energy, vol. 99, 2016, pp. 295-306.

[14] Z. Ul Abdin and A. Rachid, "Bond graph modeling of a water-based photovoltaic thermal (PV/T) collector," Sol. Energy, vol. 220, 2021, pp. 571-577.

[15] S. Prasetyo, A. Prabowo, and Z. Arifin, "The use of a hybrid photovoltaic/thermal (PV/T) collector system as a sustainable energy-harvest instrument in urban technology," Heliyon, vol. 9, 2023, pp. 1-25.

[16] D. Luca, A. Caldarelli, E. Gaudino, E. Gennaro, M. Musto, and R. Russo, "Modeling of energy and exergy efficiencies in high vacuum flat plate photovoltaic–thermal (PV–T) collectors," Energy Rep., vol. 9, 2023, pp. 1044-1055.

[17] E. Touti, M. Masmali, M. Fterich, and H. Chouikhi, "Experimental and numerical study of the PVT design impact on the electrical and thermal performances," Case Stud. Therm. Eng., vol. 43, 2023, pp. 1027-1032.

[18] M. Attia, A. Khelifa, O. Abdulmajeed, and M. Arıcı, "Thermal analysis on the performance of a finned hybrid bi-fluid PVT system," Therm. Sci. Eng. Prog., vol. 45, 2023, pp. 1021-1035.

[19] P. Bradshaw, "Turbulent secondary flows," Annu. Rev. Fluid Mech., vol. 19, 1987, pp. 53-74.

[20] I. Jurayev, I. Yuldoshev, and Z. Jurayeva, "Results of study of photovoltaic thermal battery based on thin-film module by modeling and computational methods," in Proc. Int. Conf. Appl. Innov. IT, vol. 12, no. 1, 2024, pp. 243-249. doi: 10.25673/115707.

APPENDIX

The list of abbreviations used in this paper is given below:

PV	photovoltaic panel
PVT	photothermal panel

Testing the Energy Efficiency of a New Type of Photothermal Device in Dry Climate Conditions

Muhammadjon Tursunov[1], Khabibullo Sabirov[1], Ramazon Alikulov[1], Kholov Uygun[2], Shohimardonov Jamoliddin[2] and Eshmatov Mansur[3]
[1]*Physical-Technical Institute, Chingiz Aytmatov Str. 2B, 100084 Tashkent, Uzbekistan*
[2]*Karshi Institute of Engineering Economics, Mustakillik Avenue 225, 180100 Karshi, Uzbekistan*
[3]*Turan University, Nasaf Str. 1/14, 180100 Karshi, Uzbekistan*
muhammadtursunov54@gmail.com,sabirovhabibullo@gmail.com,ramazonaliqulov209@gmail.com, shjmadu@gmail.com, uygunshams@mail.ru

Keywords: Photoelectric Battery, Photothermal Battery, Autonomous Moving, Photothermal Device, Radiator, Water Pump, Fan, Reflectors, Accumulator Battery, Invertxer, Controller.

Abstract: The results obtained on the photoelectric battery (PVB) and photothermal battery (PTB) based on a new type of autonomous moving cooling system photothermal devices (PTD) are presented in this research work. This new type of device (PTD) is self-cooling and has the ability to provide hot water for the village residents while increasing the efficiency of the PTB. There was created a new experimental copy of the device with a power of 300 W based on the design of the PTD with a new type of cooling system with a power of 60 W mentioned in previous scientific research. It consists of a 180 W PVB, a 60A·h battery, a 2kW inverter, a 50A controller, a radiator for cooling hot water, 5 cooling fans, a pump, and a cart-shaped structure for their installation. It is possible to get results in two different situations in the experimental copy with a new type of cooling system. It is possible to increase the efficiency of PTD by fully using the cooling system, and to obtain hot water for the household without a sharp decrease in efficiency by partially using the cooling system. There is mentioned a study of hot water regimes with a temperature of 40-50 ^{0}C for the agricultural sector, depending on the intensity of solar radiation and ambient temperature. Preliminary tests showed that the power of the PTD differs from the power of the PVB by up to 70 W. This new PTD showed that it is possible to use it in many other cases, such as water supply, lighting, watching TV, listening to the radio, using a refrigerator, charging computers and phones in rural areas.

1 INTRODUCTION

It is clear to all of us that our need for electricity on a global scale is increasing day by day. If we take the last five years, the world's electricity demand is increasing by 50% every year. This requires the increase and development of alternative energy types. If we pay attention to the information of the International Energy Agency, if the use of solar energy develops at such a pace, by 2050, it will be possible to meet 25% of the world's electricity needs at the expense of solar energy, and it will be possible to reduce carbon dioxide gas released into the environment by 6 billion tons per year [1]. It is clear from this that it is necessary to switch to renewable energy sources. Solar energy is the most widely used renewable energy source. We use PVBs to convert solar energy into electricity. It has been noted in many works that the use of PVBs without taking into account climatic conditions greatly affects their efficiency [2]. Especially in conditions of high atmospheric temperature, the efficiency of PVB decreases rapidly [3]. To study the phenomenon of decreasing efficiency of PVBs made of silicon-based solar cells (SC), which convert solar energy into electricity, with increasing temperature, to analyze the process of PVB efficiency decreasing to 40% when the atmospheric temperature exceeds 50^0 C, and to prevent the phenomenon it is desirable to study the possibilities. Employees of the Institute of Physics and Technology have studied this phenomenon for 15-20 years, and in order to increase the efficiency of PVBs, they have been offered cooling devices using air or water. As a result, new constructions with 20-30% more efficiency than PVBs produced in foreign countries

were created in hot climates [4]. There was studied the supply of low-power household consumers (TV, light-emitting diode lamps, laptop, telephone) with the help of an efficient small-power autonomous portable photothermal device for dry climate without water, the use of new types of photothermal devices in extremely dry areas with water shortages, The electricity generated by this device is used in desert areas. The new type of devices that we are looking for are created for these purposes. There was created a laboratory copy of PTD with a new type of cooling system of small capacity at the Institute of Physics and Technology, and based on the results, a new PTD was created for rural and water-scarce areas with a large capacity. The difference of this new type of PTD from other photoelectric devices is that it does not require a lot of water for cooling and operation in extremely dry climates.

2 EXPERIMENTAL DEVICE AND RESULTS

A crystalline silicon-based PVB with a power of 180 W was selected for the experimental tests. A photothermal device, based on the concrete PVB, was created. There was used cellular polycarbonate with 1.5 times the geometric dimensions (width and height) of parallel channels based on polycarbonate, as presented in [5-6], in order to ensure the maximum cooling of the rear surface of the PTB. This created conditions for increasing the rate and volume of water discharge from the collector and, as a result, reducing the temperature of the PVB [7-8]. With the help of an efficient small-power autonomous portable photothermal device for a dry climate without water, funds are saved from the economic side in providing the households of rural residents with the necessary electricity and partial hot water, and in order to do this, there is used a cart with a new type of collector photothermal battery, radiator, pump, water storage tank, controller, accumulator, inverter and their connecting structure. Table 1 shows the geometric dimensions, physical and technical parameters of 180 W PTB parts.

Table 1: Physical and technical characteristics of 180 W PTD parts.

Parameter	Parameter
Maximum power of PVB, P_{max}	150W
Open circuit voltage of PVB, U_{oc}	22.80V
The short circuit current of PVB, I_{sc}	10.34A
Filling factor of VAC of PVB, f_f	0,71-0,73
Reflection coefficient of the reflector, R	0,5
Water capacity of polycarbonate heat collector, V	10 l
Thermal conductivity of polycarbonate	0,2-3,9W/m·°C
Cooling radiator, dimensions (cm)	35÷45cm
Fan, power (W)	3.6W
Water storage tank, volume (V)	20 l
Water pump, power (W)	5W
Cart, dimensions (cm)	Width 80 cm, length 100 cm, height 70 cm
Water circulation hose, dimensions (cm)	Diameter 1 cm, length 300 cm
Maximum power of PVB, P_{max}	150W

Figures 1 and 2 show a view of a new type of photothermal device that enables efficient operation in dry climate conditions made by scientists of the Institute of Physics and Technology.

Figure 1: A preview of a new type of PTD.

Figure 2: Rear view of the new type of PTD.

The parts that make up the cooling system of the new type of PTD perform the following tasks:

- - a new type of heat collector ensures reduction of energy losses of the photothermal battery.
- - the fan is installed on top of the radiator and serves to increase the efficiency of the radiator with the help of wind.
- - the pump passes the heated water in the new type of collector through the radiator and ensures that it enters the collector again.
- - the water tank serves to store up to 20 liters of warm water and for the good operation of the pump;
- - a small portable structure is used to place the radiator, fan, pump, water storage tank, controller, battery and inverter in the photothermal battery and connect them to each other by switching, and when moving them from one place to another, the mobile structure acts as a mechanical protection.

The results obtained from PTB and PVB were tested at the Heliopolygon of the Institute of Physics and Technology. The new cooling system, designed to adapt to different conditions, was tested in April and May 2024 in different modes for short-circuit voltage, short-circuit current, and device power at different values of solar radiation intensity.

The new type of PTB cooling system is based on the principle of the internal combustion engine cooling system, and cooling is provided by the circulation of water through a special radiator. In order to increase the efficiency of the radiator, heat transfer from the radiator is provided with the help of an additional fan. A cooling radiator is a device designed to cool the engine system and maintain normal temperature conditions. Such a simple, uncomplicated tool is the key to long-term and reliable operation of the engine [9-10]. Not using the fan installed on the radiator of the new type of cooling system is accepted by us as the mode when the cooling system is not fully used. If all components of the cooling system, including the fan, are used, the cooling system is considered to be fully used. The efficiency of the photovoltaic battery was studied when the cooling system was not fully used and when it was operating.

Figure 3 shows the results of comparison of operating voltages of PVB when the cooling system of the PTB is fully used.

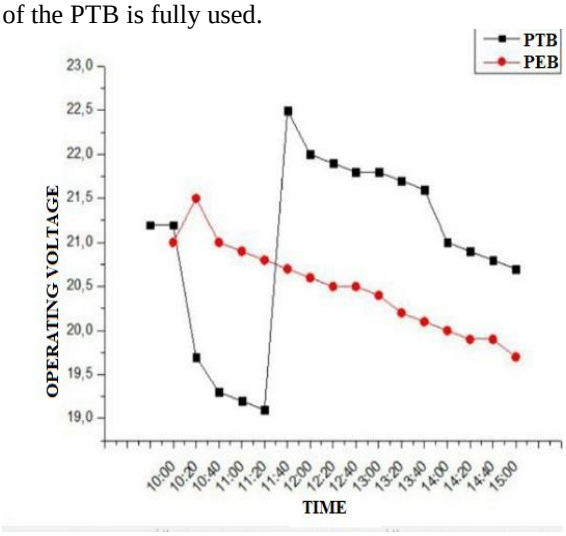


Figure 3: Time variation of the operating voltage of PVB when the cooling system of the PTB is fully used.

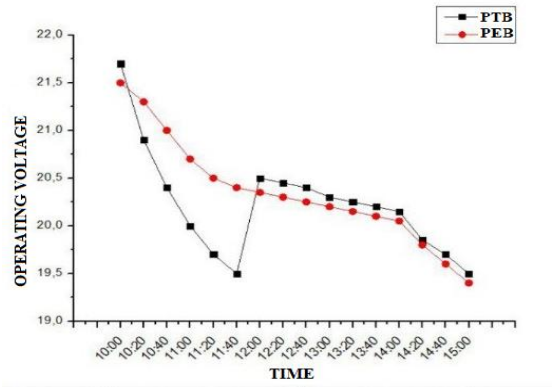


Figure 4: Time variation of the operating voltage of PVB when the cooling system of the PTB is not fully used.

Figure 3 shows the operating voltages of the PVB when the cooling system of the PTB is fully used. The black line is the value of PTB and the red line is the value of PVB. The test experiment was taken from 10:00 a.m. on April 2024 at the Heliopolygon of the Institute of Physics and Technology. Based on the graphs in Figure 3, we can see a linear decrease in voltage at PVB over time during the day. That is, the voltage of PVB has decreased over time. The main reason for this is the increase in air temperature, low wind speed (3-5 m/s), and the fact that the direction of the PTD does not correspond to the south. To overcome this loss, a heat collector based on cellular polycarbonate was installed in full thermal contact with the rear surface of the PVB (the PTB was made from the PVB) and attached to the cooling system. The results of the PTB based on this system are shown by the black line in Figure 3. Results for both PTB and PVB were taken at the same start time at 10:00 a.m. With no cooling system connected, it can be seen that the operating voltage of PTB decreases faster with time than the operating voltage of PVB. The reason for this can be explained by the fact that the heat collector is isolated from the external environment, and as a result, the temperature is maintained in a completely closed environment. As a result, at 11:20 a.m, operating voltage value of PTB decreased to 19 V, at this time, a new type of cooling system was put into operation, and at 11:40 a.m., operating voltage value of PTB increased to 22.8 V. The value of operating voltage of PVB is 20.8 V, i.e. it changes very little. Figure 4 presents the values of operating voltages of PTB and PVB when the cooling system of the new type of the PTB is not fully used. It can be seen that the graph of operating voltage of PVB is almost the same in Figures 3 and 4. When the cooling system of the new type is not fully used, the operating voltage of PTB showed higher values than the operating voltage of PVB. That is, the new type of cooling system reduces the heat accumulated in the rear part of the PTB, while reducing the operating voltage, it also allows to get hot water. Information about the hot water obtained when the new type of cooling system is fully used and when it is not fully used is presented in Figure 7.

Reflectors were installed on the sides of the new type of PTD to increase the intensity of the solar radiation current, and the short-circuit currents over time were compared with the short-circuit currents of the PVB over time.

Figures 5 and 6 show the time variation of the short-circuit current between 10:00 a.m. and 15:00, with PTB and PVB installed side by side. At 11:20 a.m., it was noted that the short-circuit current value increased from 8 A to 11 A by increasing the intensity of solar radiation by directing the reflectors to the two sides of the PTB. We can see from Figure 5 that when the natural solar radiation current intensity increases with time, the short-circuit current of PTB increases proportionally. When the sun reaches the solstice, the values of the short-circuit currents of PTB and PVB are 11.8 A and 8.2 A respectively. Figures 5 and 6 show the difference in short-circuit current values of PTB with time when the new type of cooling system is fully used and when it is not. That is, by fully using the new type of cooling system, the short-circuit current was partially increased.

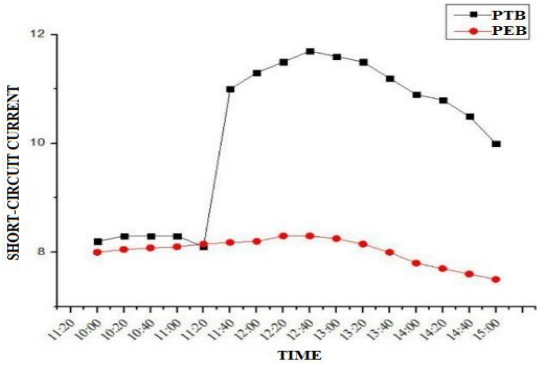


Figure 5: Time variation of short-circuit currents of PTB and PVB when the new type of cooling system is fully used.

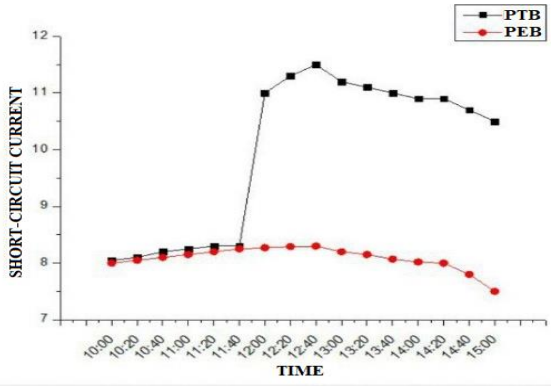


Figure 6: Time variation of short-circuit currents of PTB and PVB when the new type of cooling system is not fully used.

This new type of cooling system, along with increasing the physical parameters of PTB, allows obtaining hot water even in remote areas. PTD is equipped with a 20-liter water storage tank, and the

water is collected in the storage tank by absorbing the heat accumulated in the collector mounted on the back of the PTB. This heated water cools again through the radiator and goes to the collector, and the process continues over and over again. The hot water collected in the storage tank can be used by the village residents for many purposes (laundry, washing dishes, taking a shower, etc.). Figure 7 shows the values of the change of water temperature in the storage tank over time

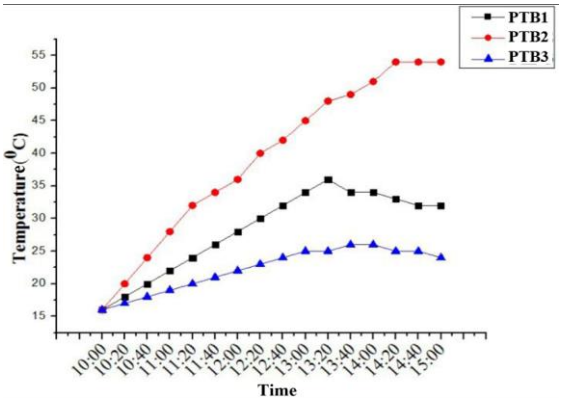


Figure 7: The values of the temperature change of hot water obtained: PTB 1 - when the new type of cooling system is fully used, PTB 2 - when the new type of cooling system is not fully used, PTB 3 - by conventional cooling.

PTB1 graph in Figure 7 shows the hot water temperature change obtained when the new type of cooling system of PTB is fully used. The temperature of the water in the water storage tank at 10:00 a.m. was 16^0 C. Gradually, as the intensity of solar radiation increases, the temperature of the water in the storage tank also rises. At 13:20, the water temperature showed its maximum value of 36^0 C. Later, the hot water temperature showed a value of 32^0 C at 15:00. The PTB2 graph also shows the values of the hot water temperature when the new type of cooling system of PTB is not fully used. At 10:00 a.m., the temperature of hot water was 16^0 C, and with the change of time, the water temperature rose to 54^0 C at 15:00. And the PTB3 graph shows the hot water temperature values obtained by conventional cooling of the PTB. The conventional method of cooling the PTB is to connect cold water to the bottom of the collector and hot water from the top of the collector. But cooling the PTB in this way brings a number of inconveniences. Firstly, it is difficult to find running water at a constant pressure in extremely dry areas. Secondly, after using the required part of the hot water from the collector, there is a problem of using the remaining part. The fact that the PTD with the new type of cooling system offered by us does not need a large amount of water for cooling, and the fact that the hot water is used as needed and the rest is not wasted, proves the efficiency of this device. But it can be seen from Figure 8 that the power of PTB cooled by conventional method is higher than the power of PTB cooled by other methods.

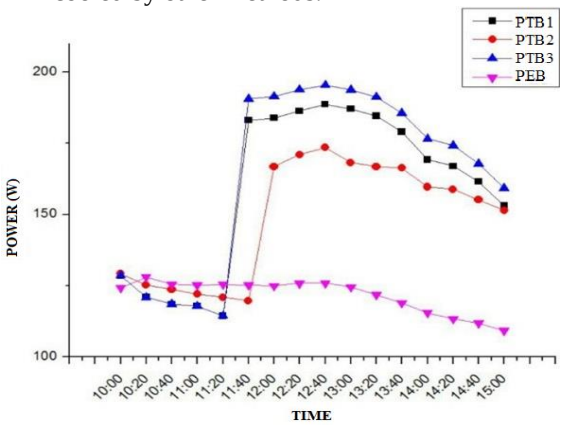


Figure 8: The values of power change over time when the new type of cooling system of PTB is fully used and is not fully used, when PTB is conventionally cooled and for PVB.

Figure 8 shows the results of power change of PVB. The natural solar radiation intensity reduced to PVB and the produced power are given. It can be seen from Figure 8 that the power values obtained in PVB showed much lower efficiency than the power values obtained in PTBs. The maximum power of PVB was 125 W. The PTB2 graph shows the power values obtained by partially using the new type of cooling system. Based on the data presented in Figure 7, it has been proven that hot water with a high temperature can be obtained with the improvement of the power of PTB compared to PVB, even with the partial use of a new type of cooling system. By partially using the new type of cooling system, it is possible to save the power required for the consumption of fans. The maximum power of PTB2 was 170 W. In Figure 8 PTB1 graph shows the results of power changes obtained when the new type of cooling system is fully used. It can be seen from the difference between the graphs PTB1 and PTB2 in Figure 8 that when the new type of cooling system is fully used the power changes have higher indicators than when the system is not fully used. The maximum power of PTB1 was 188 W.

The PTB3 graph is conventionally cooled and shows the obtained power values. The maximum power of PTB3 was 195 W. It can be seen from Figure 8 that when the PTB is cooled by continuous cold water inflow, the power values are more efficient than in all other cases. PTD with the new type of cooling system that we offer can be cooled in all three ways mentioned above. This depends on whether the PTD is installed in a very dry area or a place with a lot of water.

3 CONCLUSIONS

In this paper, it was shown that the power of PTB with a new type of cooling system is 35% to 50% more than the power of PVB. We can conclude the changes and values of other physical parameters based on the given results. By fully and partially using a new type of cooling system PTB, hot water of different temperatures was obtained. The ability to obtain hot water from 25^0 C to 55^0 C on average using the new type of cooling system was demonstrated. In the next researches, it is planned to see the possibility of developing mobile devices with sufficient capacity of PTD with a new type of cooling system to provide separately located rural facilities with electricity and partly with hot water

REFERENCES

[1] IRENA, Renewable Energy Statistics 2024, Int. Renewable Energy Agency, Abu Dhabi, 2024. ISBN: 978-92-9260-614-5. [Online]. Available: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Jul/IRENA_Renewable_Energy_Statistics_2024.pdf.

[2] M.J. Elswijk, M.J. Jong, K.J. Strootman, J.N.C. Braakman, E.T.N. De Lange, and W.F. Smith, “Photovoltaic thermal collectors in large solar thermal systems,” in Proc. 19th European PV Solar Energy Conf. and Exhibition, Paris, France, June 2004, pp. 7–11.

[3] B. Yuldoshov, E. Saitov, J. Khaliyarov, S. Toshpulatov, and F. Kholmurzayeva, “Effect of temperature on electrical parameters of photovoltaic module,” Proc. Int. Conf. Appl. Innovation IT, vol. 11, no. 1, 2023, pp. 291–295. doi:10.25673/101957.

[4] M.N. Tursunov, Kh. Sabirov, I.A. Yuldoshev, B.M. Turdiev, and I.M. Komolov, “Different constructions of photothermal batteries: comparative analysis,” Heliotechnika, no. 1, 2017, pp. 26–29.

[5] J.S. Akhatov, I.A. Yuldashev, and A. Shalimov, “Experimental investigations on PV-T collector under natural condition of Tashkent,” Energy Procedia, vol. 39, 2013, pp. 2327–2336, Proc. ISES Solar World Cong., Mexico.

[6] I.R. Jurayev, I.A. Yuldoshev, and Z.I. Jurayeva, “Experimental study of a thin-film photovoltaic thermal battery in natural conditions,” Appl. Solar Energy, vol. 59, no. 4, 2023, pp. 498–506. (CiteScore, IF=2.5).

[7] A.M. Elbreki, K. Sopian, A. Fazlizan, and A. Ibrahim, “An innovative technique of passive cooling PV module using lapping fins and planner reflector,” Case Stud. Thermal Eng., vol. 19, 2020. doi:10.1016/j.csite.2020.100607.

[8] I. Jurayev, I. Yuldoshev, and Z. Jurayeva, “Effects of temperature on the efficiency of photovoltaic modules,” Proc. Int. Conf. Appl. Innovation IT, vol. 11, no. 1, 2023, pp. 199–206. [Online]. Available: https://doi.org/10.25673/101938.

[9] I. Jurayev, I. Yuldoshev, and Z. Jurayeva, “Results of study of photovoltaic thermal battery based on thin-film module by modeling and computational methods,” Proc. Int. Conf. Appl. Innovation IT, vol. 12, no. 1, 2024, pp. 243–249. [Online]. Available: https://doi.org/10.25673/115707.

[10] I.A. Hassan, S.R. Fafraj, and A.J. Kadem, “Enhancing photoelectric conversion efficiency of photovoltaic panel using cooled water by evaporation,” Eng. Technol. J., vol. 35, part A, no. 5, 2017, pp. 525–531.

Annual Analysis of On-Grid Station with an Installed Capacity of 200.75kW: The Case of the City of Termez

Muhammadjon Tursunov[1], Khabibullo Sabirov[1], Tohir Axtamov[1], Umirbek Abdiyev[2], Azamat Narbayev[2], Sirojiddin Toshpulatov[2] and Tulkin Buzrukov[3]
[1]*Physical-Technical Institute, Chingiz Aytmatov Str. 2B, 100084 Tashkent, Uzbekistan*
[2]*Termez State University, Barkamol Avlod Str. 43, 190111 Termez, Uzbekistan*
[3]*Termez University of Economics and Service, Farovon Str. 4B, 190111 Termez, Uzbekistan*
muhammadtursunov54@gmail.com,sabirovhabibullo@gmail.com, axtamovtz@uzsci.net, umr79@mail.ru, narbaev@ro.ru, toshpulatovs@tersu.uz, tolqinbuzrukov@gmail.com

Keywords: CUF, Renewable Energy, Solar System, Solar Station, Monitoring, Energy.

Abstract: The use of solar energy, which is one of the renewable energy sources, is increasing every year. And the installed capacity of solar plants is also increasing all over the world. This shows that the transition to green energy is accelerating. It will be beneficial to monitor the useful working capacities of such solar plants installed in the world, and in our country, the amount of energy produced, and the performance of months and seasons. In the article, the annual amount of energy produced by an on-grid solar plant with a capacity of 200.75kW is equal to 198.32MWh, and calculations give the payback period is 5.6 years. The maximum value of the installed capacity utilization coefficient was 19.57% in June, the minimum value was 5.41% in January, and the average annual value was 11.25%. In addition, the specific energy coefficient peaked at 4.7 kWh/kW in June. During the winter months, it reached its lowest values, with 1.53 kWh/kW in December, 1.3 kWh/kW in January, and 1.54 kWh/kW in February.

1 INTRODUCTION

Today, the use of renewable energy sources is increasing at a high rate. Solar energy is one of the most important among renewable energy sources. According to the International Renewable Energy Agency (IRENA), it is planned to reach the capacity of 8GW of solar plants in Uzbekistan by 2030 [1]. In addition, Uzbekistan has set a plan to increase the amount of energy obtained from renewable energy sources to 40% by 2030 [2]. As a result of the construction of large-scale solar power plants in recent years, it is possible to determine the energy analysis, economic payback period and many other indicators of this type of power plant by calculating their capacity utilization factor (CUF) [3]. Many on-grid solar stations are based on efficiently using the building's roof. However, energy losses caused by high temperatures [4-8] and pollution [9] have not been eliminated. It is possible to monitor the amount of energy produced in real climatic conditions by controlling it through the Internet, which is one of the modern conveniences.

In the research that started several decades ago in European countries, scientists such as Muyiwa S. A., Emil E.T. analyzed the PV plant in Norway and obtained a value of 10.58% for the annual CUF [10]. Ramesh Chaudhary and Pratiksinh Chavda have reviewed the influence of climatic factors on CUF in their article [11]. Quantities related to the energy produced by the PV system, and system losses [12-14] were considered in detail in the works. Various economic indicators of SS were also analyzed.

In the article, the annual energy analysis of the 200.75kW on-grid network station installed in the student accommodation building of the Termez State University in the city of Termez, which has high annual radiation, was carried out.

2 METHODS AND MATERIALS

2.1 Study Location

The 200.75kW solar station in the study is located on the territory of Termez State University (longitude of 37° 13'57"N, latitude of 67° 17'8"E) (Figure 1). Termez is a city located in the southernmost part of the Republic of Uzbekistan, the climate here is very hot in summer and short and cold in winter.

Figure 1: Location of on-grid solar system.

2.2 System Configuration and Operation

In the study, the installed power utilization coefficient of the 200.75kW on-grid solar plant installed on the territory of Termez State University was analyzed. This system is installed on the roof of one of the university dormitories facing south. 550W monocrystalline panels were used in the on-grid solar system, and their electrical characteristics are given in Table 1.

Table-1: Physical parameters of photoelectric battery.

Model Type	LF550TU-36MH
Peak Power (P_{max})	550W
Module Efficiency	21.3%
Maximum Power Voltage (U_{mp})	42.11V
Maximum Power Current (I_{mp})	13.06A
Open Circuit (U_{oc})	50.28V
Short Circuit Current (I_{sc})	13.86A
Maximum System Voltage Nominal	1500V
E=1000W/m^2, T_{PEB}=25^{o}C, AM=1,5 STC conditions	

A total of 365 panels of 550W monocrystalline panels were used in the solar station. The panels were installed in a stationary position at a tilt angle of 30^0 to the south. SS has been working since July 2023.

In this solar system, solar panels generate DC electricity by absorbing sunlight, and the solar inverter converts this DC electricity into AC electricity, which can then be used directly at home or in a business. If the system produces more power than is being consumed, the surplus is fed into the main electrical grid via solar net metering. The diagram of the 200.75kW on-grid solar plant we installed is shown in Figure 2. In this setup, the inverter is connected to the Internet via Wi-Fi, allowing for monitoring of the daily energy produced by the system, the energy generated, and its consumption at any time of the day. The system uses a SUN2000-100KTL-M2 inverter and two SUN2000-50KTL-M3 inverters, which are modern equipment for solar energy applications. Full reports of daily and monthly power produced by the system, as well as daily weather information, can be accessed through the site[1] using mobile phones or computers. This information was used directly in the article.

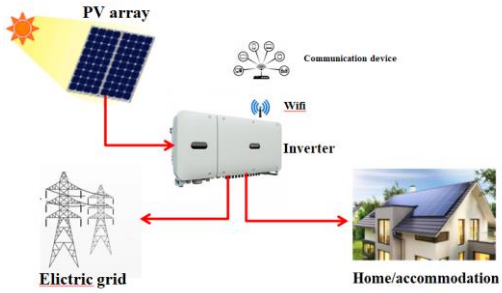


Figure 2: Schematic block circuit diagram of the SS.

2.3 Capacity Utilization Factor

The CUF (Capacity Utilization Factor) is the ratio of the actual energy output of an AC system to the energy a PV system would produce if it ran at nominal power [15]. Another way to define it is by measuring how long an electrical system operates at full 100% capacity. Coefficients calculated over short intervals can vary significantly, so the accuracy of weekly, monthly, or yearly calculations improves as the time period increases.

$$\eta_{CUF} = \frac{E_{yield}}{P_{PV,rated} \cdot Time_{Fixed}} \quad (1)$$

[1] https://region02eu5.fusionsolar.huawei.com

Where E_{yield} is SS's total energy produced for a fixed time. $P_{PV,rated}$ is the installed power of SS. $Time_{Fixed}$ is the exact time taken for SS to produce E_{yield} energy.

3 RESULTS AND DISCUSSION

The collected data from July 2023 to June 2024 (one year) used to study the photovoltaic plant's performance were carried out at the Termez State University. The annual amount of energy produced is equal to 198.32GWh, and Figure 3 gives full monthly informations.

The highest amount of energy was produced in June, which was 3.5 times more than the lowest amount of energy produced in January 8.08MWh and was equal to 28.29MWh. Considering the economic recovery period of our station and the energy produced by it, we carried out the economic analysis of SS as follows.

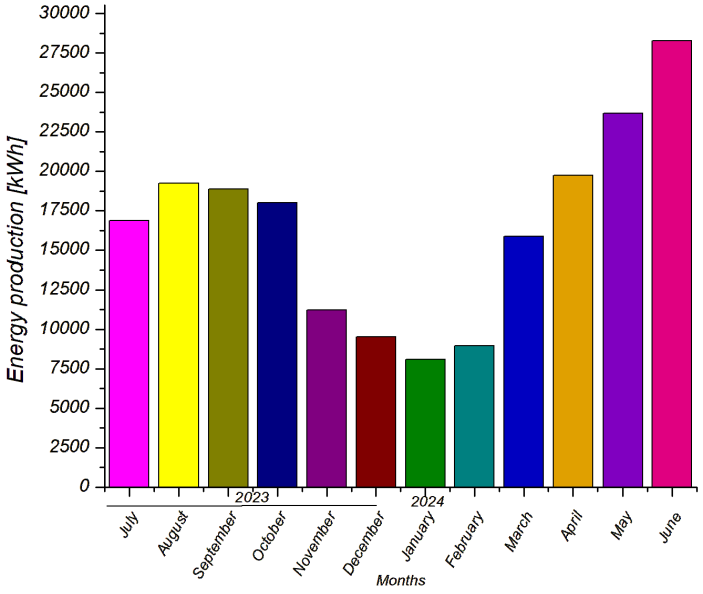


Figure 3: The energy production in one year.

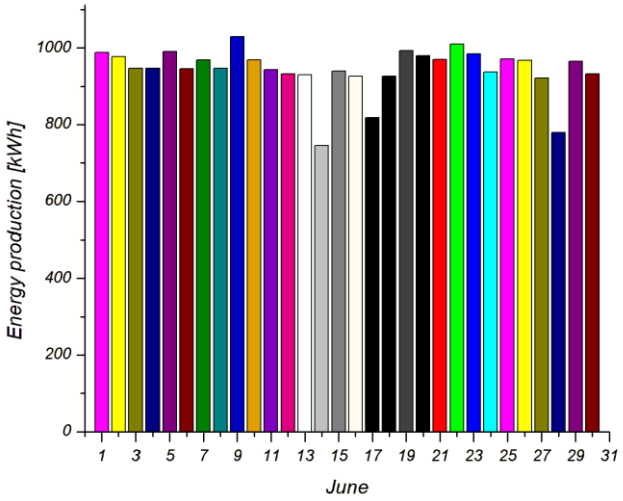


Figure 4: The daily energy reports produced in June.

$$n = \frac{P \cdot \$400}{E \cdot \$0.072} \qquad (2)$$

Here: P=200.75kW, $400 is the average price of a 1kW on-grid station, E=198320kWh annual energy production, $0.072 is the price of 1kWh energy in our country. In our calculations, we did not take into account the increase in energy prices and inflation, and we also did not take into account the costs associated with the operation of solar plant installations. The payback period of the solar plant was equal to 5.6 years. Figure 4 reports the amount of energy produced per day in June. In this month, the most daily 1.03MWh of energy was produced on June 9, and the least 745.07kWh was produced on June 14.

Annual and monthly solar plant CUF was calculated using the formula given above, and the results are reflected in graph 5.

The values of CUF are also reflected in accordance with the months of energy production. In this case, the average annual CUF value is equal to 11.25%, due to the smallest values of 6.37% for December, 5.41% for January, and 6.41% for February. In the summer season, due to the length of days, the high amount of falling radiation, and almost all days being open, the CUF value reached its maximum value in June and reached 19.57%.

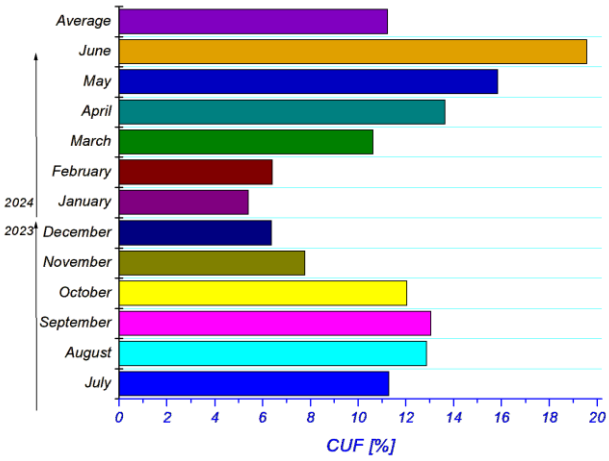


Figure 5: CUF in months.

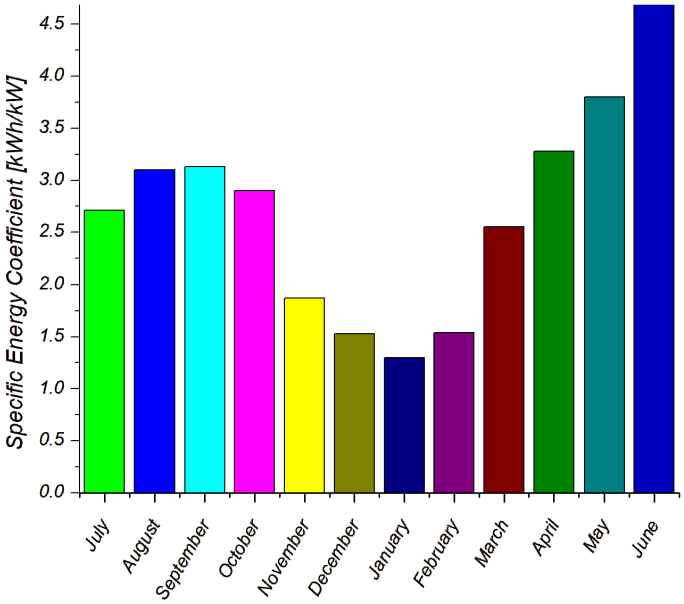


Figure 6: Specific energy coefficient in months.

As sunny days become more frequent, the CUF will undoubtedly rise. The table above indicates that the average sunlight duration in winter is 4.7 hours, which increases to 8.3 hours in spring and 12.2 hours in summer. However, it's important to note that as the temperature of the panels rises, so does the power loss [16-20]. Therefore, it is essential to consider the panels' fixed position and the level of dust accumulation. Our primary objective is to study the CUF of SS, and we have thoroughly examined these aspects earlier. In addition, to evaluate energy efficiency, how many kWh of energy are obtained during the day from 1kW of power is also important. The change in the average values of this quantity over months is given in Figure 6.

The maximum value from the graph is 4.7kWh/kW for June, this indicator reached minimal values in the winter months due to stationarity and no pollution protection. However, this high index in the summer months can be said to be a good result.

4 CONCLUSIONS

It is important to determine the efficiency of the renewable solar plant by the amount of energy it produces in natural conditions and to find the economic payback period. In the article, the annual CUF value of the rooftop on grid solar plant is equal to 11.25% and the payback period is equal to 5.6 years, the average annual energy production is equal to 198.32GWh. Various harmful gases that can be released into the environment have been prevented. Despite the fact that the panels heat up to a high temperature, due to the length of the day in the summer months, the solar plant has achieved a maximum average energy production of 1MWh. The specific energy coefficient also reached its maximum value which was 4.7kWh/kW for June. And for the winter months, this indicator reached its minimum values in December, January and February, respectively 1.53, 1.3 and 1.54kWh/kW.

5 ACKNOWLEDGMENTS

We would like to express our gratitude to Termiz State University, which greatly supports the field of renewable energy sources and scientific work.

REFERENCES

[1] E. Bellini, "Uzbekistan’s 500 MW PV tender concludes with lowest bid of $0.02888/kWh," [Online]. Available: https://www.pv-magazine.com/2023/01/02/uzbekistans-500-mw-pv-tender-concludes-with-lowest-bid-of-0-02888-kwh/.

[2] "Renewable energy sources," [Online]. Available: https://lex.uz/uz/docs/6600413.

[3] M. Tursunov, Kh. Sabirov, T. Axtamov, U. Abdiyev, M. Chariyev, B. Yuldoshov, and S. Toshpulatov, "Capacity Utilization Factor (CUF) of the 70kW on-grid solar station in the dry climate of Termez," E3S Web of Conf., vol. 401, 02059, 2023. doi: 10.1051/e3sconf/202340102059.

[4] N. Aoun, "Methodology for predicting the PV module temperature based on actual and estimated weather data," Ener. Conv. Man.: X, vol. 14, p. 100182, 2022. [Online]. Available: https://doi.org/10.1016/j.ecmx.2022.100182.

[5] M. Abd-Elhady, Z. Serag, and H. Kandil, "An innovative solution to the overheating problem of PV panels," Energy Convers. Manag., vol. 157, pp. 452-459, 2018.

[6] S. Edalati, M. Ameri, and M. Iranmanesh, "Comparative performance investigation of mono- and poly-crystalline silicon photovoltaic modules for use in grid-connected photovoltaic systems in dry climates," Appl. Energy, vol. 160, pp. 255-265, 2015. doi: 10.1016/j.apenergy.2015.09.064.

[7] A. Singh, "Analyzing policy interventions to stimulate suitable energy sources for the most polluted states of India," Renew. Sustain. Energy Rev., vol. 197, art. no. 114397, 2024.

[8] I. Jurayev, I. Yuldoshev, and Z. Jurayeva, "Effects of temperature on the efficiency of photovoltaic modules," Proc. Int. Conf. Appl. Innov. IT, vol. 11, no. 1, pp. 199-206, 2023. doi: 10.25673/101938.

[9] S. Klychev et al., "Free deposition of dust on inclined solar batteries," Appl. Solar Energy (Engl. transl. of Geliotekhnika), vol. 57, no. 5, pp. 403-408, 2021.

[10] A. Muyiwa and V. Emil, "Preliminary assessment of a small-scale rooftop PV-grid tied in Norwegian climatic conditions," Energy Convers. Manag., vol. 90, pp. 458-465, 2015.

[11] R. Chaudhary and P. Chavda, "To study the temporal variation of capacity utilization factor (CUF) of PV based solar power plant with respect to climatic condition," Curr. World Environ., vol. 11, no. 2, pp. 654-661, 2016. doi: 10.12944/CWE.11.2.38.

[12] A. Sagani, J. Mihelis, and V. Dedoussis, "Techno-economic analysis and life-cycle environmental impacts of small-scale building-integrated PV systems in Greece," Energy Build., vol. 139, pp. 277-290, 2017.

[13] A. Duman and Ö. Güler, "Techno-economic analysis of off-grid photovoltaic LED road lighting systems: A case study for northern, central and southern regions of Turkey," Build. Environ., vol. 156, pp. 89-98, 2019.

[14] O. Ayadi, R. Al-Assad, and J. Al-Asfar, "Techno-economic assessment of a grid connected photovoltaic system for the University of Jordan," Sustain. Cities Soc., vol. 39, pp. 93-98, 2018.

[15] A. Kumar et al., "Towards better performances for a novel rooftop solar PV system," Sol. Energy, vol. 216, pp. 518-529, 2021.

[16] L. Mensah, J. Yamoah, and M. Adaramola, "Performance evaluation of a utility-scale grid-tied solar photovoltaic (PV) installation in Ghana," Energy Sustain. Dev., vol. 48, pp. 82-87, 2019. doi: 10.1016/j.esd.2018.11.003.

[17] M. Tursunov et al., "Analysis of electric and thermal efficiency of crystal silicon small power suppliers," Proc. Int. Conf. Appl. Innov. IT, vol. 11, no. 2, pp. 167-171, 2023.

[18] C. Temaneh-Nyah and L. Mukwekwe, "An investigation on the effect of operating temperature on power output of the photovoltaic system at University of Namibia Faculty of Engineering and IT campus," Proc. 3rd Int. Conf. Digit. Inf., pp. 22-293, 2015.

[19] I. Jurayev, I. Yuldoshev, and Z. Jurayeva, "Experimental study of a thin-film photovoltaic thermal battery in natural conditions," Appl. Solar Energy, vol. 59, no. 4, pp. 498-506, 2023.

[20] M. Asvad, M. Gorji, and A. Mahdavi, "Performance analysis of a solar module with different reflectors and cooling flow fields," Appl. Therm. Eng., vol. 219, 2023. doi: 10.1016/j.applthermaleng.2022. 119469.

Optimal Planning of Short-Term Modes of Power Systems Containing Energy Storage Devices

Tulkin Gayibov[1] and Boburjon Murodov[2]
[1]*Tashkent State Technical University, University Str. 2, 100095 Tashkent, Uzbekistan*
[2]*JSC "National Power Networks of the Republic of Uzbekistan", Murtazayev Str. 1, 100095 Tashkent,* Uzbekistan
tulgayibov@gmail.com, b.murodov@uzbekistonmet.uz

Keywords: Power Electric System, Energy Storage, Load Schedule, Mathematical Model, Algorithm of Optimization.

Abstract: The development of the electric power industry at the present stage throughout the world is characterized by an increased rate of introduction of power plants operating on renewable energy resources, primarily solar and wind power plants. The unevenness of the load schedules of electricity consumers is increasing. To reliably provide consumers with high-quality electricity at minimal economic costs, electricity storage devices are being introduced into such energy systems. Under such conditions, the tasks of optimal planning of short-term modes of electric power systems, which consists of determining for each time interval the optimal values of all regulated parameters, in particular loads, through the use of the regulating capabilities of storage devices, become much more complicated. At present, despite the existence of several models and algorithms for optimization of the energy storage devices as part of the power system, the issues of determining the optimal operating modes of power systems in a short-term cycle through the use of their regulatory capabilities based on the use of rigorous mathematical models and algorithms have not been sufficiently studied. This paper presents a mathematical model for the problem of optimization of the modes of a power electric system containing adjustable energy storage devices and the algorithm for aligning consumer load schedules. The results of a study of the effectiveness of the proposed optimization model and algorithm using the example of aligning the daily load schedules of the Central part of the electric networks of JSC "National Power Networks of the Republic of Uzbekistan" are given.

1 INTRODUCTION

The modern development of energy throughout the world is characterized by the widespread introduction of stations powered by renewable energy sources into power systems and an increase in the unevenness of consumer load schedules [1, 4, 11]. To reliably supply consumers with high-quality electricity at minimal costs, some energy systems use energy storage devices, which serve to cover electricity imbalances resulting from random changes in plant capacities and to level out consumer load schedules. Under these conditions, the tasks of optimal planning and operational control of the modes of electric power systems become much more complicated. At the same time, it is necessary to improve existing mathematical models and algorithms for solving the problem, taking into account the alignment of consumer load schedules through the use of the regulating capabilities of energy storage devices.

The existing literature presents the results of many developments devoted to the optimization of the configuration, characteristics and operating modes of energy storage devices in electric power systems, which undoubtedly made a significant contribution to the development of the general scientific theory in this direction. In particular, in [1-6] the results of the authors' research on the selection of optimal parameters of various types of energy storage devices in power supply systems are presented. In [7-14] the results of research on optimization of operating modes of various types of energy storage devices in energy systems are presented. In particular, in [7] the problem of optimization of operating modes of energy storage devices in power supply system of buildings taking into account the uncertainty of the load, source and mechanism of rough response to demand is considered. An interval mode optimization model based on shared energy storage and refined demand response is proposed. In [8], a general overview of works devoted to the optimization of the

parameters of batteries when used for various purposes, in particular to ensure the quality of electricity, increase the efficiency of operation by regulating the required power, minimizing active losses due to accumulation and also forms of objective functions in implicit forms are given. In addition, the main approaches to using various heuristic and artificial intelligence methods to solve such a problem are presented. In [9], the results of studies on ensuring the quality of electricity through large-scale energy storage in power systems are presented.

An analysis of existing developments shows that the issues of determining the economical operating modes while ensuring the required reliability, quality of electricity and environmental protection of electric power systems with system energy storage devices based on the use of rigorous mathematical models have not been sufficiently studied. Most of these works do not study the issues of leveling the load schedules of power systems due to energy accumulation. Methods and algorithms proposed in some works for solving similar problems for electrical networks of enterprises [15, 16] and small autonomous systems cannot be directly used for large power systems. In this regard, the development and implementation of effective models and methods for optimization the modes of power systems containing large energy storage devices, taking into account all influencing factors, remains an urgent task.

In this regard, this paper presents the results of research on the development of a mathematical model and algorithm for solving the problem of optimization the modes of electric power systems, taking into account the regulatory capabilities of the system energy storage devices they contain.

2 MATHEMATICAL MODELING AND ALGORITHM OF OPTIMIZATION

The mathematical model of the problem of optimal mode planning for period T of a power system with energy storage devices can be formulated as follows:

- minimize the objective function, which is a function of the total costs associated with fuel consumption in thermal power plants (TPP):

$$B = \sum_{t=1}^{T}\sum_{i=1}^{n} B_i^{(t)}\left(P_i^{(t)}\right) \to \min \qquad (1)$$

taking into account the constraints;

- on power balance in each time interval of the period under consideration T

$$\sum_{i=1}^{n} P_i^{(t)} + P_{REN}^{(t)} + P_S^{dch(t)} = P_L^{(t)} + P_S^{ch(t)}, \quad t = 1, 2, .., T; \qquad (2)$$

- on permissible minimum and maximum power of thermal power plants:

$$P_i^{(t)\min} \le P_i^{(t)} \le P_i^{(t)\max}, \quad ; \quad i = 1, 2, n; \; t = 1, 2, ..., T \qquad (3)$$

- on permissible minimum and maximum charging and discharging power of energy storage device

$$0 \le P_S^{ch(t)} \le P_S^{ch.\max}, \quad t = 1, 2, ..., T; \qquad (4)$$

$$0 \le P_S^{dch(t)} \le P_S^{dch.\max}, \quad t = 1, 2, ..., T; \qquad (5)$$

- on permissible minimum and maximum energy (capacity) of charge and discharge of energy storage device

$$W_S^{\min} \le W_S^{(t)} \le W_S^{\max}, \quad t = 1, 2, ..., T, \qquad (6)$$

where n is the number of thermal power plants participating in optimization; T is the number of time intervals during the period under consideration; $P_i^{(t)}$, $P_S^{ch(t)}, P_S^{dch(t)}$ - power of the i-th thermal power plant, charge and discharge of the storage device in the t-th time interval of the period under consideration, respectively; $P_{REN}^{(t)}, P_L^{(t)}$ - the total power of stations operating on renewable energy sources and the load of the power system in t-th time interval; $P_S^{ch.\max}$, $P_S^{dch.\max}$ - permissible maximum charge and discharge powers of the storage device; $W_S^{(t)}$, $W_S^{\min}$, $W_S^{\max}$ - the amount of energy in the storage device in t-th time interval, as well as its permissible minimum and maximum values.

In cases where a battery is used as an energy storage device, $W_S^{\min}$ represents the amount of energy corresponding to the greatest depth of its discharge. If it is accepted that the charge energy in the storage device at the beginning and end of the planning period is equal to $W_S^{\min}$, then the amounts of charge and discharge of the storage device during the planning period T must be the same, i.e. the following conditions must be met:

$$\sum_{t=1}^{T} P_S^{ch(t)} = \sum_{t=1}^{T} P_S^{dch(t)} = W_S^{\max} - W_S^{\min}. \qquad (7)$$

Integral constraint (6) can be replaced by the following constraints, expressed in terms of the unknown charging and discharging powers of the storage device:

$$\sum_{i=1}^{t} P_S^{ch(i)} - \sum_{i=1}^{t} P_S^{dch(t)} \le W_S^{\max} - W_S^{\min},\ t = 1, 2, ..., T, \quad (8)$$

$$-\sum_{i=1}^{t} P_S^{ch(i)} + \sum_{i=1}^{t} P_S^{dch(t)} \le -W_S^{\min},\ t = 1, 2, ..., T. \quad (9)$$

Optimal planning of the power system mode for period T involves determining for each time interval the optimal powers stations, which participate in optimization and the charging and discharging powers of the storage device, at which the objective function (1) has a minimum value and the above presented constraints are met. As a result, simultaneously with solving the problem within the given constraints, the total load schedule of consumers is leveled due to the regulating capabilities of the energy storage device. Therefore, the problem under consideration can be solved by dividing it into two stages. At the first stage, the problem of leveling the consumer load schedule is separately solved by using the regulating capabilities of the storage device. And at the second stage, the optimization of the power system mode using the equalized load schedule obtained as a result of the first stage, based on the use of traditional methods and algorithms is carried out. Below we will consider the problems associated with implementing the first stage.

In the problem of leveling the load schedule, the objective function is presented in the following form:

$$f = \sum_{t=1}^{T} \left(P_{L.p}^{(t)} - \sum_{i=1}^{T} P_L^{(i)} / T \right)^2 \to \min, \quad (10)$$

where is $P_{L.p}^{(t)}$ is the calculated load in t-th time interval, obtained as a result of leveling the schedule.

Minimization of the last function is carried out on $P_{L.p}^{(t)}$, $P_S^{ch(t)}$, $P_S^{dch(t)}$, taking into account constraints on power balance in each time interval

$$P_{L,p}^{(t)} + P_S^{ch(t)} - P_S^{dch(t)} = P_L^{(t)},\quad t = 1, 2, ..., T, \quad (11)$$

permissible calculated loads

$$P_{L.p}^{(t)\min} \le P_{L.p}^{(t)} \le P_{L.p}^{(t)\max}, \quad (12)$$

$$t = 1, 2, ..., T \text{ and } (4), (5), (7), (8), (9).$$

The proposed algorithm for solving the resulting problem involves minimization the function (10) taking into account constraints in the form of equalities (11) and (7) by indefinite Lagrange multipliers, functional constraints in the form of inequalities (8) and (9) by penalty functions as in [20], and simple constraints (12), (4) and (5) through fixing at each iteration the variables that go beyond the permissible limits on the corresponding violated boundaries. Thus, the following generalized objective function is minimized

$$\begin{aligned} F = f &+ \sum_{t=1}^{T} \mu_1^{(t)} \left(P_{н.p}^{(t)} + P_S^{ch(t)} - P_S^{dch(t)} - P_н^{(t)} \right) + \\ &+ \mu_2 \left(\sum_{t=1}^{T} P_S^{ch(t)} - W_S^{\max} + W_S^{\min} \right) + \\ &+ \mu_3 \left(\sum_{t=1}^{T} P_S^{dch(t)} - W_S^{\max} + W_S^{\min} \right) + \\ &+ \sum_{t=1}^{T} Pf_1^{(t)} + \sum_{t=1}^{T} Pf_2^{(t)} \to \min \end{aligned}, \quad (13)$$

taking into account constraints (12), (4) and (5). Where $\mu_1^{(t)}, \mu_2, \mu_3$ are the undetermined Lagrange multipliers, taking into account the corresponding restrictions; $Pf_1^{(t)}$, $Pf_2^{(t)}$ – penalty functions that take into account functional restrictions in the form of inequalities (8) and (9), which have quadratic forms, as in [20].

The values of the undetermined Lagrange multipliers at each k-th iteration $\mu_1^{(t)}, \mu_2, \mu_3$ are found based on the sequential solution of the equations obtained by equating the following partial derivatives to zero:

$$\frac{\partial F}{\partial P_{L.p}^{(t)}} = 0,\quad \frac{\partial F}{\partial P_S^{ch(t)}} = 0,\quad \frac{\partial F}{\partial P_S^{dch(t)}} = 0.$$

The optimal calculated loads, charging and discharging powers for each time interval at the kth iteration are calculated using the following formulas:

$$\left.\begin{aligned} P_{L.p}^{(t)(k)} &= P_{L.p}^{(t)(k-1)} - h_L^{(t)(k)} \cdot \frac{\partial F}{\partial P_{L.p}^{(t)}}^{(k-1)}, \\ P_S^{ch(t)(k)} &= P_S^{ch(t)(k-1)} - h_S^{ch(t)(k)} \cdot \frac{\partial F}{\partial P_S^{ch(t)}}^{(k-1)}, \\ P_S^{dch(t)(k)} &= P_S^{dch(t)(k-1)} - h_S^{dch(t)(k)} \cdot \frac{\partial F}{\partial P_S^{dch(t)}}^{(k-1)}. \end{aligned}\right\}, \quad (14)$$

where $h_L^{(t)(k)}$, $h_S^{ch.(t)(k)}, h_S^{dch.(t)(k)}$ are steps in the direction of descent to the minimum in k-th iteration, defined as in [20, 21].

The condition for the convergence of the iterative process is

$$\left| f^{(k-1)} - f^{(k)} \right| \le \varepsilon_F, \quad (15)$$

when all specified constraints are met. In cases where condition (15) is met under violated constraints, penalties for such constraints are increased and the calculation proceeds to the next iteration.

An enlarged block diagram of the proposed algorithm is shown in Figure 1.

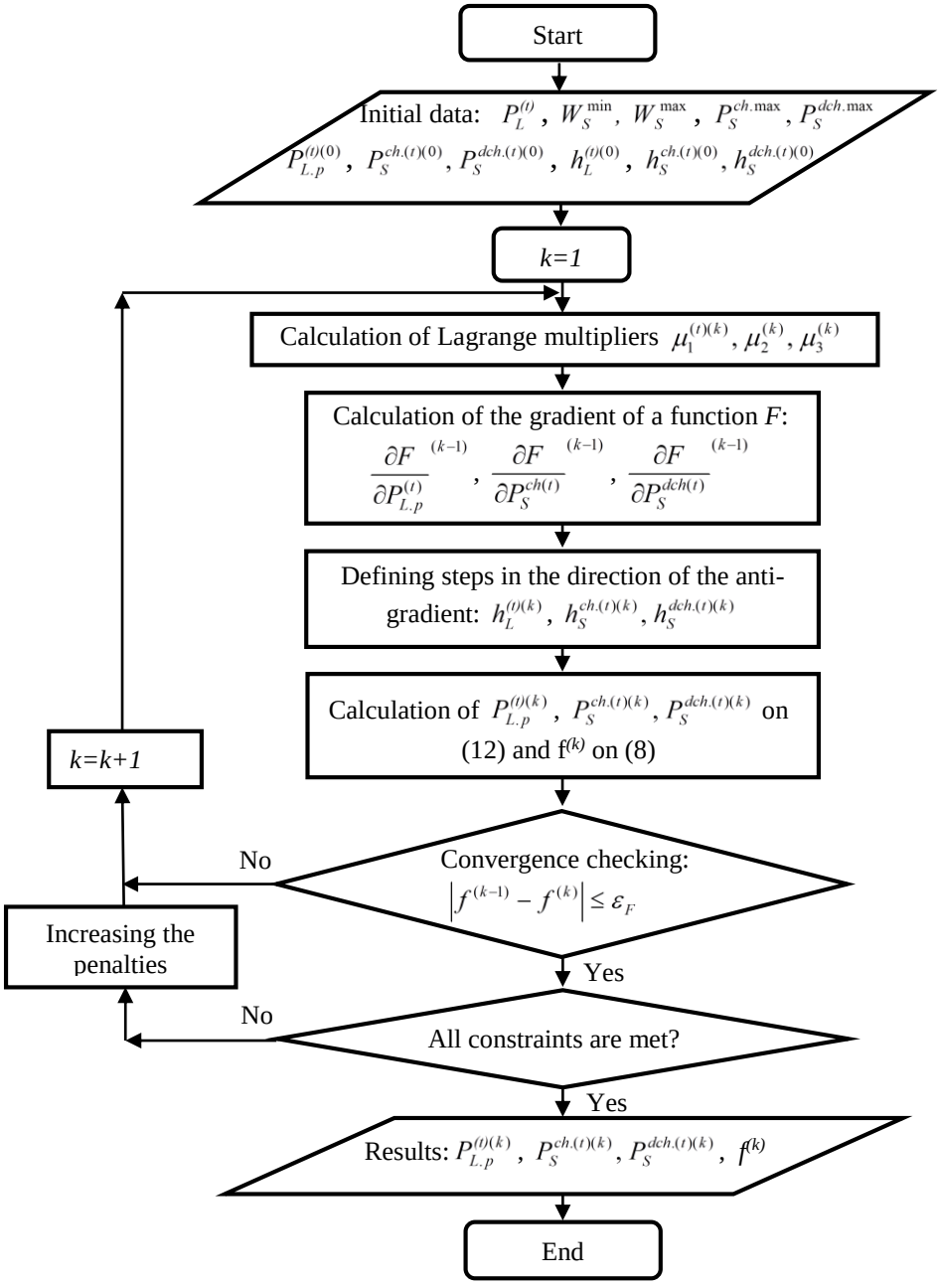


Figure 1: Enlarged block diagram of the optimization algorithm.

3 RESULTS

The effectiveness of the proposed mathematical model and optimization algorithm was studied using the example of leveling the daily schedules of total loads in the central part of the main electric networks of "National electric grid of Uzbekistan" JSC through the use of the regulatory capabilities of energy storage devices.

The day June 15.2022 for which the daily schedule of load is represented by eight characteristic time intervals (Table 1), was taken as an example.

As an example, an energy storage device with the following parameters was selected: $W_S^{\max}$ =800 MWh., $W_S^{\min}$ =160 MWh., $P_S^{ch.\max}$ =400 MW, $P_S^{ch.\min}$ =0, $P_S^{dch.\max}$ =400 MW, $P_S^{dch.\min}$ =0.

At the beginning of the day, the energy stored in the storage device is $W_S^{\min}$ = 160 MWh.

Table 1 shows the initial schedule of the total load $P_L(t)$ and optimization results: the calculated load after leveling using the storage capacity $P_{L.p}(t)$, charging $P_S^{ch.}(t)$ and discharging $P_S^{dch.}(t)$ power of the storage devices.

Figure 2 shows daily schedule of the initial and calculated total loads $P_L(t)$, $P_{L.p}(t)$, and in Figure 3 schedules of charging and discharging power of storage device. In this case, the discharge power is shown with a negative sign.

Table 1: Daily schedule of initial total loads of consumers and optimization results.

Number of the time interval	1	2	3	4	5	6	7	8
Durition of time interval, hours	0-4	4-8	8-10	10-14	14-16	16-18	18-22	22-24
$P_L^{(t)}$, MW	2196.0	2534.0	2965.0	3395.0	2965.0	2664.0	3300.0	2664.0
$P_{L.p}^{(t)}$, MW	2304.3	2585.7	2935.9	3264.1	2935.9	2980.0	3120.0	2664.0
$P_S^{ch.(t)}$, MW	108.3	51.7	0.0	0.0	0.0	320.0	0.0	0.0
$P_S^{dch.(t)}$, MW	0.0	0.0	29.1	130.9	29.1	0.0	160.0	0.0

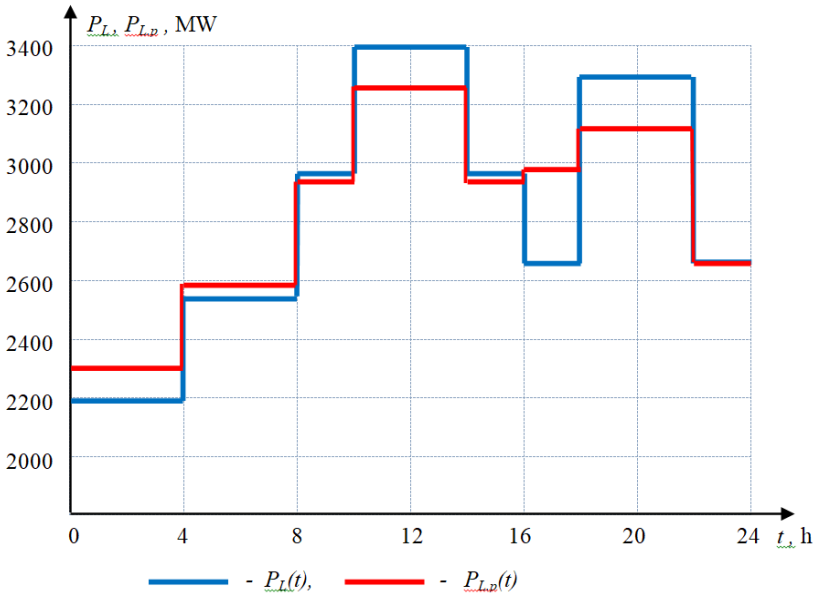


Figure 2: Daily schedules of initial and calculated loads.

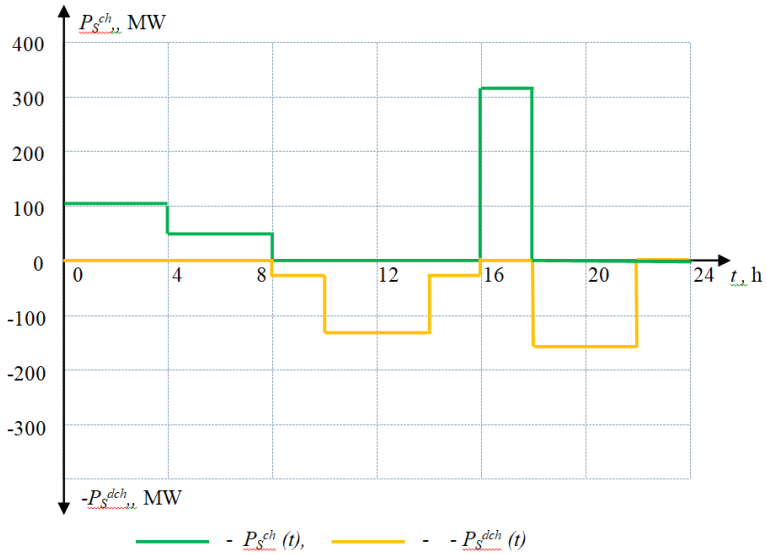


Figure 3: Daily schedules of charging and discharging the energy storage device.

The reliability of the results obtained is confirmed by comparing them with numerous results obtained based on a simple selection.

In the example considered, due to the regulating capabilities of energy storage devices, the daily load schedule of the power system is significantly leveled. In particular, the difference between the maximum and minimum total loads decreases from 1199 MW to 960 MW.

Analysis of the results shows that the use of the proposed mathematical model and optimization algorithm when planning short-term modes of power systems containing energy storage devices makes it possible to significantly align load schedules. This leads to a corresponding reduction in costs associated with the production of electricity at power plants due to the improvement of their operating modes, with the loss of electricity in networks due to their unloading. In addition, the regulatory capability of energy storage devices creates favorable conditions for the efficient use of power plants operating on renewable energy resources in the energy system.

4 CONCLUSIONS

The study addresses the optimization of short-term operating modes in electric power systems with energy storage devices. Based on the conducted research, the following key conclusions have been drawn:

1) A mathematical model for the problem of optimization of the modes of electric power systems containing energy storage devices when planning their short-term modes has been proposed.
2) An algorithm for leveling the consumer load schedules in electric power systems when optimizing their modes using the regulatory capabilities of energy storage devices is given.
3) Based on the research carried out using the example of the central part of the main electrical networks of the “National electric grid of Uzbekistan” JSC it was established that the use of the proposed model and algorithm makes it possible to significantly equalize load schedules through the use of the regulating capabilities of energy storage devices. This leads to a corresponding reduction in costs as a result of unloading elements of the electrical network and improving operating conditions of power plants.
4) The proposed mathematical model and optimization algorithm can be used when planning short-term modes of electric power systems containing energy storage devices.

REFERENCES

[1] M.M. Rana, M. Uddin, M.R. Sarkar, S.T. Meraj, G.M. Shafiullah, S.M. Muyeen, M.A. Islam, and T. Jamal, “Applications of energy storage systems in power grids with and without renewable energy integration-a comprehensive review,” J. Energy Storage, vol. 68, 2023. doi: 10.1016/j.est.2023.107811.

[2] Z. Zhang, T. Ding, Q. Zhou, Y. Sun, M. Qu, Z. Zeng, Y. Ju, L. Li, K. Wang, and F. Chi, "A review of technologies and applications on versatile energy storage systems," Renew. Sustain. Energy Rev., vol. 148, 2021. doi: 10.1016/j.rser.2021.111263.

[3] E. Hossain, H.M.R. Faruque, M.S.H. Sunny, N. Mohammad, and N.A. Nawar, "Comprehensive review on energy storage systems: Types, comparison, current scenario, applications, barriers, and potential solutions, policies, and future prospects," Energies, vol. 13, 2020, p. 3651. doi: 10.3390/en13143651.

[4] R. Byrne, T. Nguyen, D. Copp, B. Chalamala, and I. Gyuk, "Energy management and optimization methods for grid energy storage systems," IEEE Access, 2017, pp. 1–1. doi: 10.1109/ACCESS.2017.2741578.

[5] E.W. Schaefer, G. Hoogsteen, J.L. Hurink, and R.P. van Leeuwen, "Sizing energy storage in electricity grids containing flexible loads," J. Energy Storage, vol. 97, Part A, 2024. doi: 10.1016/j.est.2024.112706.

[6] M.S. Reza et al., "Optimal algorithms for energy storage systems in microgrid applications: An analytical evaluation towards future directions," IEEE Access, vol. 10, 2022, pp. 10105–10123. doi: 10.1109/ACCESS.2022.3144930.

[7] L. Zeng, Y. Gong, H. Xiao, T. Chen, W. Gao, J. Liang, and S. Peng, "Research on interval optimization of power system considering shared energy storage and demand response," J. Energy Storage, vol. 86, Part B, 2024. doi: 10.1016/j.est.2024.111273.

[8] D. Osorio, "A review in BESS optimization for power systems," TecnoLógicas, vol. 26, 2022. doi: 10.22430/22565337.2426.

[9] N. Altin, "Energy storage systems and power system stability," Proc. ISGWCP, 2016, pp. 1–7. doi: 10.1109/ISGWCP.2016.7548268.

[10] M. Tabasi and P. Asgharian, "Optimal operation of energy storage units in distributed system using social spider optimization algorithm," AIMS Electron. Electr. Eng., vol. 3, no. 4, 2019, pp. 309–327. doi: 10.3934/ElectrEng.2019.4.309.

[11] Z. Yang, F. Yang, H. Min, H. Tian, W. Hu, and J. Liu, "Review on optimal planning of new power systems with distributed generations and electric vehicles," Energy Reports, vol. 9, 2023, pp. 501–509. doi: 10.1016/j.egyr.2022.11.168.

[12] M.R. Sheibani et al., "Energy storage system expansion planning in power systems: a review," IET Renew. Power Gener., vol. 12, 2018, pp. 1203–1221.

[13] P. Balducci, K. Mongird, and M. Weimar, "Understanding the value of energy storage for power system reliability and resilience applications," Curr. Sustain. Renew. Energy Rep., vol. 8, 2021, pp. 131–137. doi: 10.1007/s40518-021-00183-7.

[14] F.N. Tsany, A.A. Widayat, D.R. Aryani, F.H. Jufri, and I.M. Ardita, "Power system stability improvement using battery energy storage system (BESS) in isolated grid," IOP Conf. Ser.: Earth Environ. Sci. doi: 10.1088/1755-1315/599/1/012025.

[15] T. Gayibov and E. Abdullaev, "Optimization of daily operation mode of photovoltaic systems of enterprises," E3S Web Conf., vol. 264, 2021, p. 04063. doi: 10.1051/e3sconf/202126404063.

[16] T.Sh. Gayibov, B.A. Uzakov, and E.A. Abdullaev, "Optimization of loading schedules of consumers with own stations based on renewable energy sources," J. Crit. Rev., vol. 7, no. 15, 2020, pp. 1738–1742.

[17] K.R. Bakhteev, "Increasing the efficiency of functioning of centralized and autonomous power supply systems through the integrated use of electrochemical energy storage devices, small-scale generation and forcing the excitation of synchronous machines," Ph.D. dissertation, 2019, 190 p. (In Russian).

[18] V.Yu. Kononenko, O.V. Veschunov, V.P. Bilashenko, and D.O. Smolentsev, "Effects of using energy storage devices in isolated power systems of Russia," Aktika: Ecology and Economics, vol. 2, no. 14, 2014, pp. 61–66. (In Russian).

[19] V.A. Karasevich, "Use of energy storage systems for energy storage in autonomous power systems," 2023, no. 7, pp. 68–70. (In Russian).

[20] T. Gayibov, B. Uzakov, and A. Shanazarov, "Algorithm of power system mode optimization taking into account losses in networks and functional constraints," AIP Conf. Proc., vol. 2612, 2023, p. 050011. doi: 10.1063/5.0117667.

Study of the Energy Efficiency of a Thermal Electrical Generator with a Hydraulic Heat Supply System

Oleg Ivanov
[1]*Poltava state agrarian university, Skovorody Str. 1/3, 36001 Poltava, Ukraine*
oleg.ivanov@pdau.edu.ua

Keywords: Thermoelectric Converter, Seebeck Effect, Power, Generator, Heat, Heat Carrier.

Abstract: The article presents the results of a study of an alternative option for generating electricity as an approach to reducing dependence on traditional methods of generating electricity. We have proposed a variant of a thermal electric generator containing a thermoelectric converter with a hydraulic system for a forced supply of cold and hot heat carriers. The research aimed to determine the thermal generator's efficiency under different variants of external load and temperature differences on the thermal sides of the thermoelectric converter. The load was selected from a variable range: 0, 0.51, 1.8, 5.1, 15 Ω. The temperature range was from 10 to 45°C. Semiconductor sensor technology was used to monitor the temperature, electrical and thermal power parameters of a thermal power generator. The study of the generator operation without load revealed a linear change in the converter's thermoelectromotive force from 600 to 2600 mV. With an external load of 15 Ω, we managed to achieve the best performance in generating the output voltage, which varied linearly from 600 to 2500 mV. In the course of comparative studies, the highest generator power was recorded at an external resistance of 1.8 Ω, which varied from 0.05 to 0.43 W depending on the temperature difference. At this load value, the best energy conversion rates were also achieved, when the efficiency coefficient varied between 0.3 and 0.55%.

1 INTRODUCTION

Today, the rapid growth in electricity consumption is outstripping the pace of its generation [1]. This significantly exacerbates the problem of growing food shortages and increases the cost of consumption [2]. This need becomes even more acute without physical access to power lines in remote areas and territories [3].

A partial solution to these problems may be the use of alternative ways of generating electricity instead of traditional ways of generating it, in particular, the use of various renewable energy sources [4].

When choosing a certain alternative technology for generating electricity, questions are always raised about their efficiency, cost, technical possibilities of their application, availability of available energy sources for their conversion into an electrical equivalent, etc [5].

Thermal energy is the type of energy that is most used for various technical and technological processes. It is of particular importance in the field of electricity generation [6]. Most power-generating capacities operate on thermal energy, the production of which is associated with potential risks of environmental degradation [7]. In particular, excessive emissions of unused heat into the atmosphere.

In addition, vehicles, production facilities, household consumers, etc. are significant sources of residual heat emissions.

It is advisable to study the technology and technical means of using residual heat to generate electricity.

A well-known approach for the direct energy conversion of heat into electricity is the use of thermoelectric converters based on the Seebeck effect. Based on this conversion method, various installations have been developed to generate electricity using geothermal sources [8], vehicle exhaust gas energy [9], utilisation of heat from air conditioning units [10], human body energy [11], etc.

Thermoelectric generators are small, lightweight and reliable energy converters. They operate without noise and vibration due to the absence of any moving mechanical parts [12].

However, these generators are characterised by low conversion efficiency and high cost. This has limited their widespread use [13]. Thus, research on thermoelectric generators in recent years has focused on optimising their performance and reducing their costs.

2 MATERIALS AND METHODS

The study aims to determine the efficiency of a thermal electric generator with a hydraulic heat supply system under various external loads.

A schematic diagram of the experimental thermal electric generator is shown in Figure 1, and a photographic image is shown in Figure 2.

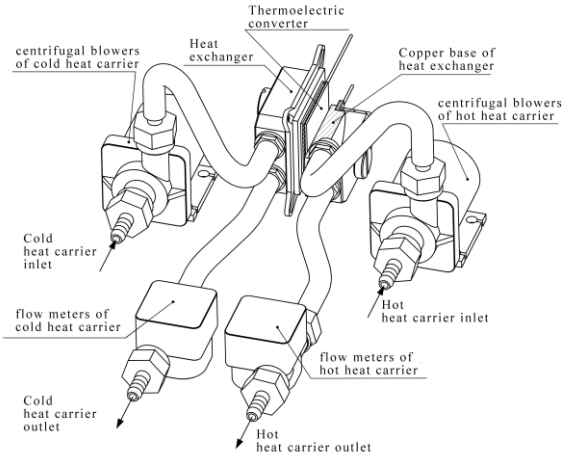


Figure 1: General scheme of a thermal electric generator.

The main component of a thermal electric generator is a thermoelectric converter based on the Seebeck effect. For a thermal electric generator to function, two heat sources with different temperature potentials are needed. Cold and hot liquid carriers are used as such sources.

Regenerative heat exchangers are used to ensure heat exchange between the heat carriers and the thermoelectric converter, which are mounted on different thermal sides of the converter.

The heat carriers in the heat exchangers are circulated by centrifugal blowers, the performance of which is controlled by a PWM controller.

DS18b20 digital temperature sensors are used in the heat exchangers to determine the change in temperature of the heat carriers before and after heat exchange, which are mounted in the heat carrier pipeline at the inlet and outlet of the heat exchangers.

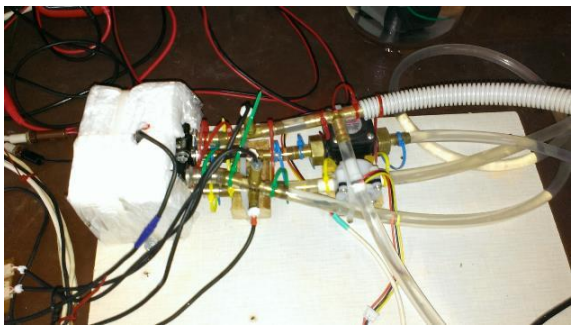


Figure 2: Photographic image of the experimental thermal electrical generator.

The volumetric flow rate of heat carriers is monitored by contact method using SEN-HZ21WA and YF-S401 volumetric flow meters with different measuring ranges.

The source of hot and cold heat carriers is calorimetric tanks.

One tank contains a thermistor heater to heat the hot carrier to the required temperature, which is controlled by an electronic thermostat. In the other tank, cold storage batteries are used to form a cold carrier with a temperature determined by the conditions of the experiment.

To monitor the temperature of the heat exchange surfaces of the thermoelectric converter, resistive temperature sensors were used, which were in direct contact with the surfaces.

A digital ammeter and a voltmeter were used to measure the electrical parameters of the thermoelectric generator.

A special microcontroller system based on the 8-bit RISK microcontroller Atmega328 is used to process signals from digital and analogue sensors.

This microcontroller system collects, stores and processes information from sensor technology and then sends it to a smart device via Bluetooth wireless communication.

The microcontroller system also displays the collected information on the LCD and stores it in the file space of the SD memory card.

The general view of the workplace for the experiments is shown in Figure 3.

To achieve this aim, a research programme was planned, which included two stages.

In the first stage, the thermal electric generator was tested without connecting an external load.

At this stage, the magnitude of the thermomotive force was measured at a step change in the temperature difference on the heat-exchanging sides of the thermoelectric converter. The temperature on

its heat-absorbing side remained constant at 10°C throughout the experiment.

Figure 3: Workplace for the study of a thermal power generator.

The required temperature difference was created by changing the temperature on the heat-absorbing side by changing the temperature of the heat carrier in the regenerative heat exchanger, which is fixed on this side.

The increase in the temperature of the heat carrier was provided by Joule heat from a thermoelectric heater of the appropriate tank capacity.

The range of temperature difference was from 20 to 45°C.

The second stage involves conducting a series of experiments to build a characteristic of changes in the output electrical parameters of the thermal electrical generator and the level of energy efficiency as a function of the external load and the temperature difference on the heat exchange sides of the thermoelectric converter.

The methodology for the second stage of experimental research was to measure the output electrical parameters of the generator when the external load varies according to a variable series: 0.51 Ohm, 1.8 Ohm, 5.1 Ohm, and 15 Ohm.

The electrical values were measured for each external load value at different temperature differences on the heat exchange sides of the converter from 10 to 45°C.

The constant temperature on the heat-absorbing side of the thermoelectric converter was a prerequisite for the experiments.

The efficiency of a thermal electric generator is determined by comparing the output power received with the amount of heat removed on the heat-absorbing side of the converter.

The amount of heat consumed was determined by calculating the heat loss of the hot heat carrier in the heat recovery exchanger.

Heat losses were calculated based on the change in the temperature of the heat transfer medium at the inlet and outlet of the heat exchanger and its current flow rate.

3 RESULTS AND DISCUSSIONS

Figure 4 shows the change in the magnitude of the thermoelectromotive force from the temperature difference on the heat exchange sides of the thermoelectric converter in the absence of an external load.

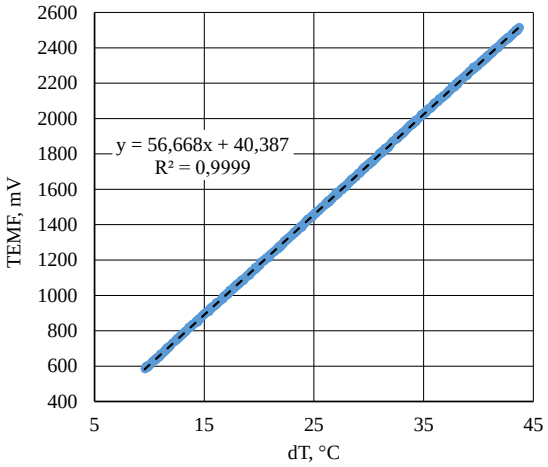


Figure 4: Dependence of the thermoelectromotive force of a thermoelectric converter on the temperature difference dT.

The presented dependence demonstrates a clear linear dependence of the thermoelectromotive force on the temperature difference.

Thus, in the range of temperature changes from 10 to 45 °C, the thermoelectromotive force varies linearly with increasing value from 600 to 2500 mV.

Based on the regression analysis of the research data, the following equation was derived:

$$y = 56.668 \cdot dT + 40.387 . \quad (1)$$

The degree of reliability of R2 of the given approximated equation to the nature of the change in the thermoelectromotive force is as close as possible to one.

This gives a good agreement between theoretical and experimental values of the thermoelectromotive force.

The main output electrical parameters that characterise the operation of any electric generator are the output voltage U, the electric current I and the power W.

The formation of all electrical parameters for a thermal electric generator is influenced by two determining factors: the temperature difference dT on the heat exchange surfaces of the thermoelectric converter and the external load $R_{external}$.

The first factor determines and shapes the width of the output voltage range, and the second factor sets the amount of current that will flow through the closed electrical circuit of a thermal power generator. Therefore, both factors are taken into account when studying electrical parameters.

Figure 5 shows the results of an experimental study of changes in the output electrical parameters of a thermal electric generator when the temperature difference varies within 10-45°C with a connected consumer of electrical energy with a resistance of 0.51 Ohm.

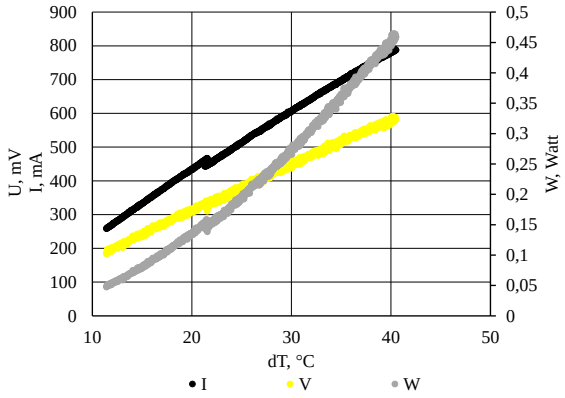


Figure 5: Changes in electrical parameters due to temperature difference dT at an external load resistance of 0.51 Ohm.

From the analysis of the pattern of changes in the output voltage and current in the electrical circuit, it is clear that these electrical parameters change in direct proportion to the value of the temperature difference dT. Their change obeys a linear law and can be described by the approximated equations of the straight line:

- for voltage:

$$U = 13.591 \cdot dT + 38.954\,; \qquad (2)$$

- for current:

$$I = 18.071 \cdot dT + 62.807\,. \qquad (3)$$

In the studied range of temperature differences on the heat exchange surfaces of the converter, the current in the circuit varies from 258 to 788 mA.

The voltage generated by the temperature difference and the level of external load on the generator is in the range of 185-586 mV.

The electrical power level of a thermal electric generator varies in a non-linear way and ranges from 0.05 to 0.46 W. The pattern of changes in the value of electric power can be described with a high degree of confidence by a regression equation:

$$W = 0.0002 \cdot dT^2 + 0.028 \cdot dT - 0.0118 \qquad (4)$$

To conduct a comparative analysis of the electrical characteristics of a thermal electric generator at other values of external load resistance, it is necessary to reduce them to common coordinate systems.

Figure 6 shows the combined characteristics of the output voltage change.

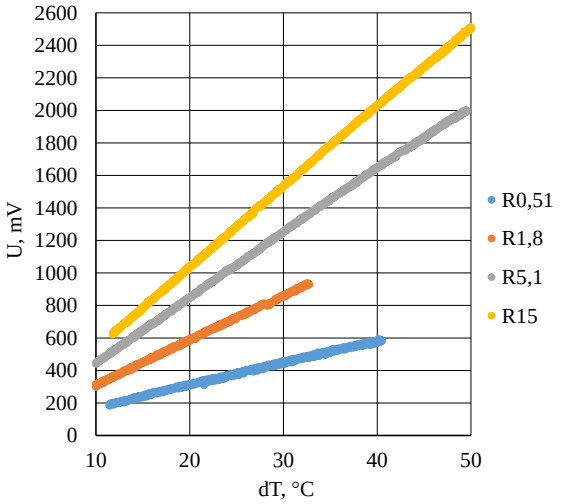


Figure 6: Comparative characteristics of electrical voltage under different external loads and temperature changes dT.

Comparison of the output voltage curves of the electric generator clearly shows that with an increase in the load level, a voltage drop is observed and its growth rate slows down as the temperature difference on the heat-exchanging sides of the thermoelectric converter increases.

Thus, when comparing the voltage levels at the maximum (0.51 Ohm) and minimum load (15 Ohm), the difference between them will be 460 mV at the minimum temperature difference with a gradual increase in this difference to 895 mV. At the same time, the voltage growth rate for the variant with a minimum load is 49 mV for each degree of temperature difference, as opposed to 13.6 mV/°C at maximum load.

For intermediate load values, the range of voltage deviation from the voltage at the minimum load within the same temperature difference is for 1.8 Ω - 264...748 mV; for 5.1 Ω - 116...274 mV.

The narrowing of the width of the voltage change ranges as the load decreases is reflected in the convergence of the voltage curves and the increase in their steepness. This helps to achieve a greater effect in generating electricity at the same temperature differences on the heat exchange surfaces of the thermoelectric converter.

Somewhat different conclusions can be drawn about the electrical characteristics of a thermal electric generator when comparing its output power under different load conditions (Figure 7).

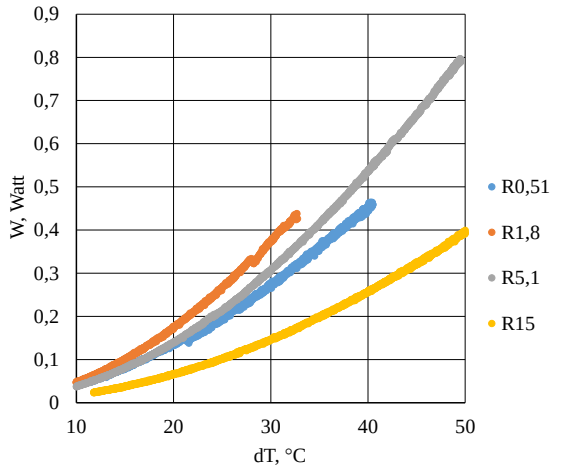


Figure 7: Comparative characteristics of electrical power at different external loads and temperature differences dT.

Thus, based on the comparison of the output power curves in Figure 7, it should be noted that the lowest power is provided by the generator when a consumer with an electrical resistance of 15 Ω is connected to it. The power level developed within the limits of dT change from 10 to 45°C is 0.023...0.256 W. This is almost two times less than the power provided by the generator when connected to a consumer with a resistance of 0.51 Ω over the entire range of dT. To supply this consumer, the generator, depending on dT, develops power in the range of 0.05...0.45 W.

In turn, for average load values, the generator power takes on higher values, in particular, for a load of 5.1 Ω at dT=20°C, the power is 0.14 W, for 1.8 Ω - 0.175 W.

A further increase in dT increases the power value, and already at dT=30°C for a 5.1-ohm load it reaches 0.31 W, and for 1.8 Ω - 0.38 W. This power level for the other load variants is only achievable at higher dT values: at 0.51 Ω, a power of 0.38 W is only achievable at dT=36°C, at 15 Ω - at dT=49°C.

Taking into account the relative position of the generator output power curves at different loads, it can be concluded that the highest power values and the highest growth rates can be achieved when using consumers with an average electrical resistance level between 1.8 and 5.1 Ω.

At the same time, the output power increases at a more significant rate as the electrical resistance of the consumer approaches the lower limit of the proposed range of electrical resistances.

Of particular interest is the level of thermal energy consumption in the production of electricity by a thermal electric generator and the degree of its efficient use.

Figure 8 shows the summary results of the level of heat energy consumption depending on the value of the electrical power output at different loads.

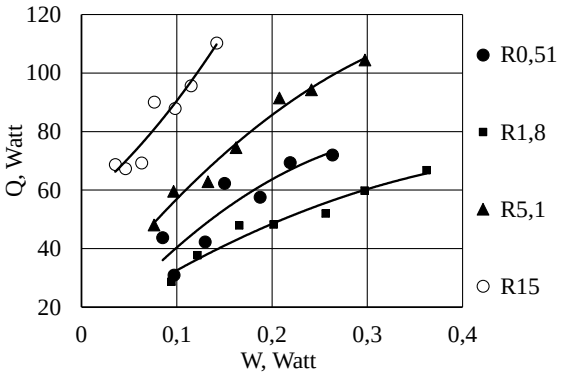


Figure 8: Comparative characteristics of heat consumption under different external loads and electrical power.

Figure 9 illustrates the character of the change in the coefficient of thermal energy use efficiency depending on the level of power generated by a thermal power generator when supplying different consumers.

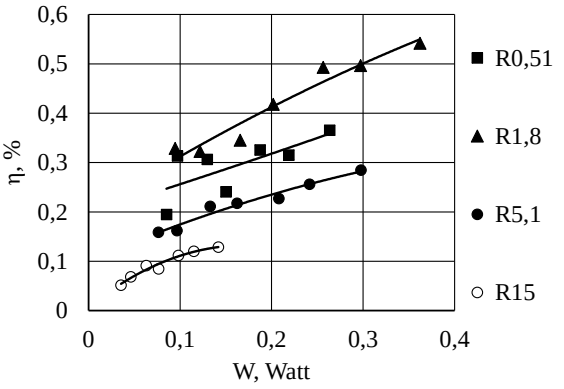


Figure 9: Comparative characteristics of the coefficient of thermal energy efficiency depending on the power level.

According to the dependencies in Figure 8, it can be seen that as the electric power increases for all variants of external loads, there is a steady increase in heat consumption. Within the framework of the experiments, the range of heat consumption is from 30 to 112 W.

The most significant level of thermal energy consumption is observed when powering a consumer with an electrical resistance of 15 Ohm, while to ensure a power level within 0.05...0.15 W, it is necessary to consume thermal energy in the amount of 67...112 W. These limits of heat consumption are the highest when ensuring identical levels of electricity use by different consumers over the entire range of output power produced by the generator.

The least amount of heat is used by a thermal electric generator when supplying a consumer with an electrical resistance of 1.8 Ω. For example, to meet the electricity needs of a given type of consumer in the range of 0.1 to 0.37 W, it is necessary to use 30 to 67 W of heat.

When consumers with an electrical resistance other than 1.8 Ω are connected to the generator, both upward and downward, the level of heat consumption increases. At the same time, a decrease in external resistance increases the amount of electricity consumption to a lesser extent.

Thus, when comparing the levels of heat consumption for resistances of 1.8 Ω and 0.51 Ω with the same generator power, the amount of heat used for a consumer with a lower resistance is 10-30 W more. This difference gradually increases as the need to generate more and more electricity increases.

As for a consumer with an electrical resistance of 5.1 Ω, the amount of thermal energy required to generate electricity in the range of 0.07 to 0.3 W increases by an average of 25...40 W compared to a consumer with the lowest level of heat consumption.

The efficiency of converting heat energy into electricity can be determined by the efficiency factor, which can be calculated as the ratio of the amount of electricity produced to the amount of heat energy consumed.

Let's consider how this coefficient changes for a thermal electric generator with different values of output power and external electrical resistance.

According to the graphical representation in Figure 9, it should be noted that the thermoelectric conversion of thermal energy into its electrical equivalent is rather low. Based on the data presented, the efficiency ratio ranges from 0.05 to 0.55%.

Despite the rather low level of the efficiency coefficient, the pattern of its change and the range of fluctuations for different variants of external load and generated power have different features.

The lowest values of the efficiency coefficient in the entire power range are achieved by a thermal electric generator with a connected consumer with a resistance of 15 Ω. In this variant of the generator's operation, with a power variation from 0.03 to 0.14 W, the efficiency factor is constantly growing, increasing by two and a half times from 0.05 to 0.13%. The character of its change is close to parabolic.

The best thermal energy efficiency is achieved by an electric generator when powered by an external consumer with a resistance of 1.8 Ω. When the power level is changed from 0.1 to 0.37 W, the thermal electric generator has a linearly increasing efficiency factor that is in the range of 0.3...0.55 %.

This high efficiency of heat energy use is fully consistent with the lowest heat consumption and the best power generation properties compared to other generator operating options. This was noted earlier.

For all other consumers, the efficiency factor has the same tendency to change as the amount of heat consumed by the power generator.

In particular, with an external resistance of 0.51 Ω, the efficiency factor takes values from 0.19 to 0.37%, which is 10-20% less when compared to a resistance of 1.8 Ω. Increasing the external resistance to 5.1 Ω reduces the efficiency coefficient and forms its change range from 0.16 to 0.285%.

3 CONCLUSIONS

A comprehensive analysis of energy costs and efficiency of their use showed a rather low level of thermoelectric conversion of thermal energy into electricity for this version of the electric generator. Its efficiency factor does not exceed 1%.

The practical use of this type of generator is possible only if there is free or residual thermal energy that is not generated specifically for the operation of this generator. Or in cases where there are no other ways to generate electricity.

A significant power generation capacity using thermoelectric converters is possible only with a radical increase in their number and the creation of favourable conditions for intensifying the processes of heat emission and heat absorption.

REFERENCES

[1] X. Xue, "Impact of electricity shortages on productivity: Evidence from manufacturing industries," Energy Eng., vol. 118, no. 4, pp. 995–1008, 2021. doi: 10.32604/ee.2021.014613.

[2] S. Pront-van Bommel, "A reasonable price for electricity," J. Consum. Policy, vol. 39, no. 2, pp. 141–158, Feb. 2016. doi: 10.1007/s10603-015-9300-x.

[3] N. Blair, D. Pons, and S. Krumdieck, "Electrification in remote communities: Assessing the value of electricity using a community action research approach in Kabakaburi, Guyana," Sustainability, vol. 11, no. 9, p. 2566, May 2019. doi: 10.3390/su11092566.

[4] T.-Z. Ang, M. Salem, M. Kamarol, H. S. Das, M. A. Nazari, and N. Prabaharan, "A comprehensive study of renewable energy sources: Classifications, challenges and suggestions," Energy Strategy Rev., vol. 43, p. 100939, Sep. 2022. doi: 10.1016/j.esr.2022.100939.

[5] S. Ahmad and R. M. Tahar, "Selection of renewable energy sources for sustainable development of electricity generation system using analytic hierarchy process: A case of Malaysia," Renew. Energy, vol. 63, pp. 458–466, Mar. 2014. doi: 10.1016/j.renene.2013.10.001.

[6] A. Kundu, A. Kumar, N. Dutt, C. S. Meena, and V. P. Singh, "Introduction to thermal energy resources and their smart applications," in Thermal Energy Systems, CRC Press, 2023, pp. 1–15. doi: 10.1201/9781003395768-1.

[7] Q. Bian, "Waste heat: The dominating root cause of current global warming," Environ. Syst. Res., vol. 9, p. 8, 2020. doi: 10.1186/s40068-020-00169-2.

[8] J.-L. Zeng, X. Liu, C.-Q. Su, Y.-P. Wang, and X. Xiong, "Numerical investigation of a geothermal thermoelectric generator using gravity heat pipe structure," Energy Rep., vol. 9, pp. 1237–1246, Oct. 2023. doi: 10.1016/j.egyr.2023.05.188.

[9] Q. Cao, W. Luan, and T. Wang, "Performance enhancement of heat pipes assisted thermoelectric generator for automobile exhaust heat recovery," Appl. Therm. Eng., vol. 130, pp. 1472–1479, 2018. doi: 10.1016/j.applthermaleng.2017.09.134.

[10] Y. Mona, C. Chaichana, K. Rattanamongkhonkun, and S. Thiangchanta, "Energy harvesting from air conditioners by using a thermoelectric application," Energy Rep., vol. 8, pp. 456–462, Nov. 2022. doi: 10.1016/j.egyr.2022.09.178.

[11] Z. Tabaie and A. Omidvar, "Human body heat-driven thermoelectric generators as a sustainable power supply for wearable electronic devices: Recent advances, challenges, and future perspectives," Heliyon, vol. 9, no. 4, p. e14707, Apr. 2023. doi: 10.1016/j.heliyon.2023.e14707.

[12] D. Champier, "Thermoelectric generators: A review of applications," Energy Convers. Manag., vol. 140, pp. 167–181, May 2017. doi: 10.1016/j.enconman.2017.02.070.

[13] J. Xiao, T. Yang, P. Li, P. Zhai, and Q. Zhang, "Thermal design and management for performance optimization of a solar thermoelectric generator," Appl. Energy, vol. 93, pp. 33–38, May 2012. doi: 10.1016/j.apenergy.2011.06.006.

Annual Monitoring Analysis of Four On-Grid Stations in Termez

Boysori Yuldoshov[1], Maxamadi Chariyev[1], Azamat Narbayev[1], Jasur Khaliyarov[1], Sirojiddin Toshpulatov[1], Sojida Raupova[2] and Tulkin Buzrukov[3]
[1]*Termez State University, Barkamol Avlod Str. 43, 190111 Termez, Uzbekistan*
[2]*Termiz State Pedagogical Institute, At-Termiziy Str. 1A, 190110 Termez, Uzbekistan*
[3]*Termiz University of Economics and Service, Farovon Str. 4B, 190111 Termez, Uzbekistan*
b.yuldoshov10@mail.ru, chariyevmaxamadi@gmail.com,narbaev@ro.ru, xjxjasur@mail.ru, toshpulatovs@tersu.uz, sojidaraupova38@gmail.com, tolqinbuzrukov@gmail.com

Keywords: CUF, Renewable Energy, Solar System, Solar Station, Monitoring, Energy.

Abstract: In the article, a one-year energy analysis of 4 on-grid solar stations with capacities of 49.5kW, 53.76kW, 39.6kW and 52.8kW installed on the roof of the building was carried out. It is important to know the efficiency of the use of such solar plants in the world, gradually abandoning the amount of traditional fuel and switching to green energy. Payback periods and CUF coefficients of the station were calculated using the annual energy production. An average payback period of 4.24 years was found for the stations. The best result for CUF belongs to IV station equal to 17.4%, and the smallest value is equal to 8.21% for II station. The annual amounts of energy produced were 58.79MWh, 38.76MWh, 50.05MWh and 80.67MWh, respectively. Despite the high temperature and pollution, the stations had their best energy production in June. In the future, such solar plants on the roof are considered promising projects, because they do not require additional space and are the reason for the cheapest initial investment.

1 INTRODUCTION

The rapid development of solar energy technologies has led to the widespread installation of photovoltaic (PV) power plants worldwide. In Uzbekistan, the adoption of on-grid rooftop solar stations is increasing as part of the national goal to achieve 7 GW of installed solar capacity by 2030. However, to maximize the efficiency of these systems, it is essential to analyze their performance metrics, including the Capacity Utilization Factor (CUF), energy output, and payback period.

This study focuses on the performance analysis of four rooftop on-grid solar power plants installed at Termez State University. The aim is to evaluate their annual energy production, compare CUF values, and identify key factors affecting system efficiency. By utilizing real-time monitoring data and performance calculations, this study provides insights into the operational characteristics of rooftop solar stations under high-temperature climatic conditions.

2 LITERATURE REVIEW

Among renewable energy sources, the installed capacity of PV modules is increasing year by year. The main reasons for this are the annual decrease in fuel energy reserves and the great importance given to green energy in the modern world. According to the sources of the International Renewable Energy Agency, in 2022, the worldwide installed capacity of photovoltaic modules exceeded 1053GW, and for our country Uzbekistan, this figure was 0.25GW, which is reflected in Figure 1 [1].

While China (393GW), the US (113GW), Japan (78.8GW) and Germany (66.5GW) have the best-installed capacity, China's relative share will be almost 37% of solar installations in 2022 [2]. Uzbekistan aims to have 7 GW of installed solar power by 2030, and many practical works and solar plants are being built for this purpose.

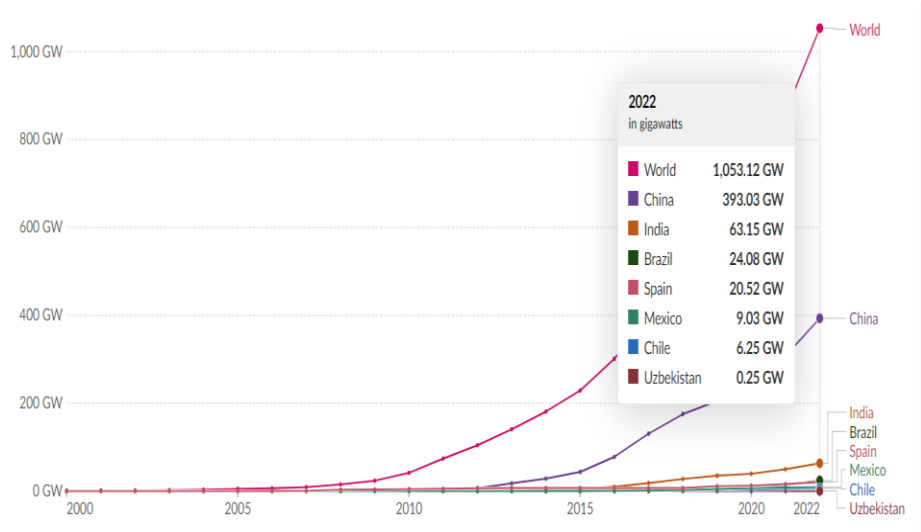


Figure 1: Installed solar power capacity from 2000 to 2022.

To effectively use the roofs of many buildings, rooftop on-grid stations were put into operation last season, and it is important to know the installed power utilization coefficients to evaluate the efficiency of these stations. It is also important to analyze annual, monthly, daily energy production. [3] in the work, in the analysis of this type of 70kW solar station (SS), the result for CUF is 10.15%. Muyiwa S. A., Emil E.T. analyzed the PV plant in Norway and obtained a value of 10.58% for the annual CUF [4]. [5-8] scientific works also states the station performance and the energy analysis are signifie important. The amount of energy produced, daily, monthly, annual monitoring and payback period of the station are important quantities that can be concluded about the station. The influence of radiation and temperature on PV modules [9] has been studied in scientific works, and high temperature has been shown as the main external factor affecting the open circuit voltage of the module.

3 METHODS AND MATERIALS

3.1 Study Location

In the article, an analysis was made based on the annual energy produced by 4 rooftop on-grid solar plants. Solar station in the study is located on the territory of Termez State University (longitude of 37° 13'57"N, latitude of 67° 17'8"E) (Figure 2). Termez is a city located in the southernmost part of the Republic of Uzbekistan, has high temperature and annually high irradiance.

Figure 2: Location of on-grid solar systems.

3.2 PV System Description and System Description

This system is installed on the roof of one of the university building facing south. 550W and 560W monocrystalline panels were used in the on-grid solar system, and their electrical characteristics are given in Table 1.

Table 1: Physical parameters of PV panel.

Model Type	YH550W-36MN	BSM560M10-72HPH
Peak Power (P_{max})	550W	560W
Module Efficiency	21.28%	21.68%
Maximum Power Voltage (V_{mp})	42.13V	42.33V
Maximum Power Current (I_{mp})	13.06A	13.23A
Open Circuit (V_{oc})	50.1V	50V
Short Circuit Current (I_{sc})	13.9A	14.14A
Maximum System Voltage Nominal	1500V	1500V

Information about the number and type of panels used for 4 stations, and the installed power is given in the second table. The panels were installed at an average tilt angle of 30 degrees.

Table 2: General information.

№	PV			Total power	Inverter power (Huawei)
	Type	Power	Quantity		
I	Ipvi Sola	550W	90	49.5kW	50kW
II	Blue Sun	560W	96	53.76kW	50kW
III	Ipvi Sola	550W	72	39.6kW	50kW
IV	Ipvi Sola	550W	96	52.8kW	50kW

3.3 On-Grid Solar System Description

In this solar system, solar panels generate DC electricity by absorbing sunlight, and a solar inverter converts this DC electricity into AC electricity, which can then be used directly at home or in businesses. If the system generates more power than is consumed, the surplus is fed into the main electrical grid through solar net metering. In this setup, the inverter is connected to the Internet via Wi-Fi, enabling monitoring of the daily energy produced by the system, the energy generated, and its consumption at any time of the day. The system utilizes a SUN2000-50KTL-M3 (50kW) inverter which are modern equipment for solar energy applications. Full reports of daily and monthly power produced by the system, as well as daily weather information, can be accessed through the site[1] using mobile phones or computers.

3.4 Capacity Utilization Factor

The CUF (Capacity Utilization Factor) is the ratio of the actual energy output of an AC system to the energy a PV system would produce if it ran at nominal power [10]. Another way to define it is by measuring how long an electrical system operates at full (100 percent) capacity. Coefficients calculated over short intervals can vary significantly, so the accuracy of weekly, monthly, or yearly calculations improves as the time period increases.

$$\eta_{CUF} = \frac{E_{yield}}{P_{PV,rated} \cdot Time_{Fixed}} , \quad (1)$$

where E_{yield} is SS's total energy produced for a fixed time. $P_{PV,rated}$ is the installed power of SS. $Time_{Fixed}$ is the exact time taken for SS to produce E_{yield} energy.

4 RESULTS AND DISCUSSION

The collected data from July 2023 to June 2024 (one year) used to study the photovoltaic plant's performance were carried out at the Termez State University. The installed capacity of each station is different, but close to each other in terms of value. Therefore, in addition to the amount of energy produced, it is possible to compare these stations with the size of CUF. The information about the annual amount of energy produced by the station is given in Table 3.

Table 3: Annual amount of energy produced.

Power	I 49.5kW [MWh]	II 53.76kW [MWh]	III 39.6kW [MWh]	IV 52.8kW [MWh]
July	7.49	4.95	4.81	9.09
August	1.49	5.16	5.14	7.95
September	3.67	1.88	4.46	7.17
October	4.79	0	4.23	6.43
November	4.03	0	3.34	4.96
December	3.95	0	2.54	4.15
January	4.09	2.7	2.87	4.5
February	4.58	3.21	3.04	4.69
March	4.99	4.22	3.8	6.24
April	7.1	5.16	4.65	7.42
May	8.3	5.8	5.55	8.92
June	8.34	5.68	5.62	9.15
Total Energy	58.79	38.76	50.05	80.67
n [year]	4.68	7.71	4.4	3.64

[1] https://region02eu5.fusionsolar.huawei.com

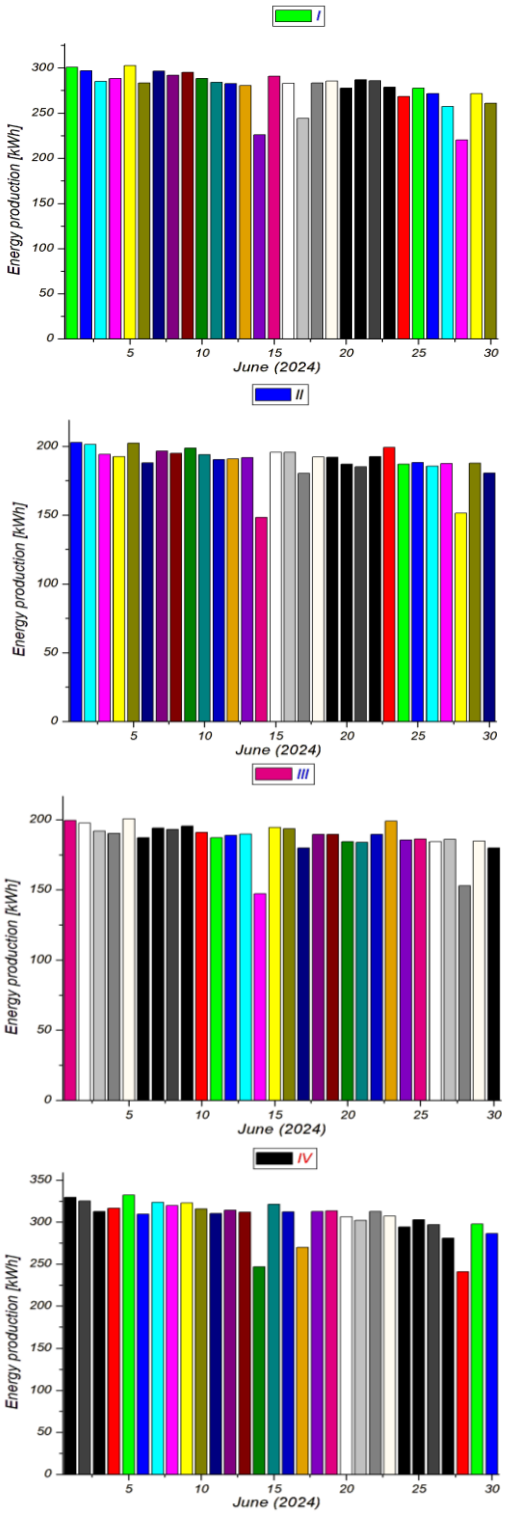


Figure 3: The energy production in June.

The best energy production belongs to the IV station and amounted to 80.67MWh. Station II produced the least amount of energy due to 3 months of technical interruptions. Due to the large impact of this technical outage on the station CUF and payback period, its value was not taken into account when calculating the average stations' payback period. Considering the evaluation of the economic payback period of the station with the energy produced by it, the economic analysis of SS, as follows, was carried out.

$$n = \frac{P \cdot \$400}{E \cdot \$0.072} \qquad (2)$$

Here, P is the installed power, $400 is the average price of a 1kW on grid station, E - annual amount of energy produced, $0.072 - this is the price of 1kWh energy in the country. In the calculations, we did not take into account the increase in energy prices and inflation, and we also did not take into account the costs associated with the operation of solar plant devices. The payback period of the solar station was 4.68, 7.71, 4.4 and 3.64 years for each station, respectively. Here, explaining the result of the second station as the absence of production for 3 months, we can get an overall average value of 4.24 years for such stations from the best results. The best production for the stations was in June. Data on daily energy production values for June are given in Figure 3.

The stations produced 8.34MWh, 5.68MWh, 5.62MWh and 9.15MWh of energy this month, respectively. The best indicator belongs to station IV, and the lack of this indicator in stations II and III can be explained by the formation of a shadow in the location of the panels. In the arrangement of the panels, it is assumed that the next row of arrays will be shaded at certain times of the day. Information about station CUF indicators is given in Figure 4.

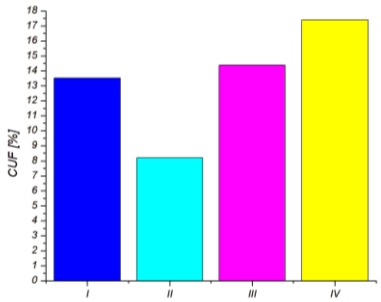


Figure 4: CUF in one year.

Since the CUF of the stations is a quantity that shows how efficiently they worked during the year, the smallest indicator of 8.21% belongs to station II. However, at station IV, this indicator is equal to 17.4%, which is one of the best indicators for SS in Termez conditions.

4 CONCLUSIONS

It is important to determine the efficiency of using the installed power of solar plants, to know the average payback period for a certain area. In the article, a one-year energy analysis of 4 on-grid solar stations with installed capacities of 49.5kW, 53.76kW, 39.6kW and 52.8kW in the Termez region was carried out. The annual amounts of energy produced were 58.79MWh, 38.76MWh, 50.05MWh and 80.67MWh, respectively. The average payback period was 4.24 years. Results of 13.52%, 8.21%, 14.38% and 17.4% were recorded for CUF, respectively. Despite the pollution and high temperature, the stations had their best energy production in June. In the future, such solar plants on the roof are considered promising projects, because they do not require additional space and are the reason for the cheapest initial investment.

REFERENCES

[1] IRENA, Renewable capacity statistics 2023. [Online]. Available: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2023/Mar/IRENA_-RE_Capacity_Highlights_2023.pdf

[2] H. Pourasl, R. Barenji, and V. Khojastehnezhad, “Solar energy status in the world: A comprehensive review,” Energy Rep., vol. 10, pp. 3474-3493, 2023.

[3] M. N. Tursunov, Kh. Sabirov, T. Z. Axtamov, U. B. Abdiyev, M. M. Chariyev, B. A. Yuldoshov, and S. F. Toshpulatov, “Capacity Utilization Factor (CUF) of the 70kW on-grid solar station in the dry climate of Termez,” E3S Web Conf., vol. 401, p. 02059, 2023. DOI: 10.1051/e3sconf/202340102059.

[4] M. Adaramola and E. Vågnes, “Preliminary assessment of a small-scale rooftop PV-grid tied in Norwegian climatic conditions,” Energy Convers. Manag., vol. 90, pp. 458-465, 2015.

[5] A. Haffaf, F. Lakdja, D. Abdeslam, and R. Meziane, “Monitoring measured and simulated performance analysis of a 2.4 kWp grid-connected PV system installed on the Mulhouse campus, France,” Energy Sustain. Dev., vol. 62, pp. 44-55, 2021.

[6] L. Zaghba, M. Khennane, A. Fezzani, A. Borni, and I. Mahammed, “Experimental outdoor performance evaluation of photovoltaic plant in a Sahara environment (Algerian desert),” Int. J. Ambient Energy, 2019.

[7] L. Zaghba, M. Khennane, A. Fezzani, I. Mahammed, and A. Borni, “Performance evaluation of 2.25 kWp rooftop solar PV plant based on experimental measurements in the desert environment, case study for Ghardaia, Algeria,” Int. J. Power Electron., vol. 13, no. 2, pp. 133-150, 2021.

[8] R. Chaudhari, B. Chaudhari, P. Chavda, and V. Leela, “To study the temporal variation of capacity utilization factor (CUF) of PV based solar power plant with respect to climatic condition,” Curr. World Environ., vol. 11, no. 2, p. 654, 2016.

[9] L. Musanga, W. Barasa, and M. Maxwell, “The effect of irradiance and temperature on the performance of monocrystalline silicon solar module in Kakamega,” Phys. Sci. Int. J., vol. 19, no. 4, pp. 1-9, 2018.

[10] A. Behura, A. Kumar, K. Rajak, C. Pruncu, and L. Lamberti, “Towards better performances for a novel rooftop solar PV system,” Sol. Energy, vol. 216, pp. 518-529, 2021.

Comparative Analysis of LSTM-Based PV Power Forecasting Models with Climate-Adaptive Feature Selection in Abuja, Nigeria

David Akpuluma[1,2], Wolf-Gerrit Früh[1], Neda Firoz[3], James Abam[4], Mohammed Umar Bello[1], Comfort Williams[1], Ambrose Onne Okpu[1]

[1]*Heriot–Watt University, Edinburgh, EH14 4AS, Scotland, United Kingdom*
[2]*Centre for Nuclear Energy Studies, University of Port Harcourt, East-West Road, PMB 5323, Port Harcourt, Nigeria*
[3]*Sam Higginbottom University of Agriculture Technology and Sciences, Allahabad, 211007 Uttar Pradesh, India*
[4]*Rivers State University, Nkpolu-Oroworukwo, PMB 5080, Port Harcourt, Nigeria*
{da2027, w.g.fruh}@hw.ac.uk, david.akpuluma@uniport.edu.ng, akpoebi@gmail.com

Keywords: PV Power Forecasting, LSTM Models, Climatic Feature Selection, Renewable Energy, Solar Energy Prediction, Abuja, Autocorrelation Analysis, Cross-Correlation Analysis, Sustainable Energy Solutions.

Abstract: In this research, we analyse how Long Short-Term Memory (LSTM) models can predict photovoltaic (PV) power output, in Abuja, Nigeria by selecting specific climate features and model configurations. The rising energy needs due to population growth and urbanisation emphasise the importance of sustainable energy sources. This study aims to improve the accuracy of PV power forecasts for integrating power into the current electrical grid and enhancing energy management strategies. By analysing data from the ERA5 dataset that includes various climatic features, we rigorously trained and assessed the LSTM models. Our results indicate that specific window sizes and combinations of features notably enhance forecasting accuracy with a window size of 6 and a mix of meteorological and solar radiation features showing the performance metrics (MAE, RMSE, R^2). The study also underscores the significance of autocorrelation and cross-correlation analyses in optimizing model setups. Our findings suggest that LSTM models can accurately predict PV power output offering insights for maximizing energy usage in urban areas with similar climates. This research contributes to efforts aimed at reducing reliance on fossil fuels and promoting sustainable energy solutions. Future endeavours will explore integrating real-time data and incorporating additional climatic features to further refine forecasting models.

1 INTRODUCTION

As population continues to grow, so does the energy demand, which has become a necessity to life in the 21st century. Like fuel to an automobile, so is energy to life, especially electrical power supply. Access to affordable, clean and reliable power supply is a sustainable development goal (SDG) of the United Nations [1]. Whilst other continents are making strides and seeking smoother transition, some West African countries seem to be lagging considerably with regards to electrical power infrastructure. Available infrastructure is outdated and requires an immediate and large-scale overhaul, which now is bereft of political will [2].

For several reasons generating power from fossil fuels is no longer sustainable, hence the need for renewable energy. In spite of the prospects of renewable energy, its reliability and integration to already existing infrastructure is dependent on a couple of factors, one of which is its availability owing to the fact that its generation is dependent on weather or meteorological factors [3, 4]. In tropical climates like Sub-Saharan Africa, renewable energy such as solar has a very high potential in addressing the electricity deficiency experienced in the region especially in the most populous African country, Nigeria. Off-grid solar homes systems (SHSs) and community microgrids seem to be the most effective and practical solution to accessing power supply for rural dwellers, with advantages of reduced cost, environmental sustainability and ease of deployment, especially when the centralized power infrastructure is not reliable and would be capital intensive for individuals without access to such funds [1].

The capital city of Nigeria, Abuja, is quite developed but not left out in electrical power infrastructure challenges the region faces. Unprecedented population growth and urbanization have resulted in an escalating need for dependable

and cost-effective electricity. Nevertheless, the current power infrastructure in Abuja is overburdened and experiences frequent interruptions, impacting economic activities and quality of life. Moreover, the utilization of fossil fuels for electricity generation is not environmentally sustainable due to ecological degradation and rising fuel costs. Renewable energy, specifically solar power, offers a feasible resolution to these challenges. Abuja, positioned in a tropical climate with ample sunshine, boasts a significant potential for harnessing solar energy. However, the successful integration of solar power into the current grid and its dependability are contingent upon precise forecasting of photovoltaic (PV) power production. Accurate PV power prediction can enhance the efficacy of solar energy systems, improve grid reliability, and enable more effective energy management approaches [5, 6].

Predicting power generation with a high level of accuracy, would enable power users in these regions without reliable power supply; to embrace this option of solar energy in photovoltaics as a reliable source of power. It can improve the efficiency of smart community microgrids where peer to peer trading of energy is possible.

Photovoltaic power generation is dependent on certain weather conditions like solar radiation intensity, temperature, wind speed and direction, cloud cover, humidity etc [4, 7]. Several models have been used in forecasting the PV power generation of renewable energy systems, some of which were traditional techniques or even hybrid models demonstrating excellent accuracy depending on the peculiarity of the prevailing conditions [7].

Long Short-Term Memory (LSTM) has been used for different time scales in predicting PV power output for several plants using climatic data. This deep learning method is robust and flexible. In some cases, providing accuracy of over 18% better than other benchmarked methods [8]. The functionality of any smart grid is dependent on the efficiency of the energy management technique employed. Energy management strategies are based on timescales (hourly, daily, weekly, monthly or yearly) depending on the purpose of its design. An RNN generates its output predictions from both current input and past data or experience, and where the distance between the cells is significant and vanishing gradient may tend to lose some information, LSTM solves this challenge by adding three gates (input gate, forget gate and output gate) to the RNN cell. So, it captures nonlinear relationships improving accuracy of the model [9]. In this study a Recurrent Neural Network (LSTM - Long Short-Term Memory) is employed to predict the power output of a PV system from climatic conditions.

The objectives of this paper are to:

- Demonstrate the possible outcome of PV power forecasting using LSTM with different climatic features for an urban city.
- Determine the most effective feature selection that enhances the accuracy of the LSTM model in forecasting PV power output in Abuja.

The importance of this investigation lies in its emphasis on refining PV power forecasting through sophisticated machine learning methodologies. By formulating and assessing Long Short-Term Memory (LSTM) models customized to Abuja's climatic conditions, this study strives to offer more precise forecasts of solar power generation. Subsequently, this can accelerate the acceptance of solar energy solutions, diminish reliance on fossil fuels, and contribute to sustainable development objectives in the region.

2 RELATED WORKS

In recent years, significant advancements have been made in the field of PV power forecasting, particularly with LSTM models and other machine learning techniques. In this overview we briefly discuss a range of research projects showcasing progress made so far, while also looking into specific areas that need improvement.

Machine Learning Model Optimization: For PV power prediction it is essential to choose relevant input features. Research highlights the significance of including weather conditions and time details to improve forecast accuracy [10, 11]. LSTM models, specifically tailored for time series data analysis, have shown superior performance in predicting power generation often surpassing other neural networks like GRU and MLP [12]. Moreover, methods such as XGBoost have also been successful in forecasting PV power output outperforming statistical methods, like SARIMA [13].

LSTM-Based Models for PV Power Forecasting: LSTM models have become quite popular for their ability to effectively handle the intermittent nature of solar irradiation and incorporate it into power grids. Studies have shown that LSTM models outperform Artificial Neural Network (ANN) models, especially when it comes to short term predictions [14]. Single variable LSTM models, which rely on PV output data demonstrate good accuracy in predicting one step ahead while multivariable models that consider weather factors excel in forecasting multiple steps ahead [15]. Additionally, research suggests that

LSTM models outshine Multi-Layer Perceptron (MLP) and Convolutional Neural Network (CNN) models when series data. According to a study by [16], employing two-week datasets produced the best outcomes. These methodologies indicate that LSTM based models are exceptionally proficient in providing dependable forecasts, for solar photovoltaic power across various time frames.

2.1 Gaps in Existing Research

Despite the progress made so far, some areas still need further investigation in the literature.

1) Selecting Features Adapted to Climate; While many studies stress the importance of choosing features, more research is required on selecting features that adapt to conditions especially tailored for regions like Abuja.
2) Ensuring Robustness Across Varied Climates; Most studies concentrate on weather patterns. Unique climate conditions in urban areas such as Abuja call for customized approaches to enhance forecasting accuracy.
3) Incorporating Local Data; Limited research exists on incorporating climate data with advanced machine learning algorithms for predicting PV power. This integration is vital for optimizing energy usage in areas with weather patterns.
4) Examining Window Sizes; Current studies often lack assessments of how different window sizes impact forecasting accuracy. This study aims to fill this void by examining the effects of window lengths on the precision of LSTM models in forecasting PV power.

By tackling these gaps this research strives to build models that utilize feature selection to climate and advanced machine learning methods customized for Abuja's specific climate conditions, in Nigeria.

3 METHODOLOGY

This section outlines the methodology employed in the study, to develop and evaluate an LSTM model for predicting surface solar radiation downwards (ssrd_5) using historical meteorological data from the ERA5 dataset, including data preprocessing, feature selection, and the development and evaluation of an LSTM model for time series forecasting.

3.1 Dataset Description

The data used in this study are sourced from the ERA5 reanalysis dataset, provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) through the Copernicus Climate Change Service (C3S) [17]. The ERA5 dataset offers a comprehensive reanalysis of global climate data, providing hourly estimates of various atmospheric, oceanic, and land surface parameters. The specific dataset used for this study is available at: ERA5 Reanalysis Dataset. The dataset covers the period from January 1, 2020, to December 31, 2022, with hourly time stamps. The dataset for the years 2023 and 2024 was not available at the time of the analysis.

This dataset provides valuable insights into the climatic conditions of the Abuja region over the specified period. It can be used for various analyses including wind energy potential assessment, solar radiation analysis, temperature and humidity studies, and cloud cover observation, which are crucial for understanding local weather patterns and for planning renewable energy projects.

The dataset includes the following meteorological variables measured at two different heights (10 meters and 100 meters), solar radiation parameters, temperature, dew point, surface pressure, and cloud cover. These variables are recorded for nine different locations in the vicinity of Abuja, Nigeria. The dataset was in csv file format which contained the time series data for the specified period and locations, with each variable labelled accordingly with the location suffix (_i). Table 1 shows variables and description of the dataset.

The data were collected for nine specific locations, each identified by a unique suffix (i where i ranges from 1 to 9). These locations are situated around Abuja, Nigeria, with the central point being at Latitude 9.0° N and Longitude 7.5° E, as shown in Fig. 1.

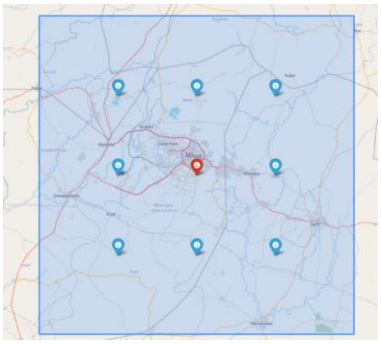


Figure 1: The geolocations of the study sites in Abuja, and a map of Nigeria.

Table 1: Data description.

Variable Type	Variable	Description
Meteorological Variables (Wind)	u10	10m_u_component_of_wind
	v10	10m_v_component_of_wind
	fg10	Wind gusts at 10m
	u100	100m_u_component_of_wind
	v100	100m_v_component_of_wind
Solar Radiation Variables	cdir	Clear sky direct solar radiation at surface
	fdir	Total sky direct solar radiation at surface
	ssrdc	Surface solar radiation downward clear sky
	ssrd	Surface solar radiation downwards
Temperature and Pressure Variables	t2m	2m temperature
	d2m	2m dew point temperature
	sp	Surface pressure
Cloud Cover Variables	lcc	Low cloud cover
	mcc	Medium cloud cover
	hcc	High cloud cover

3.2 Data Pre-Processing

Rows with missing values were dropped to ensure the dataset is complete and ready for analysis. The 'Timev' column, containing date and time information, was utilized to extract separate 'Date' and 'Time' columns. The 'Date' and 'Time' columns were converted to datetime formats to facilitate time series analysis.

1) Autocorrelation: Autocorrelation is a statistical measure used to analyze the degree of similarity between a given time series and a lagged version of itself over successive time intervals. It helps in understanding the repeating patterns, cycles, or trends within the data. For time series forecasting tasks, especially with LSTM (Long Short-Term Memory) models, autocorrelation analysis is crucial as it informs the selection of important features and appropriate model parameters. Before selecting features and building the model, it is essential to examine the autocorrelation of the target variable, which is, the surface solar radiation downwards (ssrd_5). By analyzing the autocorrelation plot, we can identify significant lags that exhibit strong correlations with the current time step. This information is valuable in determining the appropriate window size and the number of units for the LSTM model. The autocorrelation plot (shown in the Figure 2) indicates a strong correlation at a 24-hour repetition, suggesting a daily cycle in the data. The plot shows a clear 24-hour cycle, which means that the solar radiation values are highly correlated with the values from the previous day.
2) Feature Selection: Relevant features for the analysis were selected. These include meteorological parameters such as wind speed, temperature, solar radiation, and cloud cover at different time intervals for location 5 and location 8, since they had the highest correlation in our analysis.
3) Normalization: The numerical features were normalized to scale the data within a range of 0 to 1. This step was crucial for ensuring that the neural network training process converges more efficiently.

The dataset was further then transformed into sequences using a sliding window approach. The window size, which determines the number of time steps used to predict the next value, was experimented with different sizes ranging from 2 to 6. For each window size, the data was reshaped into overlapping sequences to create the input (X) and output (y) datasets.

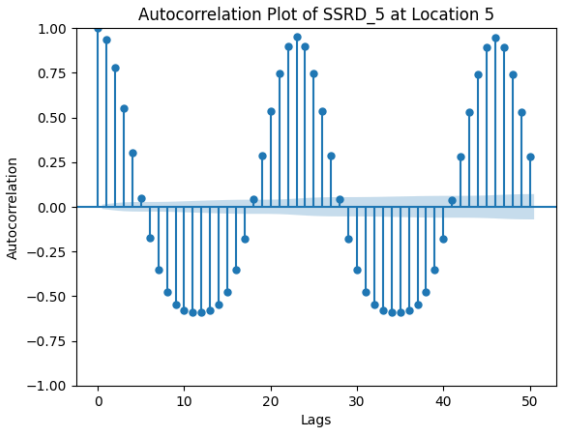


Figure 2: Autocorrelation of ssrd_5 at location 5.

4) Dataset Split: The dataset was divided into training, validation, and testing sets. The splitting approach used was chronological to preserve the temporal structure, ensuring the model is tested on unseen future data. This split helps in evaluating how well the model can predict future PV power output based on past data. Specifically, each sequence of weather parameters (X) was used to predict the subsequent solar radiation value (ssrd_5) (y). The sequences were split into training and testing sets using an 80-20 split ratio. The training set was further split into training and validation sets with an 80-20 split ratio for model evaluation.

3.3 Description/LSTM Model Architecture

The Long Short-Term Memory (LSTM) model is a type of recurrent neural network (RNN) designed to handle time series data and sequences. It addresses the vanishing gradient problem faced by traditional RNNs, enabling it to learn long-term dependencies.

Figure 3 represents the flow diagram of the LSTM approach proposed in this study.

3.3.1 Layers of the Proposed Model

1) Input Layer: The model takes sequences of shape (window_size, num_features), where window_size is the number of time steps in each input sequence, and num_features is the number of features in each time step.
2) LSTM Layer: The LSTM layer with 50 units processes the input sequences, capturing temporal dependencies:

$$\text{LSTM}(x_t) \rightarrow (h_t). \quad (1)$$

3) Dense Layer: A Dense layer with a single unit produces the final output, predicting the target variable:

$$\text{Dense}(h_t) \rightarrow (y_t). \quad (2)$$

4) Loss Function: The Mean Squared Error (MSE) is used as the loss function, measuring the average squared difference between the predicted values and the actual values:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2. \quad (3)$$

3.3.2 Experimental Setup

The LSTM model was constructed using the Keras library, with an architecture comprising a single LSTM layer with 50 units and a ReLU activation function. The choice of 50 units was guided by the autocorrelation analysis. The number 50 was chosen to approximately cover two days (48 hours) of data, providing the model with sufficient information to capture the daily cycles effectively. While 50 is also a common default setting in many Keras examples, the decision to retain this number was supported by the autocorrelation findings from Figure 2. The strong 24-hour cycle observed in the data justifies using a window that spans multiple days.

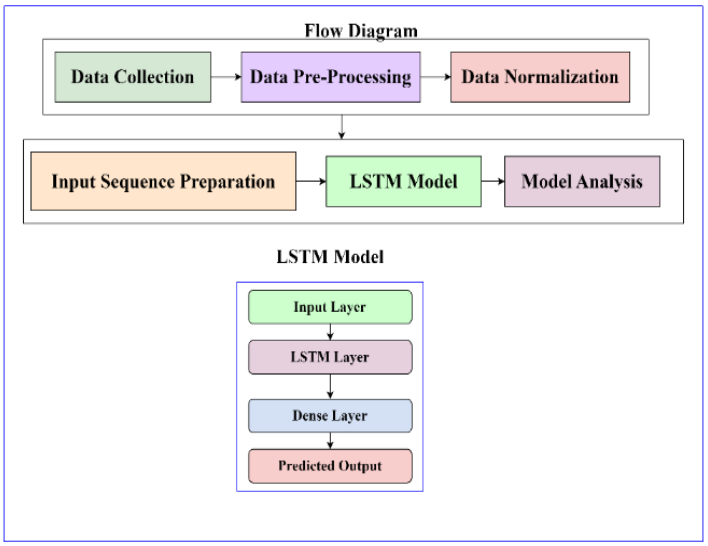


Figure 3: Flow diagram of the proposed model.

A Dense layer with a single neuron was added as the output layer to predict the solar radiation value. The model was compiled using the Adam optimizer and mean squared error (MSE) loss function. The model was trained on the training dataset with a batch size of 32 and a maximum of 100 epochs. Early stopping was employed with a patience of 5 epochs to prevent overfitting, ensuring the model retained the best weights based on validation loss. During cross-validation, different hyperparameters (e.g., number of LSTM units, batch size, learning rate) was tested to identify the optimal configuration that minimizes the validation loss. The model with the best performance across the cross-validation folds was selected as the final model. This model was then evaluated on the test set to confirm its performance. The dataset was divided into k folds (10). The model was trained and validated k times, each time using a different fold as the validation set and the remaining folds as the training set. For each fold, the same data preparation steps are followed: normalization, sequence preparation, and splitting into input (X) and target (y). The model is trained and evaluated on each fold, and performance metrics (e.g., MSE) are recorded. The performance metrics from all folds are averaged to provide a more accurate estimate of the model's performance. This average performance metric is considered the final evaluation of the model.

The trained model was evaluated on the test set to assess its performance using the mean squared error (MSE) metric. The training and validation loss over epochs were plotted to visualize the model's learning process and to identify any signs of overfitting. By following this methodology, the study aimed to develop a robust LSTM model for forecasting solar radiation using various meteorological features, contributing to more accurate weather predictions for the Abuja region. Table 2 presents the hyperparameters used in the proposed model.

Table 2: Hyperparameters description for the proposed model.

Hyperparameter	Window Size 2-6
Number of LSTM Units	50
Batch Size	32
Activation	ReLU
Epochs	100
Optimizer	Adam
Loss Function	Mean Squared Error (MSE)
Early Stopping Patience	5
Input Shape	(2, num_features)

3.4 Evaluation Metrics

1) Mean Absolute Error (MAE): It is the average of the absolute differences between the predicted and actual values. Mathematically, it can be represented as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \qquad (4)$$

where:
y_i is the actual value,
$\hat{y}_i$is the predicted value, and
n is the total number of data points.

2) Root Mean Squared Error (RMSE): It is the square root of the average of the squared differences between the predicted and actual values. Mathematically, it can be represented as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (5)$$

3) R^2 (Coefficient of Determination): It is a statistical measure that represents the proportion of the variance for a dependent variable that is explained by an independent variable or variables in a regression model. Mathematically, it can be represented as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \qquad (6)$$

where:
SS_{res} is the sum of squares of residuals, i.e., $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$,
SS_{tot} is the total sum of squares, i.e., $\sum_{i=1}^{n}(y_i - \bar{y})^2$, and
$\bar{y}$ is the mean of the observed data.

4 RESULTS AND DISCUSSION

In this section we discuss the results of the experiments performed in this study.

4.1 Months Analysis

The prediction analysis focuses on several months throughout the dataset. The specific months highlighted include January 1 to February 29 and July 1 to August 31, for the years 2020 to 2022. These months were chosen to represent the dry and wet seasons characterised by the location. The varying weather conditions during these seasons were considered to test the robustness of the LSTM models.

4.2 Auto Correlation and Cross Correlation Analysis

Figure 2 shows the autocorrelation for location 5 with significant lags that influence the PV power output, determining the optimal number of input nodes for the LSTM model. This lag is crucial in setting up the sequence length for the LSTM.

The cross-correlation analysis in Figure 4 shows that locations 5 and 8 for SSRD variable are closely related. This relationship suggests that these locations experience similar climatic conditions, which can be leveraged in the model for improved accuracy by incorporating spatial dependencies.

For consistency and based on the cross-correlation findings, location 5 was used for both input and output. This consistency helps in reducing complexity and improving model stability, ensuring that the predictions are reliable and based on closely related data.

The histogram of the ratio of `SSRD_5` (Surface Solar Radiation Downwards at location 5) and `SSRDC_5` (Surface Solar Radiation Downwards Clear Sky at location 5) shown in Figure 5 provides insights into cloud cover variability. High variability in this ratio indicates significant cloud cover, which affects solar radiation and PV output. The histogram helps in understanding the distribution and frequency of different cloud cover conditions over the studied period.

The analysis was conducted using different window sizes (Window 2, 3, 4, 5, 6):

1) Window 2: Considered a very short-term prediction, capturing immediate past data.
2) Window 3, 4, 5, 6: Incrementally larger windows capturing longer past data sequences, helping in understanding how far back the model needs to look to make accurate predictions. Figure 6 visualises the analysis of the MSE vs Window lengths across the entire year, dry and wet seasons.
3) SSRD_5 and FDIR_5: Predictions for these features was performed. The results are shown in Figure 6.

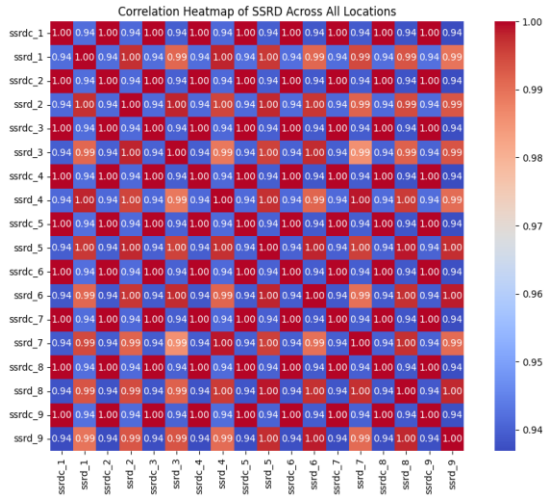


Figure 4: Correlation heatmap of SSRD across all locations.

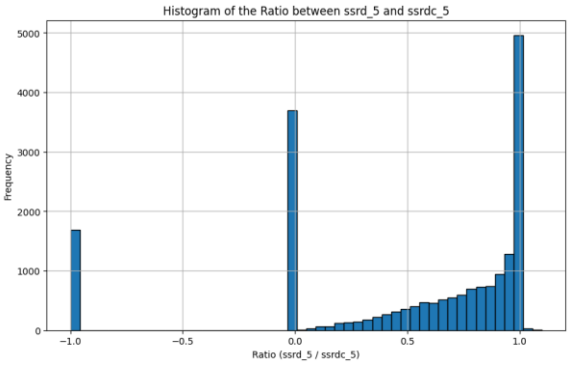


Figure 5: Ratio between ssrd_5 and ssrdc_5.

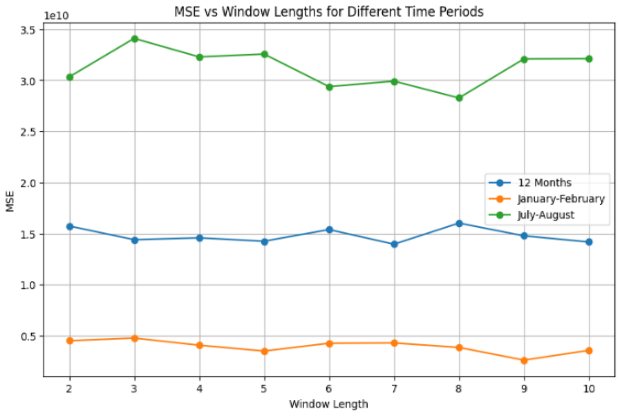


Figure 6: MSE vs window lengths for different time periods for year 2022.

4.3 Results Interpretation

The analysis demonstrates that LSTM models are highly effective in predicting PV power output in Abuja, leveraging weather data to account for both temporal and spatial factors. The model exhibits superior performance during stable weather seasons, such as the dry and rainy periods, which are characterized by consistent solar radiation patterns. This consistency allows the model to capture predictable dynamics, whereas transitional seasons pose greater challenges due to increased variability in weather conditions.

The 24-hour cycle identified through the autocorrelation study serves as a key insight for designing the model, providing a logical foundation for determining optimal window sizes. This alignment enables the model to effectively account for diurnal variations, improving accuracy and reliability. Moreover, the analysis highlights that incorporating data from multiple locations, especially those with strong spatial interdependencies like locations 5 and 8, enhances the model's ability to generalize across a wider range of conditions.

Figures 7, 8, and 9 provide evidence of how temporal and spatial configurations influence model performance. Window sizes tailored to the diurnal cycle result in reduced errors, as seen through lower MAE and RMSE metrics. Including features such as direct irradiance (FDIR_5) alongside solar radiation (SSRD_5) across all locations further boosts model accuracy, illustrating the value of combining complementary data sources.

The R^2 metric underscores the importance of spatial diversity in the input data, revealing that models trained with features from multiple locations achieve higher predictive accuracy and explanatory power. This reinforces the idea that broader spatial data not only improves generalization but also stabilizes performance across varying temporal scales.

While smaller window sizes fail to capture long-term dependencies and larger windows risk incorporating irrelevant information, an intermediate window size aligned with the 24-hour cycle strikes a balance between model complexity and accuracy. The non-linear relationship between window size and performance metrics, as observed in Figures 7 and 8, reflects this trade-off, emphasizing the need for careful tuning of input parameters.

Additionally, models leveraging multi-location data demonstrate greater resilience across changing window sizes, making them more robust for long-term forecasting. By incorporating direct measures of sunlight intensity, such as FDIR_5, the models further reduce errors and improve their explanatory power, highlighting the importance of precise feature selection [18].

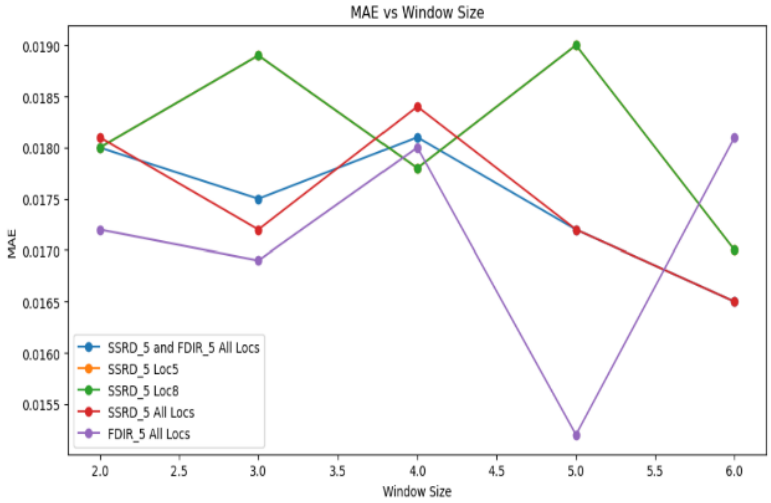


Figure 7: Performance metrics vs window sizes.

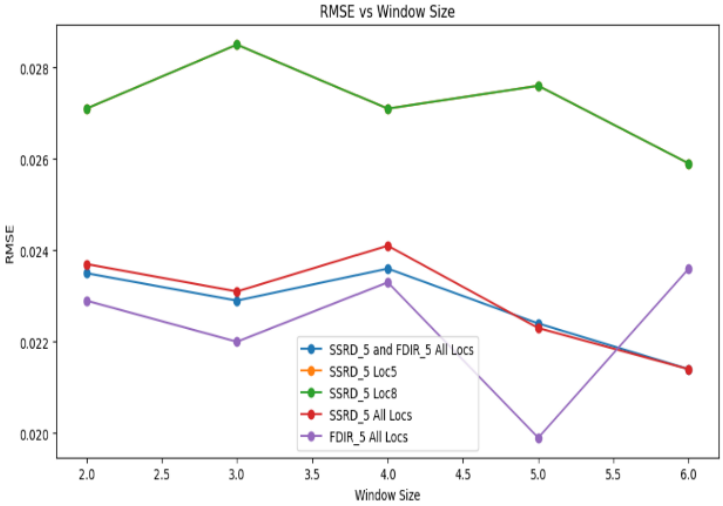


Figure 8: Performance metrics vs window sizes.

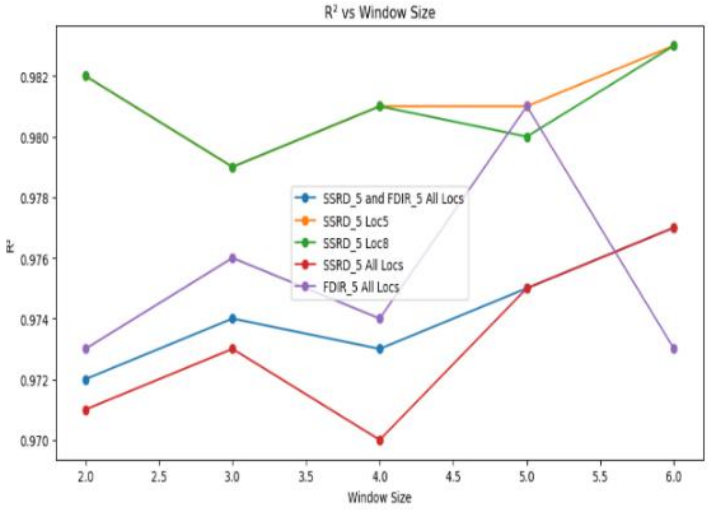


Figure 9: Performance metrics vs window sizes.

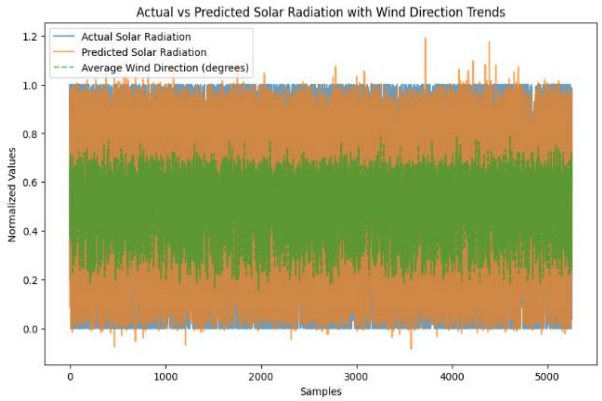


Figure 10: Actual vs Predicted Solar Radiation with Wind Direction Trends.

Figure 10 illustrates the relationship between actual solar radiation, predicted solar radiation, and average wind direction trends across the test dataset samples. The variations in solar radiation and wind direction trends reflect the dynamic behaviour of meteorological data. The green wind direction trendline indicates fluctuations in atmospheric conditions, which may have a subtle influence on solar radiation that the model captures. The predicted solar radiation values closely align with the actual values, showcasing the LSTM model's ability to effectively utilize temporal and meteorological features, including wind direction. Although wind direction does not display a direct cyclic pattern with solar radiation in this plot, its inclusion likely enhances the model's performance by accounting for

atmospheric effects such as cloud movement or aerosol dispersion, which impact solar radiation.

Our findings reveal that a hybrid approach combining deep learning techniques with statistically correlated features outperforms a hybrid model that relies solely on machine-learning techniques without leveraging the statistical relationships between features, as demonstrated in the recent results reported by Alkandari and Ahmad [18].

Seasonal variability is another important factor, as spatially diverse configurations are better equipped to handle the uncertainties of transitional periods. These findings highlight that the integration of temporal cycles, spatial interdependencies, and diverse weather features is key to building a reliable and accurate LSTM model for solar energy forecasting in Abuja. Figures 7, 8, and 9 validate this approach, offering practical guidance for optimizing predictive performance through feature selection and model design.

4.3 Implications for PV Power Forecasting in Abuja and Similar Urban Settings

The use of LSTM models, for predicting PV power in Abuja and similar urban areas has shown promise for application in cities with comparable weather conditions. Precise forecasts of PV power generation can enhance the efficiency and reliability of energy systems leading to seamless energy management and grid stability. The findings from this research, which highlights the significance of choosing relevant features and optimal time windows can be utilized in other locations to accelerate the integration of renewable energy sources into current power systems. This in turn aligns with the overarching objectives of reducing dependence on fossil fuels and promoting sustainable development. Our study introduces an innovative method for real-time, low-cost solar energy forecasting on resource-constrained edge devices, focusing on the evaluation and optimization of deep learning models to improve energy management for both residential and industrial applications, aligning with findings reported in [19].

5 CONCLUSIONS

This study effectively showcased the use of LSTM models to predict PV power output by utilizing weather features specific to Abuja, Nigeria. Overall, the analyses revealed that LSTM models, with a window size of 6 and a combination of meteorological and solar radiation features resulted in the most accurate predictions. The analyses of autocorrelation and cross correlation played a role in identifying the optimal feature selection and window sizes thereby significantly enhancing the model's predictive accuracy. Through comparing setups, it became clear that certain combinations of features and window sizes produced better performance metrics (MAE, RMSE and R^2). These results emphasize the significance of choosing relevant features and appropriate window lengths to optimize the model configuration for predicting PV power in urban environments. The implications of this study go beyond Abuja as it provides insights for urban areas facing similar climate conditions. Accurate PV power prediction can lead to improved energy management, grid stability and greater adoption of energy sources aligning with sustainable development objectives.

Future research could explore real-time data feeds, larger datasets, and additional climatic features to further improve the model's accuracy and applicability in different regions.

ACKNOWLEDGEMENTS

This research paper would not have been possible without the support of various individuals and organizations. We gratefully acknowledge the Nigerian Petroleum Technology Development Fund (PTDF) for its immense support in bringing this work to fruition.

REFERENCES

[1] C. Etukudor, B. Couraud, V. Robu, W. Fruh, D. Flynn, and C. Okereke, “Automated negotiation for peer-to-peer electricity trading in local energy markets,” Energies, vol. 13, no. 4, 2020. [Online]. Available: https://doi.org/10.3390/en13040920.

[2] S. M. Kadiri, A. O. Bagre, M. B. Camara, B. Dakyo, and Y. Coulibaly, “Electrical power distribution status in west assessment and perspective overview,” in Proc. IEEE Int. Conf. Renewable Energy Res. Appl. (ICRERA), 2019. [Online]. Available: https://doi.org/10.1109/ICRERA47325.2019.8997112.

[3] W. Tsai, C. Tu, C. Hong, and W. Lin, “A review of state-of-the-art and short-term forecasting models for solar PV power generation,” Energies, vol. 16, no. 14, 2023. [Online]. Available: https://doi.org/10.3390/en16145436.

[4] K. Tu, H. Xiong, and F. Chen, “Comprehensive prediction of generation power in photovoltaic microgrid,” in Proc. IEEE Joint Int. Inf. Technol. Artif.

Intell. Conf., 2019. [Online]. Available: https://doi.org/10.1109/ITAIC.2019.8785746.

[5] O. T. Ibitoye, O. S. Agunbiade, T. W. Ilemobola, A. B. Oluwadare, P. C. Ofodu, K. O. Lawal, and J. O. Dada, "Nigeria electricity grid and the potentials of renewable energy integration: A concise review," in Proc. IEEE 7th Int. Energy Conf. (ENERGYCON), 2022, pp. 1-4. [Online]. Available: https://doi.org/10.1109/ENERGYCON53164.2022.9830349.

[6] Y. N. Chanchangi, F. Adu, A. Ghosh, S. Sundaram, and T. K. Mallick, "Nigeria's energy review: Focusing on solar energy potential and penetration," Environ. Dev. Sustain., vol. 25, no. 7, pp. 5755-5796, 2023. [Online]. Available: https://doi.org/10.1007/s10668-022-02308-4.

[7] C. K. Rao, S. Sahoo, and F. F. Yanine, "Forecasting electric power generation in a photovoltaic power system for smart energy management," in Proc. Int. Conf. Intell. Controller Comput. Smart Power (ICICCSP), 2022. [Online]. Available: https://doi.org/10.1109/ICICCSP53532.2022.9862396.

[8] M. N. Akhter et al., "An hour-ahead PV power forecasting method based on an RNN-LSTM model for three different PV plants," Energies, vol. 15, p. 2243, 2022. [Online]. Available: https://doi.org/10.3390/EN15062243.

[9] Y. Wang, Y. Shen, S. Mao, X. Chen, and H. Zou, "LASSO and LSTM integrated temporal model for short-term solar intensity forecasting," IEEE Internet Things J., vol. 6, pp. 2933-2944, 2019. [Online]. Available: https://doi.org/10.1109/JIOT.2018.2877510.

[10] F. Jin, "Photovoltaic power generation prediction based on k-means clustering analysis and GRO-CNN-LSTM attention," Math. Model. Algorithm Appl., vol. 2, no. 1, pp. 61-68, 2024. [Online]. Available: https://doi.org/10.54097/qdcrpc46.

[11] R. Kabir, A. Elmouatamid, H. Elkhoukhi, and P. W. Pong, "Photovoltaic power forecasting using neural networks for short and medium-term dependencies," in Proc. IEEE Texas Power Energy Conf. (TPEC), 2024, pp. 1-6. [Online]. Available: https://doi.org/10.1109/TPEC60005.2024.10472207.

[12] T. Magesh et al., "Prediction of solar power irradiance using machine learning and deep learning for smart grid integration," in Proc. Int. Conf. Recent Adv. Electr. Electron. Ubiquitous Commun. Comput. Intell. (RAEEUCCI), 2024, pp. 1-6. [Online]. Available: https://doi.org/10.1109/RAEEUCCI61380.2024.10547750.

[13] C. K. Rodriguez-Leguizamon et al., "PV power generation forecasting based on XGBoost and LSTM models," in Proc. IEEE Workshop Power Electron. Power Qual. Appl. (PEPQA), 2023, pp. 1-6. [Online]. Available: https://doi.org/10.1109/PEPQA59611.2023.10325757.

[14] H. He, R. Hu, Y. Zhang, Y. Zhang, and R. Jiao, "A power forecasting approach for PV plant based on irradiance index and LSTM," in Proc. 37th Chin. Control Conf. (CCC), 2018, pp. 9404-9409. [Online]. Available: https://doi.org/10.23919/ChiCC.2018.8483960.

[15] T. Limouni et al., "Univariate and multivariate LSTM models for one step and multistep PV power forecasting," Int. J. Renew. Energy Dev., vol. 11, no. 3, pp. 815-824, 2022. [Online]. Available: https://doi.org/10.14710/ijred.2022.43953.

[16] R. Ahmed et al., "Computationally expedient photovoltaic power forecasting: A LSTM ensemble method augmented with adaptive weighting and data segmentation technique," Energy Convers. Manage., vol. 258, p. 115563, 2022. [Online]. Available: https://doi.org/10.1016/j.enconman.2022.115563.

[17] H. Hersbach et al., "ERA5 hourly data on single levels from 1940 to present," Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2023. [Online]. Available: https://doi.org/10.24381/cds.adbb2d47.

[18] M. AlKandari and I. Ahmad, "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods," Appl. Comput. Informat., vol. 20, no. 3/4, pp. 231-250, 2024. [Online]. Available: https://doi.org/10.1016/j.aci.2019.11.002.

[19] A. M. Hayajneh, F. Alasali, A. Salama, and W. Holderbaum, "Intelligent solar forecasts: Modern machine learning models and TinyML role for improved solar energy yield predictions," IEEE Access, vol. 12, pp. 10846-10864, 2024. [Online]. Available: https://doi.org/10.1109/ACCESS.2024.3354703.

Environmental Sustainability Assessment of Green Hydrogen Production from Seawater Using Life Cycle Assessment

Halina Falfushynska, Cornelia Scott and Markus Holz
Anhalt University of Applied Sciences, Strenzfelder Allee 28, 06406 Bernburg, Germany
galyna.falfushynska@hs-anhalt.de, cornelia.scott@hs-anhalt.de, markus.holz@hs-anhalt.de

Keywords: Water Desalination, Electrodialysis, Reverse Osmosis, Global Warming Potential, Renewable Energy.

Abstract: The electrolysis of water to produce green hydrogen fuel is pivotal for the transition to renewable energy sources. Predominantly, this process utilizes low-temperature alkaline or proton exchange membrane electrolysers, which require high-purity water, posing challenges for large-scale adoption due to freshwater scarcity. Seawater, comprising 96.5% of Earth's water reserves, presents an almost inexhaustible alternative. However, its complex composition, including various salts and organic compounds, complicates direct electrolysis. Specialized anodes and highly efficient electrocatalysts are essential to prevent corrosion and counteract undesirable chlorine evolution reactions. This study investigates the use of low-saline water from the Baltic Sea for hydrogen production via osmotic desalination and alkaline electrolysis using Life Cycle Assessment, focusing on economic, social, and environmental impacts. Findings indicate that electrodialysis is more energy-efficient compared to reverse osmosis, exhibiting lower environmental impacts across most categories, including global warming potential, ecotoxicity, and eutrophication. Reverse osmosis showed higher impacts, especially in fine particulate matter production and water-related parameters. Despite higher operational costs, integrating seawater desalination presents a promising method for renewable energy storage and hydrogen production. Optimizing electrodialysis could enhance its economic feasibility and performance, supporting sustainable green hydrogen production. Our research underscores the significant potential of seawater desalination coupled with electrolysis for sustainable energy transitions, particularly for regions with abundant seawater access but limited freshwater resources.

1 INTRODUCTION

The electrolysis of water to produce green hydrogen fuel is a cornerstone for the future of renewable energy. Currently, the most widespread technology of hydrogen production based on low-temperature alkaline electrolysers or proton exchange membrane electrolysers, primarily rely on high-purity water as feedstock [1]. However, if water electrolysis were to be adopted on a massive scale, as anticipated for the global energy landscape in the near future, issues related to water resource availability, particularly freshwater could emerge. This presents a significant challenge due to the exhaustive nature of freshwater sources. More than half of the world's population faces water scarcity for at least one month a year. Only in last 60 years renewable internal freshwater resources in Germany decrease in around 15% [2].

Seawater, constituting 96.5% of the Earth's water reserves, offers nearly unlimited availability as a natural electrolyte feedstock. Despite this vast potential, direct seawater splitting for hydrogen production is still in its early stages due to the complex composition of natural seawater. Seawater contains various dissolved salts and organic compounds, posing significant challenges for electrolysis. Efficient seawater splitting needs highly Oxygen evolution reaction-selective anodes to counteract chlorine evolution reactions, alongside highly efficient electrocatalysts to protect the electrolysis cell, particularly the anode, from chloride-induced corrosion [3]. Nonetheless, seawater electrolysis may yield more hydrogen from the same volume of water compared to freshwater, thanks to its enhanced conductivity due to the presence of alkali and alkaline earth metal cations [3].

Some studies have suggested that coupling seawater reverse osmosis or forward osmosis systems with conventional electrolysers could be a feasible solution for seawater electrolysis [4]. However, this approach requires additional purification steps to bring the treated seawater to the necessary purity

levels for current electrolysers, which increases the system's complexity and cost. This underscores the need for more efficient and direct methods of seawater electrolysis. On the other hand, some findings indicate that the increase in the levelized cost of hydrogen production is insignificant, as the capital and operating costs of seawater reverse osmosis are negligible [5]. Nonetheless, green hydrogen production from seawater has been largely overlooked, due to the belief that it is too expensive for industrial-scale implementation. To our knowledge, in Germany, there is currently only a pilot project, "OffsH2ore," which aims to produce green hydrogen in the North Sea.

To address these issues, the present study aimed to assess the economic, social, and environmental performance of producing green hydrogen through osmotic desalination and alkaline electrolysis of low-saline water from the Baltic Sea.

2 MATHERIALS AND METHODS

In this study, a comprehensive Life Cycle Assessment (LCA) using Open LCA software (https://www.openlca.org/) was performed to analyze green hydrogen production utilizing low-salt water from the Baltic Sea. This LCA study follows the four phases outlined in ISO 14040:2006 (ISO 2004). The first phase is the goal and scope definition. The second phase involves data collection and inventory analysis. The third phase focuses on assessing the life cycle impact of the two products. Finally, in the fourth phase, the results are interpreted, including a sensitivity analysis, and thoroughly discussed.

The system's scope incorporated set of input and output data, including resources, raw materials, machinery, power, the main and secondary products, waste, and contaminants. While key input parameters were adjusted for this study, majority of data have taken from Ecoinvent, needs_18 databases and references [6,7]. To evaluate the environmental impacts of H_2 production processes, the ReCiPe 2016 model as well as IMPACT 2002+ were employed [8]. ReCiPe 2016 model includes 22 midpoint impact categories related to global warming potential, ozone depletion, ionizing radiation, photochemical oxidant formation, human toxicity potential, ecotoxicity and eutrophication potentials and resource scarcity.

3 RESULTS AND DISCUSSION

In our study, we utilized life cycle inventory flows of each technology to compute life cycle midpoint impact category indicators for brackish water desalination might be used for green hydrogen production. This allowed us to shed light on the key midpoint environmental performance indicators and then identify the processes responsible for the potential impacts of the water desalination technologies required for subsequent green hydrogen production (Table 1).

Table 1: Life cycle environmental impacts of seawater desalination in terms of reverse osmosis and electrodialysis. The impacts are expressed per 1000 m^3 of potable water in ReCiPe 2016 model.

Name	RO	ED	Unit
Fine particulate matter	3.00E-5	1.27E-5	DALY
Fossil resource scarcity	3.11	1.33	USD2013
FW ecotoxicity	1.43E-11	5.87E-12	species.yr
FW eutrophication	1.47E-11	4.44E-12	species.yr
GW, FW ecosystems	3.54E-11	1.53E-11	species.yr
GW, Human health	6.49E-4	2.80E-4	DALY
GW, Terrestrial ecosystems	1.30E-6	5.60E-7	species.yr
Human carcinogenic toxicity	2.03E-6	3.79 E-7	DALY
Human non-carcinogenic toxicity	4.50E-5	1.88E-5	DALY
Ionizing radiation	1.68E-7	6.40E-8	DALY
Land use	1.11E-10	7.22E-8	species.yr
Marine ecotoxicity	2.51E-8	1.07E-8	species.yr
Marine eutrophication	1.91E-10	1.94E-13	species.yr
Mineral resource scarcity	0.002	0.010	USD2013
Ozone formation, Human health	1.14E-7	4.88E-8	DALY
Ozone formation, Terrestrial ecosystems	1.67E-8	7.15E-9	species.yr
Stratospheric ozone depletion	6.47E-8	3.08E-8	DALY
Terrestrial acidification	3.45E-8	1.48E-8	species.yr
Terrestrial ecotoxicity	1.46E-9	6.27E-10	species.yr
WC, Aquatic ecosystems	2.42E-15	3.41E-14	species.yr
WC, Human health	8.90E-9	1.25E-7	DALY
WC, Terrestrial ecosystem	5.41E-11	7.62E-10	species.yr

GW: Global warming, Freshwater: FW, Water consumption: WC

One significant finding was that the potential for fine particulate matter production, which serves as an indicator for NO_x and SO_x emissions, was markedly higher for Reverse Osmosis (RO) compared to Electrodialysis (ED). Similar trends were observed for other impact categories as well: the Global Warming Potential indicators, ecotoxicity-related indicators, and eutrophication indicators all demonstrated more intensive effects for RO than for ED.

Conversely, parameters associated with water consumption and related conjunction effects, as well as land occupation, were found to be higher for ED. However, these parameters contributed less than 5% across all impact categories for both RO and ED technologies.

Furthermore, both ED and RO processes contribute to the acidification of aquatic reservoirs, albeit to a marginal extent. Of the two, RO has a more significant potential to support water eutrophication (Table 2).

Table 2: Life cycle environmental impacts of seawater desalination in terms of reverse osmosis and electrodialysis. The impacts are expressed per 1000 m^3 of potable water in IMPACT2002+ model.

Name	RO	ED	Unit
Aquatic acidification	0.240	0.089	kg SO2 eq
Aquatic ecotoxicity	1783.3	509.8	kg TEG water
Aquatic eutrophication	0.336	0.0018	kg PO4 P-lim
Carcinogens	0.239	0.258	kg C2H3Cl eq
Global warming	52.59	22.73	kg CO2 eq
Ionizing radiation	476.08	160.64	Bq C-14 eq
Land occupation	0.013	8.576	m2org.arable
Mineral extraction	0.563	0.650	MJ surplus
Non-carcinogens	1.069	0.462	kg C2H3Cl eq
Non-renewable energy	566.59	253.15	MJ primary
Ozone layer depletion	1.86E-7	2.19E-7	kg CFC-11 eq
Respiratory inorganics	0.027	0.011	kg PM2.5 eq
Respiratory organics	0.025	0.011	kg C2H4 eq
Terrestrial acid/nutri	0.769	0.335	kg SO2 eq
Terrestrial ecotoxicity	235.71	99.2	kg TEG soil

The integration of data using a Sankey diagram highlights that the energy supply required for both ED and RO has the most significant impact (see Fig. 1).

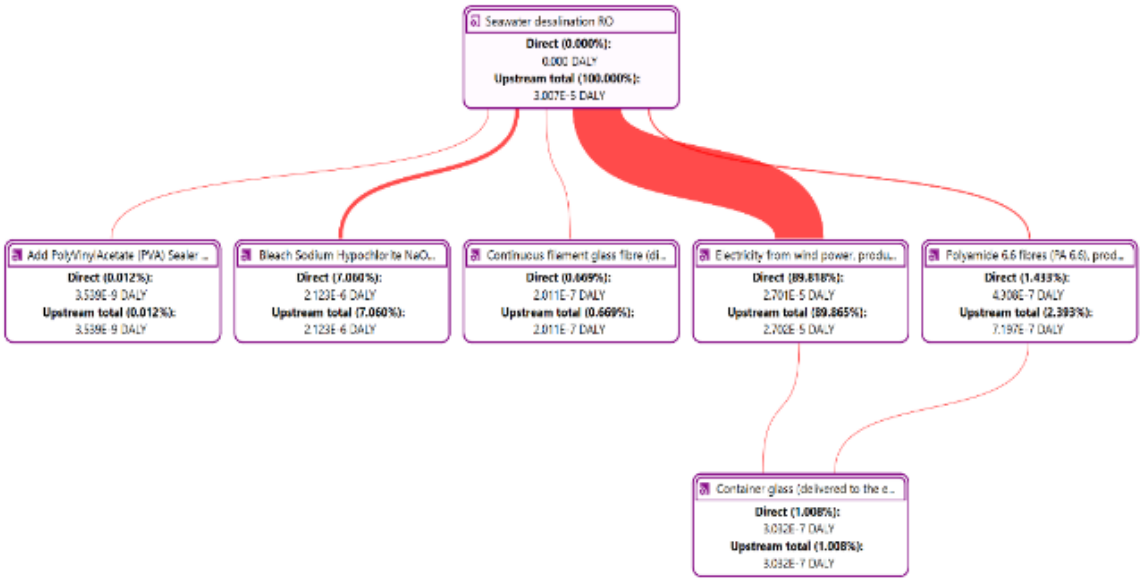
Our investigation into utilizing seawater at coastal locations highlights promising opportunities to harness sea water as abundant resource for storing surplus renewable electricity. While the cost of water from large-scale desalination plants is relatively minor compared to the overall expenses of producing green hydrogen through electrolysis, developing and maintaining desalination and deionization plants for water electrolysis demands significant capital investment and incurs ongoing operational and maintenance costs

Direct seawater use for water electrolysis without deionization presents several challenges that necessitate water desalination prior to hydrogen production. The most prevalent technique for this purpose is reverse osmosis [9]. However, only a limited number of studies explore alternative methods, such as electrodialysis or thin-film composite forward osmosis [10]. Some research indicates that electrodialysis may be a cost-effective alternative to reverse osmosis for low-salt desalination. If the costs of ion-exchange membranes are reduced and their performance is enhanced, electrodialysis could become economically preferable to reverse osmosis across the entire range of brackish water salinity [11].

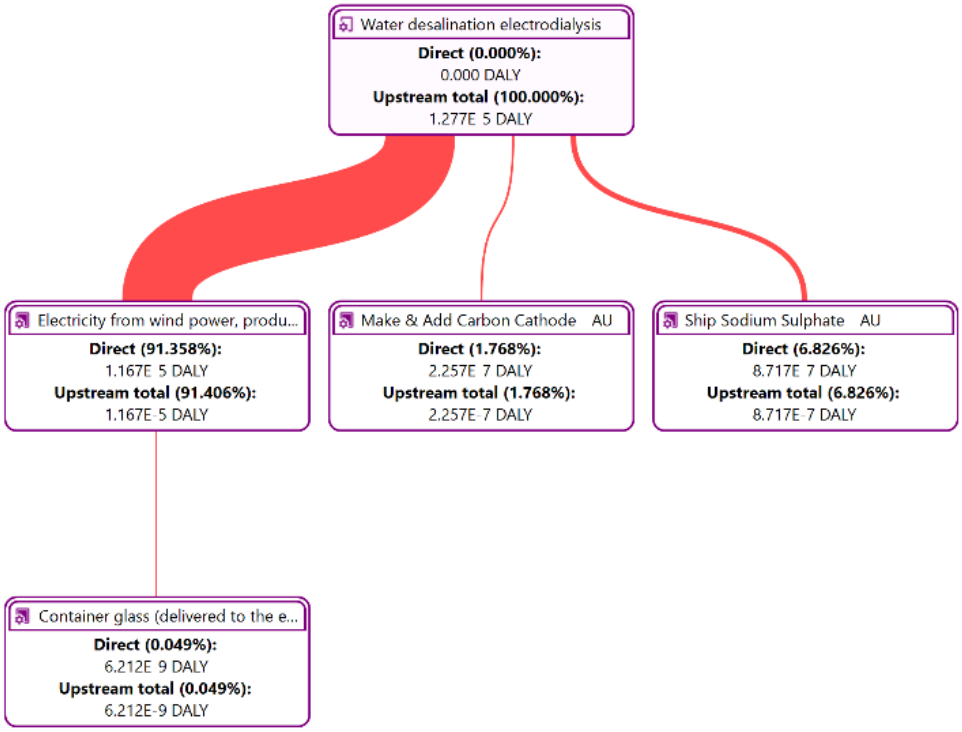
Germany, with access to the low-salinity Baltic Sea, has the potential to tap into an endless source of sustainable energy carriers. Proactive steps in this regard have already been made by research teams in Poland [12] who revealed that combining reverse osmosis with multiple-effect desalination systems can achieve water quality suitable for both alkaline and proton exchange membrane electrolytic cells. Our findings support that the electrodialysis treatment of brackish water from the Baltic Sea might reduce climate change potentials compared to reverse osmosis across nearly all considered scenarios. This advantage is largely attributed to lower energy requirements for desalination, measured at 2-3 kWh/m^3 compared to 3-4 kWh/m^3 for reverse osmosis [13].

Electrodialysis desalination showed to be more energy-effective related to non-renewables and therefore less carbon emission than reverse osmosis. This consumption is in good agreement with previous findings [11,14].

The IMPACT 2002+ model identifies water desalination as a key factor potentially contributing to aquatic ecotoxicity, particularly in the case of reverse osmosis due to brine discharge and associated salinity



a)



b)

Figure 1: Sankey diagram of environmental impacts of seawater desalination using reverse osmosis a) and electrodialysis b) technologies. The impacts are expressed per 1000 m^3 of potable water

changes. Increased salinity is known to induce more significant physiological and biochemical changes in marine organisms than chemical stressors [15]. Osmotic stress can lead to immune disorders, metabolic alterations, and increased oxidative stress [15, 16]. Nonetheless, the range of aquatic toxicity observed in our study is pretty low, favoring the use of brackish water for hydrogen production via potable water splitting.

Water consumption is critically important in our study, especially in regions with high potential for solar energy but limited water resources, such as desert areas. It is noteworthy that water consumption in our study is quite low, aligning well with previous findings [13]. As a cost-effective and environmentally friendly option, using seawater for hydrogen production offers a sustainable resource for renewable energy accumulation, thereby optimizing energy transition. This approach holds potential for further exploration and could support the economic and social development of low-income countries, such as those in Africa and South America, while also benefiting European nations.

Both studied methods for water desalination demonstrated a low impact on human health and may not induce carcinogenic pathologies, making them user-friendly options for producing potable water. These methods are suitable not only for green hydrogen production as a renewable energy carrier but also for supplying potable water in areas facing scarcity. Specifically, a photovoltaic-powered desalination system using time-variant electrodialysis reversal technology has been shown to provide brackish water desalination in India with a 22% cost reduction, making it competitive with fossil fuel-powered alternatives [14].

3 CONCLUSIONS

In conclusion, our study highlights the significant differences in environmental impacts between reverse osmosis and electrodialysis desalination technologies of brackish water for green hydrogen production. Integrating seawater desalination in coastal regions presents a viable pathway for renewable energy storage and hydrogen production. Electrodialysis pretends to be a more energy-efficient option, achieving lower environmental footprints in categories such as global warming potential, ecotoxicity, and eutrophication. Meanwhile, reverse osmosis exhibited higher impacts particularly in fine particulate matter production and water-related parameters. Using nanofiltration as pre-treatment for reverse osmosis might increase fresh-water output. Meantime forward osmosis and freeze desalination are expected to be promising for brine post-treatment, reducing environmental harm. Additionally, innovative nanomaterial membranes improve selectivity, reduce pressure, and lower costs [17]. These findings advocate for further exploration into optimizing electrodialysis, potentially reducing costs and enhancing performance, thereby supporting sustainable and economically feasible hydrogen production.

ACKNOWLEDGMENTS

This research was funded by the BMBF Project “Professor*innengewinnung und Nachwuchsentwicklung zur Etablierung eines Centers of Advanced Scientific Education (CASE).

Also, we acknowledge support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Open Access Publishing Fund of Anhalt University of Applied Sciences.

REFERENCES

[1] H. Falfushynska and M. Holz “Power-to-X Strategies: A Key Driver for Decarbonization and Renewable Energy Integration in Economies,” Lecture Notes in Networks and Systems, in press, 2024.

[2] Renewable internal freshwater resources per capita (cubic meters) – Germany. https://data.worldbank.org/indicator/ER.H2O.INTR.PC?locations=DE.

[3] J. Mohammed-Ibrahim and H. Moussab. “Recent advances on hydrogen production through seawater electrolysis,” Materials Science for Energy Technologies, vol. 3, pp. 780-807, 2020.

[4] S. S. Veroneau, A. C. Hartnett, A. E. Thorarinsdottir, and D. G. Nocera. “Direct seawater splitting by forward osmosis coupled to water electrolysis,” ACS Applied Energy Materials, vol. 5(2), pp. 1403–1408, 2022.

[5] M.A. Khan, T. Al-Attas, S. Roy, M.M. Rahman, N. Ghaffour, V. Thangadurai, S. Larter, J. Hu, P.M. Ajayan, and M.G. Kibria. “Seawater electrolysis for hydrogen production: A solution looking for a problem?” Energy & Environmental Science, vol. 9, 2021.

[6] S. Fayyaz, S. Khadem Masjedi, A. Kazemi, E. Khaki, M. Moeinaddini, and S. Irving Olsen. “Life cycle assessment of reverse osmosis for high-salinity seawater desalination process: Potable and industrial water production,” Journal of Cleaner Production, vol. 382, pp. 135299, 2023.

[7] E. Najjar, M. Al-Hindi, M. Massoud, and W. Saad. "Life Cycle Assessment of a seawater reverse osmosis plant powered by a hybrid energy system (fossil fuel and waste to energy)," Energy Reports, vol. 7(Supplement 5), 448-465, 2021.

[8] M.A.J. Huijbregts, Z.J.N. Steinmann, P.M.F. Elshout, G. Stam, F. Verones, M. Vieira, M. Zijp, A. Hollander, R. van Zelm. "ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level," Int. J. Life Cycle Assess., vol. 22, pp. 138-147, 2017.

[9] V.S.S. Katakam and V. Bahadur. „Reverse osmosis-based water treatment for green hydrogen production," Desalination, vol. 581, pp. 117588, 2024.

[10] B. E. Logan, L. Shi, and R. Rossi. "Enabling the use of seawater for hydrogen gas production in water electrolyzers," Joule, vol. 5(4), pp. 760-762, 2021.

[11] S. K. Patel, B. Lee, P. Westerhoff, and M. Elimelech. "The potential of electrodialysis as a cost-effective alternative to reverse osmosis for brackish water desalination," Water Research, vol. 250, pp. 121009. 2024.

[12] K. Sztekler, W. Kalawa, T. Bujok, P. Boruta, E. Radomska, Ł. Mika, ..., and A. Widuch. "Hybrid desalination system for Baltic Sea water: A preliminary study," Desalination, vol. 574, pp. 117269, 2024.

[13] C. Skuse, R.R.Z. Tarpani, P. Gorgojo, A. Gallego-Schmid, and A. Azapagic. "Comparative life cycle assessment of seawater desalination technologies enhanced by graphene membranes," Desalination, vol. 551, pp. 116418, 2023.

[14] W. He, A.-C. Le Henaff, S. Amrose, T. Buonassisi, I. M. Peters, and A. G. V. Winter. "Flexible batch electrodialysis for low-cost solar-powered brackish water desalination," Nature Water, vol. 2(370–379), 2024.

[15] H. Falfushynska, F. Wu, E.P. Sokolov, and I.M. Sokolova. "Salinity variation modulates cellular stress response to ZnO nanoparticles in a sentinel marine bivalve, the blue mussel Mytilus sp." Marine Environmental Research, vol. 183, pp. 105834, 2023.

[16] M.N. Noor, F. Wu, E.P. Sokolov, F. Haider, and I.M. Sokolova. "Salinity-dependent effects of ZnO nanoparticles on bioenergetics and intermediate metabolite homeostasis in a euryhaline marine bivalve, Mytilus edulis," Science Total Environment, vol. 774, pp. 145195, 2021.

[17] E.J. Okampo and N. Nwulu. "Optimisation of renewable energy powered reverse osmosis desalination systems: A state-of-the-art review," Renewable and Sustainable Energy Reviews, vol. 140, p. 110712, 2021.

SECTION 5

Information Technology in Economics

Proceedings of the 12th International Conference on Applied Innovations in IT (ICAIIT)

NOVEMBER 2024

Modeling the Determinants of Global Economic Uncertainty

Olha Kovalchuk[1], Kateryna Berezka[2], Oksana Tulai[3], Bohdan Ferens[4], Ludmila Babala[5] and Oleksandr Zapolskyi[3]

[1]*Theory of Law and Constitutionalism Department, West Ukrainian National University, Lvivska Str. 11, 46009 Ternopil, Ukraine*

[2]*Department of Applied Mathematics, West Ukrainian National University, Lvivska Str. 11, 46009 Ternopil, Ukraine*

[3]*Department of International Relations and Diplomacy, West Ukrainian National University, Lvivska Str. 11, 46009 Ternopil, Ukraine*

[4]*Committee on Integration of Ukraine into EU, Verkhovna Rada of Ukraine, Hrushevskogo Str. 5, 01008 Kyiv, Ukraine*

[5]*Department of Economic Cybernetics and Informatics, West Ukrainian National University, Lvivska Str. 11, 46009 Ternopil, Ukraine*

o.kovalchuk@wunu.edu.ua, k.berezka@wunu.edu.ua, o.tulai@wunu.edu.ua, ferens@rada.gov.ua, ludaduma7@gmail.com, zapolskyolexandr@gmail.com

Keywords: Economic Uncertainty, World Uncertainty Index, Cross-Country Analysis, Correspondence Analysis, Economic Development, Decision-Making, Global Risks.

Abstract: Economic uncertainty poses a significant challenge for decision-makers across the public and private sectors. While uncertainty is an inherently nebulous concept, developing consistent quantitative measures allows for rigorous analysis of its impacts. The World Uncertainty Index (WUI) provides a standardized quarterly index of uncertainty levels for 143 countries dating back to 1952 based on language in Economist Intelligence Unit reports. This study applies correspondence analysis to examine the relationship between countries' WUI values and their levels of economic development classified by the IMF's income groups. The results reveal distinct associations - advanced, high-income economies exhibit relatively low uncertainty while emerging markets and developing economies face higher uncertainty levels. Low-income countries experience moderate uncertainty. These findings underscore how economic instability can impede development progress. By quantifying uncertainty through empirical measures and analyzing its linkages with other economic factors, researchers can derive valuable insights for policymakers aiming to cultivate confidence and stability.

1 INTRODUCTION

Several major events in recent years, such as the global financial and ecological crisis, the growth of cybercrime, increasing political polarization and trade tensions, the COVID-19 pandemic, and the war in Ukraine have heightened concerns about rising levels of economic uncertainty worldwide [1–3]. However, quantifying uncertainty in a consistent way that allows for comparisons across different periods and countries poses an intrinsic challenge. Uncertainty is an inherently nebulous concept, reflecting the state of uncertain minds among consumers, business leaders, and policymakers about future potential events and outcomes. It is also a broad concept, relating to macro-economic phenomena like GDP growth rates as well as micro-level aspects like the growth trajectories of individual firms. Moreover, uncertainty extends beyond just economic factors to encompass other major events and issues such as elections, wars, and climate change. Despite its nebulous and wide-ranging nature, developing robust measurements of uncertainty is crucial for research and analysis.

WUI is a panel index that measures uncertainty levels across 143 countries every quarter, going back to 1952. It provides uncertainty data for a wide range of developed and developing nations over an extensive period. The WUI is calculated by counting how frequently the word “uncertainty” and its variants (like “uncertain”) appear in the country reports published by the Economist Intelligence Unit. To enable cross-country comparability, these raw uncertainty word counts are scaled by the total word count of each report. In other words, the WUI reflects the number of “uncertainty” words expressed per thousand words in each report [4].

Having a standardized panel index of uncertainty covering many countries over multiple decades represents a major new resource. It allows researchers to analyze uncertainty trends over time and compare uncertainty levels between different nations and regions in a consistent quantitative manner [5–7].

The series of major disruptive events that have rocked the global economy in recent years, sometimes stemming from political rifts between nations, have ushered in a new era of heightened turbulence and volatility. These turbulent episodes have caused uncertainty levels to skyrocket to exceptionally high levels worldwide, according to the research. Elevated uncertainty, in turn, has acted as a drag on economic growth.

As depicted in the Chart of the World Uncertainty Index (Fig. 1), while the index fell slightly in December 2023, it has remained at elevated levels in recent times due to the compounding effects of successive shocks. Among the most recent shocks were Russia's invasion of Ukraine and the associated cost-of-living crisis rippling across the world [4].

The WUI illustrates the rise in war-related uncertainty spanning the globe. While uncertainty initially peaked in European nations versus other regions, this geographic gap has narrowed over time, underscoring the war's widening economic spillover effects.

Despite the December 2023 dip, the index continues to reflect the new normal of higher uncertainty that has taken hold amid the tumultuous global conditions over the past few years. The turbulence from repeated economic, political, and geopolitical shocks has kept worldwide uncertainty readings much higher than historical levels.

Analyzing and quantifying uncertainty becomes particularly valuable in the context of the multitude of major shocks that have impacted the global economy over the past several years. Measuring uncertainty through indices like the World Uncertainty Index allows researchers to disentangle and examine the various potential sources and drivers contributing to heightened uncertainty levels during this tumultuous period.

With the global economy being repeatedly buffeted by disruptive events ranging from the financial crisis to trade conflicts, political upheaval, the pandemic, and geopolitical tensions, uncertainty analysis provides a lens to assess the relative impacts of each of these shocks. By dissecting how uncertainty levels responded to specific shocks, we can better understand which events or sources of turbulence were most destabilizing and detrimental to economic certainty worldwide.

Analyzing uncertainty through empirical measures is crucial for policymakers, businesses, and economists to comprehend the prevailing global economic climate. During periods of concurrent, compounding crises, uncertainty monitoring offers insights into disentangling the complex mix of factors fueling business and consumer uncertainty that weighs on economic growth and decision-making.

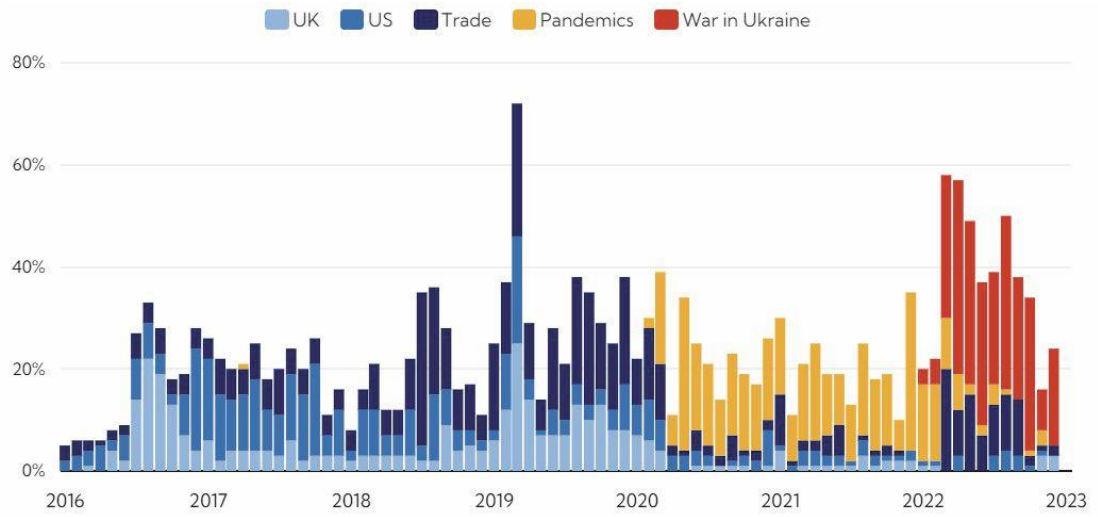


Figure 1: 3D plot of row and column coordinates [6].

2 RELATED WORKS

Recently, scholars and practitioners have been paying close attention to the study of serious challenges that create political and economic uncertainty on a global scale. The authors E. Bouri et al. analyzed the role of monetary policy uncertainty in predicting jumps in nine advanced equity markets [8]. Researchers H. Chen et al. studied the impact of economic policy uncertainty on capital investment by Australian firms [9]. H. Almustafa et al. examined the effect of economic policy uncertainty on firm-level investment and corporate financial leverage [10]. I. Khandokar and A. Serletis investigated the dynamic empirical relationship between modern risk/uncertainty indicators and leverage [11]. M. S. Kaviani et al. studied the relation between changes in policy uncertainty and changes in credit spreads. They found that macroeconomic conditions, including general uncertainty, do not explain this result, which also holds when they use instrumental variables to address endogeneity issues [12]. H. Ahir et al. introduced the WUI, which utilizes Economist Intelligence Unit reports [13]. However, reports on developed economies demonstrate a more detailed and technical presentation style with an emphasis on monetary policy and financial markets, while reports on developing countries often focus more on political risks and institutional changes [6]. The reports also vary in their depth of regional coverage: larger economies and key regional players receive comprehensive analysis with quarterly updates, while reports for smaller economies may be less detailed and updated less frequently [14]. S. R. Baker et al. developed the Economic Policy Uncertainty Index (EPU), which is based on the analysis of newspaper articles in developed countries [15]. While both the EPU and WUI measure uncertainty, the WUI demonstrates broader global coverage by capturing major international events like the 9/11 attacks, the SARS outbreak, and Brexit, whereas the EPU tends to focus more on domestic policy uncertainty. The WUI shows that uncertainty levels are generally lower in advanced economies compared to developing countries, with uncertainty spikes being more synchronized among advanced economies and those with stronger trade and financial connections. Despite the existence of several publications on this topic [16], further scientific research on economic uncertainty is needed due to the emergence of new risks and challenges that continue to shape the global economic landscape.

3 METHODOLOGY

Uncertainty is an important factor in making many decisions [17]. It negatively impacts economic development, while stability and confidence in the future foster growth. In this work, an attempt is made to identify non-obvious relationships between the level of uncertainty and the economic development of countries around the world.

The WUI quantifies the level of uncertainty mentioned in the Economist Intelligence Unit's country reports. It is calculated by finding the percentage of times the word "uncertain" (or its variants like "uncertainty") appears in the reports. This percentage is then scaled up by a factor of 1,000,000 to obtain the WUI value. A higher WUI number signals greater uncertainty expressed in the reports, while a lower value indicates less uncertainty being mentioned. For example, a WUI of 200 means the word "uncertain" accounted for 0.02% of all words used in the reports. Given that these reports average around 10,000 words, a WUI of 200 roughly corresponds to the word "uncertain" appearing twice per report on average. Therefore, the WUI provides a quantitative measure of the degree of uncertainty discussed and projected in these country reports from the Economist Intelligence Unit [18].

The WUI is computed as follows:

$$Q_1 = \frac{0.6 \cdot Q_4 + 0.3 \cdot Q_3 + 0.1 \cdot Q_2}{3},$$

where Q_1, Q_2, Q_3, Q_4 are the quarters of the World Uncertainty Index for 143 countries from the 1950s onwards. The 3-quarter weighted moving average of the index serves as the preferred measure when analyzing data at the country level [18].

The IMF classifies countries into three broad groups primarily based on their income levels and stages of economic development:

- Advanced Economies. These are high-income countries with well-developed economies and advanced economic structures. Examples include the United States, Japan, Germany, the United Kingdom, France, etc.
- Emerging Market and Developing Economies: This group includes both emerging market economies and other developing economies. Emerging markets tend to have rapid economic growth and increasing economic liberalization, like Ukraine, China, India, Brazil, etc. Other developing economies have lower incomes and are at earlier stages of development.
- Low-Income Developing Countries. These are the poorest countries in the world with low gross

national income per capita. Many are faced with severe economic, political, and social challenges. Examples include Afghanistan, Haiti, Yemen, South Sudan, etc.

To identify non-obvious interdependencies between WUI levels and the economic development of countries around the world, correspondence analysis was used [19]. It is an exploratory technique used to visualize and analyze the associations within high-dimensional contingency tables. Its computational objective is to represent the distances between rows and columns in a lower-dimensional space while preserving the relationships as accurately as possible.

The method utilizes the Pearson chi-squared statistic to evaluate how well a lower-dimensional representation captures the structure of the original high-dimensional table. Essentially, correspondence analysis performs factor analysis on categorical data, acting as a dimensionality reduction technique.

The rows and columns of the initial table are mapped to points in space, with the chi-squared distance calculated between them. The goal is to find a low-dimensional (typically 2D) space that minimizes the distortion of these distances, thereby reproducing the structure of the original table faithfully.

Correspondence analysis operates on frequency tables, consisting of rows representing one set of categorical variables and columns representing another set. The following terminology is associated with this technique:

- Mass. The observations in the table are normalized by calculating relative frequencies. The sum of all elements in the table becomes equal to 1 (each element is divided by the total number of observations). This resulting standardized table shows how the mass is distributed across the cells or points in space. The row and column sums in the matrix of relative frequencies are referred to as the row mass and column mass, respectively.
- Quality: In correspondence analysis, quality refers to how well a row or column point is represented in the coordinate system defined by the selected number of dimensions. The quality of a point is defined as the ratio of the squared distance from the point to the origin in the chosen dimensionality, divided by the maximum squared distance in the full-dimensional space. A low quality indicates that the selected dimensionality inadequately represents the corresponding row or column.
- Relative Inertia. Inertia is defined as Pearson's chi-squared statistic for a 2x2 table, divided by the total number of observations. Relative inertia represents the dimensionality's contribution. A partial solution may represent a point reasonably well (high quality), but that point may contribute little to the overall inertia.
- Row & Column Profiles. If the rows and columns of the table are completely independent, then the elements can be represented using the row and column sums or, in correspondence analysis terminology, using the row and column profiles.
- Relative Dim n. This column displays the relative contribution of the corresponding row or column point to the inertia accounted for by dimension n. This value is provided for each point (row or column) across all dimensions.
- Cosine^2 – Quality. This column contains the quality of representation for each point in the corresponding dimension. The cosine squared can be interpreted as the "correlation" between the point and that dimension. It is the square of the cosine of the angle formed by the point and the dimension's axis.
- Metric Coordinate System. In correspondence analysis, the term "distance" refers to the differences between the rows and columns of the relative frequency matrix, represented in a lower-dimensional space. The coordinates in this reduced space represent these distances. However, unlike standard Euclidean distances calculated directly from the frequencies, these distances are weighted.

Graphical analysis is the most crucial part of correspondence analysis. Typically, the horizontal axis represents the dimension accounting for maximum inertia. The plot shows the percentage of total inertia explained by each eigenvalue. The smaller the distance between points of the same type (rows or columns), the stronger their association.

To assess the relationship between points of different types, one must consider the angles they form with the vertex at the centroid (0,0) coordinates.

The general rules for visually assessing the degree of dependence are:

- Draw line segments from two points of different types to the centroid.
- If the angle formed is acute, the row and column are positively correlated.
- If the angle is obtuse, the correlation between the variables is inverse/negative.
- If the angle is straight (90 degrees), there is no correlation.

- The angles between row and column points relative to the centroid reveal the nature and strength of their associations in the low-dimensional representation.

4 RESULTS AND DISCUSSION

We applied correspondence analysis [19] to identify non-obvious relationships between IMF income (advanced economies, emerging economies, low-income economies) [20] and the WUI for 143 countries [18].

For 143 countries, the values of the World Uncertainty Index were determined in the range from 0 to 0.31. One-third of the countries (31) have a WUI value less than 0.025, one-third (31 countries) – have a WUI value in the range from 0.025 to 0.057, and one-third – a WUI value greater than 0.057 [14]. We ranked the WUI for the 143 analyzed countries into the following levels (Table 1) [8].

Frequency tables were constructed for the distribution of the analyzed countries into groups by levels of economic development (advanced economies, emerging economies, and low-income economies) and levels of WUI (low, medium, high) (Table 2).

Among the 143 analyzed countries, the following groups were identified:

- 12 countries with advanced economies and low WUI;
- 7 countries with advanced economies and medium WUI;
- 11 countries with advanced economies and high WUI;
- 19 countries with emerging economies and low WUI;
- 15 countries with emerging economies and medium WUI;
- 27 countries with emerging economies and high WUI;
- 12 countries with low-income economies and low WUI;
- 18 countries with low-income economies and medium WUI;
- 24 countries with low-income economies and high WUI.

Table 1: Grouping of countries worldwide based on their values.

Ranking	Range
low	<0.025
medium	between 0.025 and 0.057
high	> 0.057

Table 2: Frequency table for IMF income and WUI.

	Observed Table (Frequencies) Row variables: IMF_income (3) Column variables: WUI_Rank (3)			
Edvanced Economies	12	11	7	30
Emerging Economies	19	27	15	61
Low-Income Economies	10	24	18	52
Total	41	62	40	143

Correspondence analysis is fundamentally a component decomposition of the chi-squared (χ^2) statistic. Its primary objective is to identify the lowest dimensional space that can adequately represent the deviations from the expected values. Table 3 presents the calculated eigenvalues, which indicate the minimum number of dimensions required to qualitatively capture the information contained in the data tables for each pair of analyzed variables.

For a qualitative representation of the contingency table between the IMF income and WUI, two dimensions are sufficient. First dimension accounts for 56% of the total inertia, while the second accounts for the remaining 44%. Pearson's chi-squared (χ^2) statistic objectively assesses how close the empirical distributions are to the theoretical ones. The obtained level of 0.001 indicates statistical significance. With 9 degrees of freedom (df = 9), the calculated χ^2 value is 26139.8, which exceeds the critical value of 28.88 at the 0.001 level. Therefore, it can be stated that the predicted values closely match the observed ones.

Special statistics are employed to assess the quality of the solution obtained from correspondence analysis. Ideally, all or most of the points should be accurately represented – the distances between them should not be significantly distorted as a result of applying the dimensionality reduction procedure. Table 4-5 presents the calculation results of these statistics based on the row and column coordinates.

Table 3: Eigenvalues and inertia for all dimensions.

	Total Inertia = 1.00				
Number of Dims.	Singular Values	Eigen-Values	Pers. of Inertia	Cumulative Percent	χ^2
1	0.7487	0.5605	56.0521	56.0521	14784.78
2	0.6629	0.4395	43.9479	100.0000	11592.09

Table 4: Row coordinates and contributions to inertia.

Row Name	Row Coordinates and Contributions to Inertia Input Table (Rows × Columns): 3×3 Standardization: Row and column profiles Row Num.	Coordin. Dim. 1	Coordin. Dim. 2	Mass	Quality	Relative Inertia	Inertia Dim. 1	Cosine² Dim. 1	Inertia Dim. 2	Cosine² Dim. 2
advanced economies	1	-0.2444	0.0550	0.2098	1	0.3919	0.3916	0.9518	0.3986	0.0481
emerging economies	2	-0.0671	-0.0438	0.4266	1	0.0815	0.0560	0.7016	0.5134	0.2984
low-income economies	3	0.2197	0.0196	0.3636	1	0.5266	0.5484	0.992	0.0879	0.,0079

Table 5: Column coordinates and contributions to inertia.

Row Name	Column Coordinates and Contributions to Inertia Input Table (Rows × Columns): 3×3 Standardization: Row and column profiles Column Number	Coordin . Dim. 1	Coordin. Dim. 2	Mass	Quality	Relative Inertia	Inertia Dim. 1	Cosine² Dim. 1	Inertia Dim. 2	Cosine² Dim. 2
advanced economies	1	-0.2741	0.01500	0.2867	1	0.6430	0.6729	0.9970	0.0404	0.0030
emerging economies	2	0.0697	-0.0429	0.4336	1	0.0864	0.0658	0.7256	0.5006	0.2744
low-income economies	3	0,172932	0.0511	0,2797	1	0.2707	0.2607	0.9197	0.4590	0,0803

For the two-dimensional solution obtained from correspondence analysis of the IMF income and the World Uncertainty Index, a high-quality value of 1 was achieved for all groups of convicts. This indicates that the selected two dimensions adequately represent all the rows and columns of the original data table.

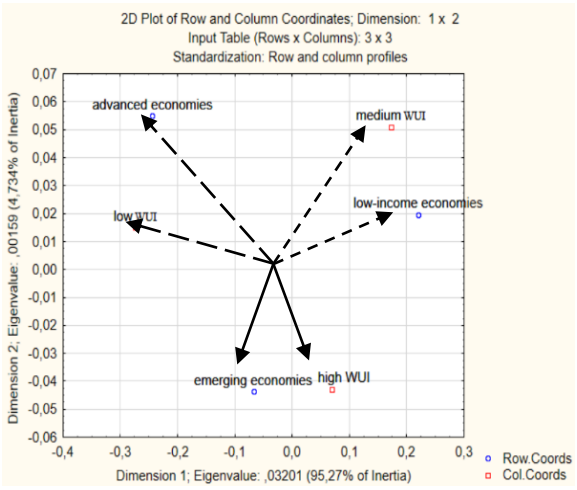


Figure 2: 3D plot of row and column coordinates.

A graphical analysis was conducted to examine the relationship between IMF income and the WUI used in this analysis. The analysis of 2-dimensional plots depicting the row and column coordinates for the corresponding pairs of variables provides the basis for the following conclusions (Fig. 2):

- most high-income countries have a low World Uncertainty Index;
- emerging market and developing economies are characterized by a high World Uncertainty Index;
- low-income developing countries have a medium World Uncertainty Index.

The analysis revealed distinct associations between countries' levels of economic development based on IMF income classifications and their degrees of economic uncertainty captured by the WUI. The results indicate that advanced, high-income economies tend to exhibit relatively low levels of uncertainty while emerging markets and developing economies are more prone to higher uncertainty. Low-income developing countries fall somewhere in between, with moderate uncertainty levels. These findings underscore how economic uncertainty can act as a headwind to development, with more economically advanced nations better

positioned to cultivate stability and confidence. By quantifying uncertainty through empirical indices like the WUI and employing analytical techniques like correspondence analysis, we can gain deeper insights into the complex interplay between economic turbulence and a country's stage of development. Such analysis can help policymakers and economists better understand and address the factors fueling economic uncertainty across the global landscape.

5 CONCLUSIONS

This study applied correspondence analysis to investigate the relationship between countries' economic development levels, as classified by the IMF's income groups, and their degrees of economic uncertainty measured by the WUI. The results revealed clear associations, with advanced, high-income economies exhibiting relatively low uncertainty levels, emerging and developing economies facing higher uncertainty, and low-income countries experiencing moderate uncertainty. These findings underscore the importance of economic stability and confidence in fostering growth and development. Nations with more established, robust economies appear better able to mitigate uncertainty and its detrimental impacts. Conversely, emerging markets and less-developed countries grapple with greater volatility, which can hinder investment, spending, and overall economic progress.

Quantifying uncertainty through empirical measures like the WUI and rigorously analyzing its relationships with other economic indicators represents a valuable contribution. It provides empirical insights into the complex dynamics between uncertainty, instability, and a nation's economic standing on the global stage. Future research could focus on investigating the relationship between prolonged periods of high uncertainty and key economic indicators, as well as evaluating the effectiveness of various policy measures aimed at reducing economic instability and bolstering confidence during turbulent periods.

REFERENCES

[1] L. He, L. Huang, and L. Fang, "Institutional conditions, economic policy uncertainty and foreign institutional investment in China," Emerg. Mark. Rev., vol. 50, p. 100823, 2022.

[2] O. Kovalchuk, K. Berezka, L. Babala, R. Ivanytskyy, N. Karpyshyn, and N. Zhuk, "Modeling Country Economic Security: a Machine Learning Approach," in Proc. 14th Int. Conf. Adv. Comput. Inf. Technol. (ACIT), 2024, pp. 370–375.

[3] K. M. Berezka and O. Ya. Kovalchuk, "Associative Rules for Modeling International Security Decisions in the Context of the Ukrainian-Russian War. Preliminary Evaluations," Folia Oeconomica, vol. 23, no. 2, pp. 67–85, 2023.

[4] H. Ahir, N. Bloom, and D. Furceri, "Global Economic Uncertainty Remains Elevated, Weighing on Growth," IMF Blog, Jan. 26, 2023.

[5] N. Liu and F. Gao, "The world uncertainty index and GDP growth rate," Finance Res. Lett., vol. 49, p. 103137, 2022.

[6] H. Ahir, N. Bloom, and D. Furceri, "The World Uncertainty Index," NBER Working Paper 29763, 2022, p. 114.

[7] H. Ahir, N. Bloom, D. Furceri, R. Hossain, S. Das, and D. Roy, "The World Uncertainty Index And The Macroeconomic Indicators; A Case For India," Educ. Admin. Theory Pract., vol. 30, no. 5, pp. 5719–5731, 2024.

[8] E. Bouri, K. Gkillas, R. Gupta, and C. Kyei, "Monetary policy uncertainty and jumps in advanced equity markets," J. Risk, vol. 23, pp. 101–112, 2020.

[9] X. Chen, C. H. A. Le, Ya. Shan, and S. Taylor, "Australian policy uncertainty and corporate investment," Pac.-Basin Finance J., vol. 61, p. 101341, 2020.

[10] H. Almustafa, I. Jabbouri, and P. Kijkasiwat, "Economic Policy Uncertainty, Financial Leverage, and Corporate Investment: Evidence from U.S. Firms," Economies, vol. 11, no. 2, p. 37, 2023.

[11] I. Khandokar and A. Serletis, "Risk, uncertainty, and leverage," Econ. Modell., vol. 91, pp. 257–273, 2020.

[12] M. S. Kaviani, L. Kryzanowski, H. Maleki, and P. Savor, "Policy uncertainty and corporate credit spreads," J. Finance Econ., vol. 138, no. 3, pp. 838–865, 2020.

[13] Uncertainty Index, IMF, 2018. [Online]. Available: http://dx.doi.org/10.2139/ssrn.3275033.

[14] World Uncertainty Index, Federal Reserve Bank. [Online]. Available: https://fred.stlouisfed.org/release?rid=470.

[15] S. R. Baker, N. Bloom, and S. J. Davis, "Measuring Economic Policy Uncertainty," Q. J. Econ., vol. 131, no. 4, pp. 1593-1636, 2016.

[16] D. Deepa et al., "Taxonomy of Global Risk, Uncertainty, and Volatility Measures," Int. Finance Discuss. Pap., no. 1216, 2017. [Online]. Available: https://www.researchgate.net/publication/.

[17] S. Sniazhko and E. Muralidharan, "Uncertainty in decision-making: A review of the international business literature," Cogent Bus. Manag., vol. 6, no. 1, 2019.

[18] World Uncertainty Index. [Online]. Available: https://worlduncertaintyindex.com/.

[19] M. Riani, A. C. Atkinson, F. Torti, and A. Corbellini, "Robust Correspondence Analysis," J. R. Stat. Soc. C Appl. Stat., vol. 71, no. 5, pp. 1381–1401, 2022.

[20] IMF/World Economic Outlook, United Nations. [Online]. Available: https://media.un.org/.

Marketing Logistics and its Digitalization

Iryna Polishchuk[1], Yuriy Dovhan[1], Iryna Kramar[2], Larysa Dovhan[1] and Oksana Yakushevska[1]
[1]*Vinnytsia Institute of Trade and Economics of the State University of Trade and Economics, Soborna Str. 87, 21050 Vinnytsia, Ukraine*
[2]*Economics and Finance Department, Ternopil Ivan Puluj National Technical University, Ruska Str. 56, 46001* Ternopil, *Ukraine*
i.polishchuk@vtei.edu.ua, y.dovhan@vtei.edu.ua, iryna_stankevych@odaba.edu.ua, l.dovhan@vtei.edu.ua, o.yakushevska@vtei.edu.ua

Keywords: Consumer, Management, Marketing Logistics, Digitalization, Digital Tools.

Abstract: The article highlights the essence and role of marketing logistics in the conditions of the modern market environment and overviews the trends in its digitalization. The study is relevant due to the strategic importance of marketing logistics for business competitiveness, improving operational efficiency, and meeting modern consumer demands. The purpose of the article is to highlight the historical process of forming marketing logistics as a modern management concept and the current trends in its digitalization. The theoretical foundations of marketing logistics summarized in this work will facilitate the integration of digital technologies into marketing logistics, improve business processes, and support the sustainable development of companies in the context of dynamic changes in the global market. Comparative, institutional, systemic, and structural-functional methods are used to conduct research. Considerable attention is paid to highlighting the reasons for the need for digitization of marketing logistics and the problems that arise in its process. The main reason for the need for digitization of marketing logistics is defined as obtaining competitive advantages by companies. The main problems that arise in the process of marketing logistics digitalization include a lack of a digitalization strategy, a lack of technology awareness, etc. The necessity of training marketing logistics specialists who have deep professional knowledge and skills in the field of economics and management, as well as in the field of digital information and communication technologies is emphasized. The main modern trends in the use of digital tools of marketing logistics are characterized. It is concluded that marketing logistics is an effective concept of business process management, and its digitalization is an integral component of modern economic development, which contributes to increasing the efficiency of companies and maintaining their competitiveness.

1 INTRODUCTION

In modern conditions of highly competitive product markets and constant changes in the marketing environment, one of the important ways of achieving success for companies is to speed up the time goods go through supply chains and build flexible logistics systems. This can be ensured by transitioning the management of companies' business processes to marketing logistics approaches.

Marketing and logistics integration allow companies to focus on finding better ways to serve consumers as a source of competitive advantage [1]. Speed, quality of service, and customer satisfaction are considered to be a crucial factor that unites marketing and logistics.

Today, there is a rapid digitization of the economy, and marketing logistics in particular, which creates favorable conditions for increasing the efficiency of companies, increasing the productivity of employees, better identifying and meeting the needs of consumers, and ensuring the appropriate level of service.

Modern automation tools, capabilities of computer and information-and-communication systems amaze the imagination and make the prospects of using digital solutions in marketing logistics extremely significant.

2 MARKETING LOGISTICS AS A MODERN CONCEPT OF COMPANY MANAGEMENT

Economic science has a large number of theories, approaches and concepts of company management. With the development of socio-economic relations, constant changes in the market environment and the constant growth of competition, management concepts disappeared or improved. Thus, in the middle of the 20th century, two relatively young concepts of company management emerged, firstly, a marketing one and then, a little bit later, a logistics one.

Scientists and practitioners understood an increasing need to deepen the integration of both concepts, which gave rise to the emergence of the concept of marketing logistics. Converse (1954) noted that physical distribution should be considered in the context of marketing activities, marking one of the first steps toward integrating marketing and logistics [2]. Kotler (1967) emphasized the importance of planning, implementing, and controlling the physical flow of goods from their point of origin to the point of consumption to meet customer needs [3]. Drucker (1962) substantiated the importance of an integrated approach to distribution and marketing, contributing to the development of the concept of marketing logistics [4]. Christopher and Peck (1997) focused on the strategic integration of marketing and logistics processes to create value for consumers, considering the integration of marketing and logistics within the supply chain context [5].

A new management concept is based on the following ones:

1) a marketing concept of management focused on consumers and the market;
2) a logistics concept of management focused on logistics flow.

According to the concept of marketing logistics, company management should focus on both the market and consumers, as well as the logistics flow.

Subsequently, researchers expanded the concept of marketing logistics to the level of integration of marketing management with supply chain management [6; 7]. Marketing logistics seeks to manage the interface between the marketing and logistics activities of the company to align their respective strategies within the context of the wider supply chain [1].

The concept of marketing logistics is aimed at achieving optimal interaction of marketing and logistics processes. The main goal of marketing logistics is to meet the needs of consumers and achieve a high level of efficiency of the entire supply chain [8]. Marketing logistics is considered as a strategic tool that helps companies to effectively interact with the market, predict, and quickly adapt to changes in the market environment. Effective integration of the company's marketing and logistics, and as a result, the implementation of marketing logistics approaches, is the factor that can ensure its competitiveness and success [9].

Generally, marketing logistics interaction is the following [10]:

1) interrelated management concepts (logistics and marketing as a streaming market construct);
2) logistics is the basis for the application of a marketing as the basis for developing a logistics strategy and optimizing the logistics system;
3) tools for achieving and maintaining competitive advantages.

Among the main functions of marketing logistics, there can be distinguished demand forecasting, order management, customer service, procurement management, inventory management, warehousing, transportation, distribution management, management of relationships with participants in the logistics supply chain (suppliers, partners, consumers).

In the approaches to company management, the synthesis of marketing and logistics is based on the contact points of its marketing and logistics mixes.

Stock and Lambert [11] pay considerable attention to the study of the relationship between marketing activities of companies and logistics. The "Place" element of the marketing mix is defined as a point of contact between marketing and logistics activities. This is explained by the fact that at the point of sale, customer service is carried out, the requirements of which are determined by marketing, but fulfilled by logistics. The authors claim that consumer service is the result of the joint work of marketing and logistics, which are connected by a common goal, namely satisfaction of requests and needs of consumers.

We share this view. However, we consider that the number of such contact points involved in the interaction of the classic marketing mix "4 Ps" ("Product", "Price", "Place", "Promotion") and the logistics mix "8 Rs" ("Product", "Quantity", "Quality", "Consumer", "Place", "Time", "Expenses", "Service") is much more (Figure 1).

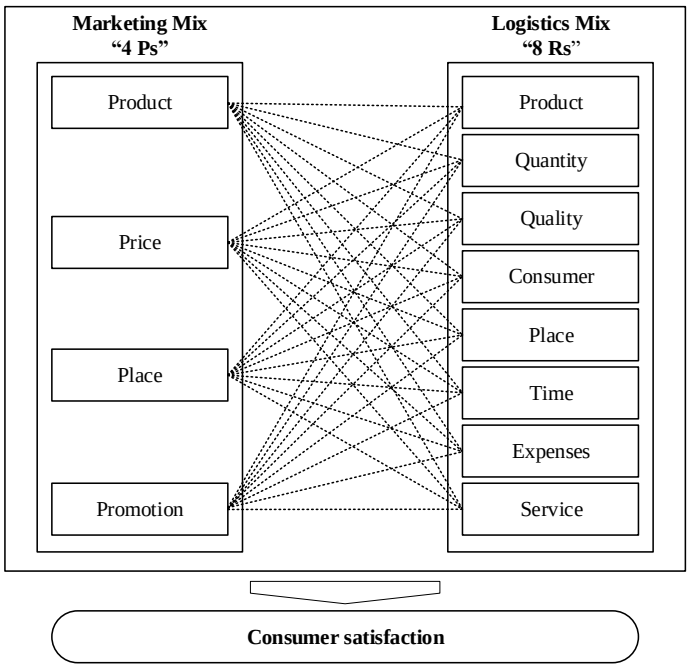


Figure 1: Interaction of the marketing and logistics mix.

Studying the market and identifying consumer needs takes place through marketing research. But even before they are held, logistics already play an important role. Since certain markets and consumers may be logistically inaccessible or limited in access for the company, it is impossible to achieve the fulfillment of the rules of the logistics mix "8 Rs". The reasons why logistics is limited or unavailable at all can be different, e.g. distant geographical location, political, legal, economic or other barriers (wars, pandemics, etc.). Logistical limitations significantly affect the decision-making of companies to conduct such marketing research, at least in the short term.

If marketing research has been conducted, consumer needs have been identified, and target market segments have been determined, and then the stage of marketing development of a product to satisfy them begins. This is the traditional area of the company's marketing product policy ("Product" marketing mix element), which is provided in cooperation with R&D departments. But even here, companies cannot do without a logistical approach in determining what the product should be to become logistically available to consumers, in particular, what should be its physical and chemical properties, dimensions, packaging, etc. A similar situation occurs in the assortment management field – the logistics largely determines the possibility of its formation, expansion or narrowing. The role of logistics in making decisions about eliminating products and their practical implementation is also important.

There are much more contact points between marketing and logistics than those referring to the "Place" element of the marketing complex. From our point of view, they are widely present in the "Product", "Price" and "Promotion" elements as well.

Maximal satisfaction of requests and needs of consumers is achieved under the condition that consumers are fully met with each of the elements of both the marketing and logistics mix. In our opinion, this is the goal of marketing logistics.

3 REASONS AND PROBLEMS OF DIGITALIZATION OF MARKETING LOGISTICS

Digitalization is changing the world, industries, companies, business models, business processes, and it is changing marketing logistics as well.

One of the main tasks of marketing logistics is to create and develop a marketing information system for effective distribution, conduct market research, and organize information support for the movement of logistics flow in supply chains. The fulfillment of these tasks should be ensured by the information and communication component of marketing logistics, which is currently most affected by digitalization.

Digitalization of marketing logistics involves using digital technologies and implementing digital information and communication innovations to increase the flexibility and productivity of the company's marketing and logistics system.

Digitization of the company's business processes and marketing logistics, in particular, requires the introduction of a complex of information, computing and communication technologies [12].

In tough competitive conditions, digital innovations allow companies to automate many marketing logistics processes, reduce the number of managers' routine tasks, improve communication within the company due to the structuring and ordering of data, and reduce marketing logistics costs, i.e. gain competitive advantages.

However, digital innovations in marketing logistics are implemented somewhat slower than in other areas of activity, which is a significant risk not only for an individual company but for all its supply chain partners as well [13]. Some reasons are as follows [14]:

1) digitalization is perceived as unattainable;
2) lack of a digitalization strategy;
3) ack of technological awareness.

The success of digitalization of marketing logistics depends on the success of elimination of these causes by company management.

It should be noted that one of the most important factors of a company's success or failure is people and their decisions. Therefore, it is not enough for managers to offer digital innovations and equip the company's specialists with digital information and communication technologies of marketing logistics, they need to be trained to use all this in a high-quality way.

Informatization and digitalization change the conditions of professional activity of specialists, which leads to the need to combine different types of activities in professional activity. Therefore, the professional training of marketing logistics specialists should be based on an approach in which, in addition to economic and managerial issues, an important place in training will be given to technological and informational components [15]. Today, specialists who have not only a certain specialization with deep professional knowledge and abilities, but also an understanding of other key functions and the ability to perform them are considered to be effective [1]. The implementation of digital innovations and digital information and communication technologies in the management of marketing logistics processes of companies requires appropriately trained specialists.

4 KEY TRENDS OF MARKETING LOGISTICS DIGITALIZATION

Digital tools of marketing logistics in the field of marketing include the following ones: E-commerce; Internet platforms and aggregators; Internet advertising and search engine optimization; social media marketing, web promotion and brand reputation management in social networks and the Internet; big data, blockchain, and marketing analytics. In the field of logistics, electronic data interchange; satellite technologies; automated inventory and warehouse management systems; 3D printing; organization of deliveries by drones; unmanned deliveries; blockchain; big data, and logistics activity analytics can be used [10]. This list is not complete and is constantly updated.

Let's focus on the main modern trends in the use of digital marketing logistics tools (Figure 2):

- Electronic Data Interchange Technologies. They are able to quickly and effectively ensure the implementation of internal (within the company between departments and employees) and external document flow (with consumers, business partners and with government bodies). Modern electronic document management systems increase the efficiency of work processes and decision-making in the field of marketing logistics, provide a high level of data security, reduce data loss, and simplify document management. Walmart, one of the largest retail chains in the world, actively uses these technologies to optimize its logistics system. This includes integration with suppliers through the receipt of electronic orders, invoices, and delivery confirmations, rapid inventory updates through real-time data synchronization, reduced order processing times, and increased sustainability by significantly reducing paper usage, aligning with the company's sustainable development goals.
- Cloud Storage. These technologies allow data processing and storage due to convenient and free access to IT resources through the provider in remote access mode. Cloud Storage services are capable of covering the entire range of marketing logistics processes, e.g. demand research, customer search, resource procurement, goods delivery, customer settlements, after-sales service, reverse logistics, etc.

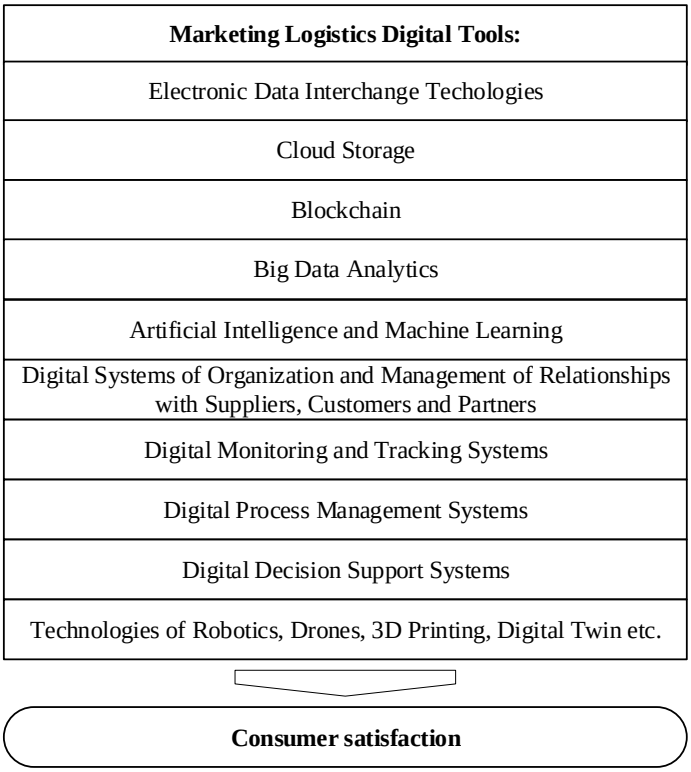


Figure 2: Marketing logistics digital tools.

- Blockchain. The implementation of blockchain technology makes it possible to store the required amount of data, protect it from illegal access, prevent the possibility of changing digital information about the process of movement of the logistics flow, and save the financial resources of all participants in logistics processes. Security, anonymity, and integrity of data without any third-party organization controlling transactions are the main aspects of the great popularity of blockchain technology in marketing logistics. Leading companies (CMA CGM, MSC Mediterranean Shipping Company, Hapag-Lloyd, and others) utilize blockchain-based freight platforms (such as TradeLens until 2023) to optimize and digitize global supply chains. These platforms address key logistics challenges, including lack of transparency, delays, and the complexity of coordination among multiple participants.
- Big Data Analytics. These technologies are actively used by companies when conducting marketing and logistics research, in particular, for collecting information, forming databases, processing and analyzing large data sets, forming forecasts and strategies for marketing and logistics activities. Big Data methods made the foundation for personalization systems that are widely used in almost any business today. Supply chain participants are also using big data to improve business performance, in particular, Big Data analysis to optimize the use of warehouse space, routes, and transport loading allows to reduce logistics costs.
- Artificial Intelligence and Machine Learning. Artificial intelligence technology is changing the field of marketing by providing a personalized experience to consumers. Artificial intelligence provides efficient analysis of large volumes of data and can help in accurately forecasting the demand for goods and services, which is important to prevent the formation of excessive or insufficient stocks of resources and finished products. Artificial intelligence is also used to automate logistics processes such as order processing, delivery tracking, route planning, etc. With Artificial Intelligence and machine learning, business analysts can quickly assess productivity, shipment speed, customer satisfaction and

other important variables. UPS leverages artificial intelligence to optimize delivery routes through its ORION platform, saving millions of liters of fuel. DHL uses machine learning for inventory management, demand forecasting, and parcel sorting.

- Digital Systems of Organization and Management of Relationships with Suppliers, Customers and Partners (Supplier Relationship Management Systems, Customer Relationship Management Systems, Partner Relationship Management Systems etc.). Digital systems integrated into the company's corporate information environment based on a single database are a crucial component of a modern business strategy. Such systems allow companies to improve interaction with consumers and partners through personalization. They also help organize work between consumers, performers, intermediaries, and other participants, find free freight carriers, calculate the cost of services, update current information about delivery times, control all orders, and remind managers about the need for their timely processing, etc.
- Digital Monitoring and Tracking Systems. In particular, the Internet of Things technology involves connecting physical devices to the Internet, which ensures the connection of the material world with the digital one. Smart devices (smartphones, laptops, tablets, computers, smart TVs, vehicles, smart home appliances, door locks, smart watches, heart and blood pressure monitors, fitness bracelets, etc.) use sensors and mechanisms to collect and data exchange with the environment and with each other. Such digital systems can collect more data about the target audience, data about consumers and their behavior, which makes it possible to create better products and services to meet their needs, successfully sell these products and services, and personalize marketing. The Internet of Things technology is widely used in logistics, allowing to improve the efficiency of processes and provide more accurate control over the movement of goods from the supplier to the final consumer. It is possible to monitor traffic, optimize routes, track the movement of goods in warehouses, control their location and condition, as well as automate the processes of picking and packing orders. For example, a sensor that is installed to control the temperature inside the car when transporting perishable products collects information and allows you to adjust the temperature automatically. Radio Frequency Identification systems allow tracking by using radio frequency to transmit information using tags (microchips) attached to physical objects. With their help, fast and clear tracking of where and how cargo is moved is carried out, acceptance and shipment are accelerated, and the level of transparency and reliability of operations is increased.
- Digital Process Management Systems and (Production Management Systems, Transportation Management Systems, Warehouse Management Systems, Inventory Management Systems etc.). These systems increase the efficiency of marketing logistics processes at various stages of the movement of the logistics flow, contribute to the optimization of supply chains, direct the company's orientation to demand and consumers, facilitate management decision-making, and improve communication between participants in business processes. DHL actively implements digital process management systems to optimize its operations. DHL utilizes the MySupplyChain system, which provides real-time transparency for all logistics operations, integrating processes from warehousing to delivery while minimizing human involvement.
- Digital Decision Support Systems. These information systems support business or organizational decision-making activities. They serve the level of management, operations and planning of the company and help to make decisions on the problems that can change quickly and are not easily defined in advance, i.e. unstructured and semi-structured decision-making problems in marketing logistics. The use of digital technologies and algorithms for data collection, processing and analysis allows effective management decisions to be made.
- Technologies of Robotics. Companies are actively implementing the robotization of production and warehouse logistics, which is a global trend caused by the need to speed up logistics processes. Robotic Process Automation is a tool for automating business processes, it is the technology that allows automating a large volume of tasks for marketing logistics.
- Drones. Drones allow them to create unique and engaging content, collect and analyze data

that will help them improve decision-making and optimize campaigns. For example, a drone can measure the traffic and attendance of a particular location and help determine the best time and place to display your ad. Drones can also track consumer reactions and emotions and help measure the effectiveness and return of your investments (ROI) in your promotional campaigns. Drones are becoming increasingly popular autonomous means of delivery, particularly in "last mile" logistics.

- 3D Printing. Thanks to the use of 3D printing technology, supply chains are becoming much faster, stable, and flexible. 3D printing allows almost any company to create devices or parts of devices, spare parts from various materials directly at consumers', eliminating the need for delivery. 3D-printing shortens the time of manufacturing objects and costs, expands the possibilities of consumers by bringing them closer to the desired products.
- Digital Twin. Digital twin can significantly increase the operational efficiency of marketing logistics processes, ensure a constant flow of operational data, collect product and packaging data, identify potential flaws and trends, develop products and experiment with changes in their layout without any risks. Logistics centers can create digital twins and use them to test different scenarios and improve operational efficiency. This technology can be used to provide real-time delivery information, which will improve delivery times and impact the efficiency of autonomous vehicles.

5 CONCLUSIONS

Marketing logistics is an effective management concept, and its digitalization is an integral component of modern economic development, which contributes to increasing the efficiency of companies and maintaining their competitiveness.

Companies continue to develop and create improved products and provide improved services, constantly implementing the latest digital methods and marketing logistics tools. They are approaching an innovative digital future, better satisfying the needs of consumers, creating the most convenient and effective marketing logistics systems.

REFERENCES

[1] M. Christopher and H. Peck, "Marketing Logistics," Oxford: Butterworth-Heinemann, 2nd edition, 2003.

[2] P. Kotler, "Marketing Management: Analysis, Planning, and Control", Prentice-Hall, 1967.

[3] P.D. Converse, "The Other Half of Marketing," Proceedings of the 26th Annual Boston Conference on Distribution, pp. 310-314, 1954.

[4] P.F. Drucker, "The Economy's Dark Continent," Fortune, vol. 103, pp. 265-270, 1962.

[5] M. Christopher and H. Peck, "Marketing Logistics," Oxford: Butterworth-Heinemann, 1997.

[6] U. Juttner, M. Christopher, and J. Godsell, "A strategic framework for integrating marketing and supply chain strategies," The international journal of logistics management, vol. 21, no. 1, pp. 104-126, 2010. doi: 10.1108/09574091011042205.

[7] J. Managan and C. Lalwani, "Global logistics and supply chain management," Cornwall: Wiley, 3rd edition, 2016.

[8] V. Lagodiienko, I. Perevozova, L. Bakhchivanzhi, K. Ozarko, V. Milcheva, and O. Orlova, "The role of digital technologies in optimizing the functioning of the marketing and logistics system of the enterprise," Review of economics and finance, vol. 21, pp. 2019-2026, 2023.

[9] S. Vujović and S. Milosavljević, "The fate of marketing logistics in the conditions of the 4.0. revolution," Novi ekonomist, vol. 13, no. 26, pp. 76-83, 2019. doi: 10.7251/NOE1926076V.

[10] H.A. Al-Ababneh, M.A.A. Dalbouh, S.A.S. Alrhaimi, I.M. Siam, and T. Ibragimkhalilova, "Digitalization, innovation and marketing in logistics," Acta logistica, vol. 10, issue 4, pp. 615-624, 2023. doi: 10.22306/al.v10i4.440.

[11] J. Stock and D. Lambert, "Strategic logistics management," McGraw-Hill/Irwin, 4th edition, December 28, 2000.

[12] P. Golinska-Dawson, K. Werner-Lewandowska, K. Kolinska, and A. Kolinski, "Impact of market drivers on the digital maturity of logistics processes in a supply chain," Sustainability, vol. 15, 3120. 2023.

[13] L. Cadavid and A. Valencia-Arias, "Analysis and categorization of studies of digital marketing in small and medium enterprises," Intangible Capital, vol. 18 (2), pp. 263-289, 2022. doi: 10.3926/ic.1809.

[14] J. Macias-Aguayo, D. McFarlane, B. Schönfuß, and L. Salter, "A catalogue of digital solution areas for logistics SMEs," IFAC PapersOnLine, vol. 55-10, pp. 1828-1833, 2022. doi: 10.1016/j.ifacol.2022.09.664.

[15] N. Zamkova, I. Polishchuk, Yu. Dovhan, L. Dovhan, and R. Novytskyi, "Training of future logistics and supply chain managers: a competency approach," Financial and credit activity: problems of theory and practice, vol. 1 (48), pp. 427-440, 2023. doi: 10.55643/fcaptp.1.48.2023.3946.

Formation of Product Promotion Strategies of Companies Producing Computer Games

Olha Polous, Inna Mykhalchenko, Hanna Radchenko, Olena Lysenko and Kateryna Pysmenna
[1]*National Aviation University, 1, Liubomyra Huzara ave., 03058, Kyiv, Ukraine*
polousolgavictorovna@gmail.com, inna.mykhalchenko@npp.nau.edu.ua, hanna.radchenko@npp.nau.edu.ua, olena.lysenko@npp.nau.edu.ua, 6914653@stud.nau.edu.ua

Keywords: Computer Games, Market, Promotion, Strategy, Changes, Evaluation, Globalization.

Abstract: The research actuality of the study is due to the fact that it is vital to understand the economic implications of promotional strategies, companies can use modern insights to optimize their marketing budgets, ensuring that their promotional activities yield the highest possible return on investment. Aim of the paper is to evaluate the evolving strategies used by companies to promote their products for the release of computer games, aiming to identify effective digital marketing tools, understand regional and cultural marketing nuances, and uncover innovative promotional tactics. In order to achieve the research goal, general scientific methods of analysis and synthesis, logical and situational analysis and specific methods inherent in management sciences (subjective-objective approach, economic and statistical analysis, method of ranking, grouping, comparison, classification) were used. The practical and theoretical relevance of this study is that it can be valuable resource for understanding the complexities of video game promotion in the digital age. By analysing EA's approach and considering emerging trends, we aim to offer insights that can help companies adapt and thrive as the industry evolves.

1 INTRODUCTION

The gaming market is dynamic, with trends and consumer preferences shifting frequently, necessitating that companies continuously adapt their promotional strategies to maintain competitiveness. Understanding these changes is crucial for developing effective marketing campaigns that resonate with target audiences, ensuring that new game releases achieve maximum visibility and sales impact [8, 13]. The competitive nature of the gaming industry means that companies must continuously innovate their promotional tactics to stand out.

To analyze the dynamic gaming market, methods like trend analysis and content synthesis are crucial for identifying patterns and integrating insights from various promotional strategies. Logical analysis, including comparative and causal methods, helps evaluate the effectiveness of tactics and their impact on key performance indicators across demographics. Situational analysis, such as SWOT and PESTEL, examines internal strengths and external factors, allowing companies to tailor campaigns to regional and cultural contexts. Scenario planning on the example of some company enables businesses to prepare for market uncertainties by crafting adaptable strategies based on hypothetical shifts. Together, these methods provide a structured approach to understanding market trends, optimizing promotional tactics, and enhancing engagement with diverse global audiences.

Electronic Arts (EA) stands out as a successful example of implementing adaptive promotional strategies in the dynamic gaming market. The company excels at leveraging trend analysis and consumer insights to craft campaigns that resonate with their global audience, ensuring maximum visibility for new releases. EA's continuous innovation in tactics, such as pre-release hype, influencer collaborations, and tailored regional promotions, keeps them competitive in an ever-changing industry. By addressing player preferences and market demands, EA effectively maintains its position as a leader in engaging and expanding its audience base. Their ability to adapt to technological advancements, such as VR and cloud gaming, make this company a good example for modeling a scenario for improving the product promotion strategy of a company that produces computer games which can be adapted for other companies in the future.

2 GLOBAL COMPUTER GAMES MARKET: EVALUATION OF THE EFFECTIVENESS AND FEATURES OF THE FORMATION OF PRODUCT PROMOTION STRATEGIES OF COMPANIES PRODUCING COMPUTER GAMES

The video game industry has emerged as a cultural and economic powerhouse. Beyond mere entertainment, video games offer meticulously crafted universes designed to deliver immersive experiences. This immersive quality is a testament to the strategic artistry employed by video game developers and marketers. In 2020, the industry's financial prowess was undeniable, generating a staggering $159.3 billion in revenue, surpassing the combined revenue of the global film industry and North American sports industry [13]. This remarkable success necessitates a deeper exploration of the underlying strategies that fuel the video game market's explosive growth (Table 1).

Table 1: Strategic marketing stages used for promotion campaign for games.

Stage	Phase	Description
Stage 1	Pre-Launch Hype	Building anticipation through trailers, teasers, and influencer partnerships sets the stage for a successful launch
Stage 2	Launch Day Strategies	Strategic release timing and special promotions make the launch day a landmark event in the gaming world
Stage 3	Post-Launch Engagement	Fostering an active community and continuously delivering fresh content keeps the game relevant and engaging
Stage 4	Measuring Success	Tracking KPIs, analysing data, and adapting strategies ensure the game stays in tune with player preferences and market trends

Unlike a monolithic entity, gamers represent a diverse tapestry of preferences. Some relish heart-pounding action adventures, while others lose themselves in sprawling RPGs or the casual convenience of mobile gaming. Unveiling the art of video game product strategy begins with a deep understanding of the target audience. This necessitates delving into demographics – age, gender, gaming preferences, and digital habits. Unravelling these attributes empowers developers to tailor their strategies, ensuring the marketing message resonates deeply with the players they aim to captivate [5].

Most of promotional strategies around upcoming games are built around pre-launch stage. To achieve commercial success, developers and publishers must cultivate anticipation and excitement amongst potential players – a process that begins long before the official launch [6].

The pre-launch stage, also known as the pre-production marketing phase, is a crucial period in the video game development cycle. It encompasses all the activities and strategies employed by developers and publishers to generate anticipation and excitement for a game before its official release. This stage typically begins months, and sometimes even years, before the game hits store shelves or digital marketplaces [5].

After successfully passing the pre-launch stage, second stage of pre-registration begins, also being used as a strategy to gain interest of the main audience. Pre-registration, a relatively recent phenomenon in the gaming industry (emerging in the late 2000s with the rise of mobile gaming), allows players to express interest in an upcoming game before its official launch [6].

Launch day is all about reaching a critical mass of players. Timing is everything, with developers aiming for periods when players are most active and competition from other major releases is minimal. \A strong marketing push, encompassing social media campaigns, targeted online ads, and partnerships with gaming influencers, ensures the game reaches its target audience. Livestreams showcasing gameplay and developer Q&A sessions can further build excitement and anticipation on launch day itself.

The journey doesn't end with launch day. Effective post-launch strategies are essential for retaining players, fostering a loyal community, and ensuring the game's long-term success. A well-orchestrated launch campaign elevates a game's visibility, while continuous efforts to retain players through compelling content and community engagement ensure a thriving and long-lasting gaming experience for everyone involved [5, 6].

The last step in this process is always measuring success, which in video game marketing requires a multifaceted approach, encompassing both quantitative and qualitative assessments. At the core lie key performance indicators (KPIs), which provide valuable insights into a game's performance and player engagement. A comprehensive evaluation of

an online promotion strategy requires a multifaceted approach, encompassing various metrics categorized into three key areas, presented in Table 2.

Table 2: Three key areas for evaluation of online promotion strategy.

Metric	Description
Reach and Awareness	1. Website Traffic. Tools like Google Analytics track unique visitors, page views, and traffic sources (organic search, social media referrals, paid advertising clicks), offering insights into audience demographics and campaign reach. 2. Social Media Engagement. Analysing likes, shares, comments, and mentions helps quantify the level of audience engagement with the promotional content. 3. Brand Mentions and Online Reviews.
Engagement and Interest	1. Click-Through Rate (CTR). CTR measures the percentage of users who click on a call to action (CTA) within a promotional message. 2. Time Spent on Site (the average time users spend on a webpage or landing page). 3. Bounce Rate. The bounce rate reflects the percentage of visitors who leave a webpage after viewing only one page.
Conversions and Sales	1. Conversion Rate. The conversion rate measures the percentage of website visitors who complete a desired action, such as making a purchase, subscribing to a newsletter, or downloading a white paper. 2. Sales Revenue Generated. 3. Customer Acquisition Cost (CAC). CAC measures the average cost of acquiring a new customer through the online promotion campaign.

As it can be seen from a Table 2, divided in three groups, evaluation can be started from analysing reach and awareness, which can be seen through website traffic (monitoring website traffic provides a fundamental understanding of the campaign's ability to attract visitors), social media engagement (SME platforms provide valuable metrics for gauging campaign awareness), brand mentions and online reviews (monitoring brand mentions across the web, including social media platforms, online forums, and review sites, provides valuable insights into brand awareness and sentiment) [1, 3]. Positive mentions indicate successful awareness generation, while negative mentions necessitate intervention and potential adjustments to the promotion strategy.

Beyond individual metrics, leveraging data analysis techniques provides a deeper understanding of campaign performance and audience behaviour. Dividing the audience into distinct segments based on demographics, interests, or online behaviour allows for targeted analysis [9]. By analysing metrics within each segment, marketers can identify which demographics or behaviours respond most favourably to specific promotional messages and channels [12]. Attribution modelling assigns credit for a conversion across various touchpoints within the customer journey and A/B testing involves presenting different versions of a promotional message, landing page, or website banner to a portion of the target audience [11].

Evaluating the effectiveness of an online promotion strategy is an ongoing process. By leveraging a combination of metrics, data analysis techniques, and real-world examples, companies can gain valuable insights into campaign performance and audience behaviour [10]. This data-driven approach empowers marketers to continuously optimize their online promotion strategies, ensure maximum reach and engagement, and ultimately, achieve their business objectives.

The COVID-19 pandemic has translated to significant financial gains for video game companies. Nintendo, for example, reported a staggering $1.4 billion in second-quarter profits, a fivefold increase compared to the same period in 2019. This surge is largely attributed to the phenomenal success of "Animal Crossing: New Horizons," a social simulation game experiencing global popularity since its March launch [2].

Electronic Arts' renowned sports franchise, FIFA, witnessed a doubling of new players in the second quarter compared to the previous year, adding 7 million users according to Bloomberg Intelligence. Similarly, "NBA 2K20," a popular basketball game, saw a remarkable 82% increase in active players during the same period. These statistics suggest that for many, virtual sports experiences served as a substitute for the lack of real-world competition [8].

In 2022, the leading revenue-generating markets and countries were the United States, China, Japan, South Korea, Germany, the United Kingdom, France, Canada, Italy, and Brazil. Despite the United States leading in revenue with $46.4 billion, China boasted the highest number of players, totalling 696.5 million (Fig. 1) [4]. This highlights the varying dynamics between revenue potential and player volume across regions, emphasizing the need for companies to tailor

their strategies accordingly using opportunities in both high-revenue and high-engagement markets.

Analysing the game industry, it is important not only to study the main regions, where this field improves the most, but also provide a critical examination of platform preferences within the game development landscape for the year 2023 (Fig. 2). The data incorporates two fundamental metrics: developer interest and current development.

Figure 1: Top 10 countries by game revenue, 2022 [4].

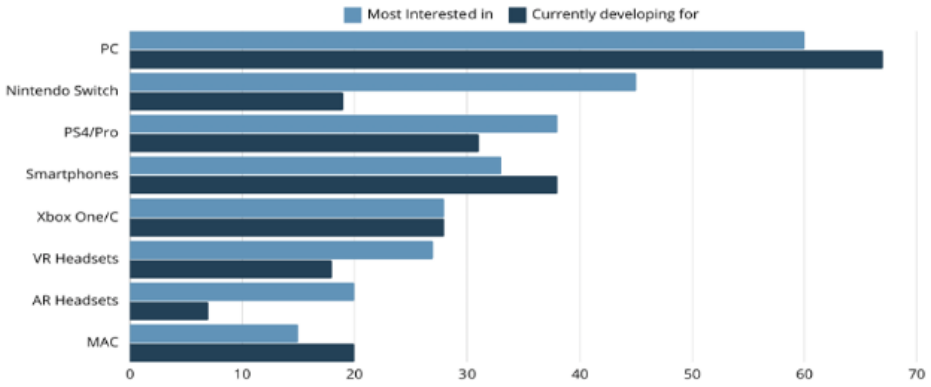


Figure 2: The most important Gaming Platforms in % of game developers most interested/currently developing in the following platforms [13].

Figure 3: Net Bookings Results of EA company by composition [7].

By meticulously dissecting these two metrics, this table aims to provide a comprehensive understanding of the platforms that are most attractive and actively pursued by game developers. As it can be seen, the platform with the highest percentage of developers interested in developing for it is the PC, at 39%. This is followed by the Nintendo Switch (37%) and PlayStation 4/Pro (35%). Interestingly, these three platforms also hold the top three spots for platforms that developers are currently working on, with PC at 33%, Switch at 31%, and PS4/Pro at 29% [13].

Founded in 1982, Electronic Arts (EA) is a preeminent American video game developer and publisher. The company has established itself as a global leader in interactive entertainment, boasting widely recognized franchises like Madden NFL, EA SPORTS™ FIFA, Battlefield™, and Dragon Age™. EA's core business model encompasses the development, marketing, publishing, and distribution of these popular video games [7].

Electronic Arts (EA) reported solid financial results for fiscal year 2023, with net revenue reaching $7.42 billion, reflecting a 6% increase year-over-year. However, a closer look to Figure 3 reveals a shift in revenue sources [7].

Despite the decline in full game sales, EA's net income still managed to grow by 1.6% to $802 million. However, net bookings, a metric that combines net revenue with the change in deferred net revenue (revenue recognized but not yet earned), came in at $7.3 billion, a 2.3% decrease year-over-year. This indicates that while some players are engaging with live services, overall spending on new games might be slowing down.

Beyond its impressive portfolio, EA is guided by a well-defined strategic framework cantered on three core pillars: Players First, Commitment to Digital, and One EA. The Players First approach emphasizes cultivating strong and lasting relationships with players by ensuring consistent product quality and fostering engagement with established franchises. Recognizing the growing shift towards digital game sales, the Commitment to Digital pillar underscores EA's dedication to expanding its offerings of live services for customers. Finally, the One EA model leverages the company's extensive scale to bolster its player network and streamline the development pipeline, fostering greater efficiency.

3 PRACTICAL RECOMMENDATIONS FOR IMPROVING STRATEGY OF PROMOTING THE PRODUCTS OF THE COMPANY FOR THE RELEASE OF COMPUTER GAMES (ON THE EXAMPLE OF ELECTRONIC ARTS ACTIVITIES)

Electronic Arts EA utilizes a freemium model where the base game is free, attracting players with a low initial barrier to entry. This allows them to try the core gameplay before potentially spending money on additional features and content. High-quality, immersive experiences are a core focus for EA. They invest heavily in research and development to create visually stunning, technically advanced games with engaging gameplay, aiming to deliver unforgettable experiences that keep players hooked Microtransactions are a significant revenue driver for EA. Players can purchase in-game items, virtual currency, or cosmetic enhancements, allowing for customization and offering a monetization strategy beyond the initial game purchase [7].

Reaching a broad audience is crucial for EA. They employ extensive marketing and advertising efforts across various channels, including traditional and digital platforms, to maximize game exposure and attract a large player base. However, it's important to note that EA has faced criticism for their use of loot boxes and in-game purchases, with concerns regarding the potential promotion of gambling habits among players.

SWOT Analysis of Electronic Arts (EA) will examine Electronic Arts (EA) through the lens of a SWOT framework (Table 3).

Understanding a company's competitive landscape is crucial for strategic decision-making. From the table below, it can be seen that EA boasts several advantages that contribute to its success – its strength is mostly shown in powerhouse franchises, financial might and marketing budget that allows the company to have comprehensive campaigns and broad audience reach:

- blockbusters like FIFA and Madden consistently generate significant revenue, building a dedicated fanbase and forming a solid foundation for future growth;

- with a robust financial performance exceeding EA can invest heavily in marketing and maintain a strong market presence;
- EA has significant marketing budget.

Table 3: SWOT analysis of Electronic Arts (EA).

Strengths	Weaknesses
Strong game franchises (FIFA and Madden). Financial Might/Robust financial Performance. Significant marketing budget.	Community Backlash. Criticisms around microtransactions and perceived lack of innovation have tarnished EA's reputation, potentially hindering marketing success. Franchise Dependence. While franchises like FIFA are strong, overreliance on a few titles can expose EA to risks if popularity declines or competition intensifies.
Opportunities	**Threats**
Booming Market. The gaming industry's explosive growth, with projections reaching 3.3 billion players by 2024 (up from 2 billion in 2015). Differentiation is Key. Focusing on creating unique and engaging content will be paramount. Embracing Diversity.	Cutthroat Competition. Because of the popularity of industry, company faces a lot of strong competitors like These include companies like Activision Blizzard, Take-Two Interactive and Ubisoft. Evolving preferences of consumers.

Electronic Arts employs a multifaceted promotional strategy that has achieved considerable success. An evaluation of its effectiveness, considering both strengths and potential areas presented in a Table 4.

Overall, EA's promotional strategies are demonstrably effective in reaching a broad audience, generating excitement for new releases, and building brand recognition. However, continuous adaptation and a focus on player satisfaction are crucial to maintain success in the ever-evolving gaming industry.

So, EA boasts a long and successful history in the gaming industry. However, recent controversies surrounding microtransactions and perceived stagnation in some franchises have dented player trust. Future promotion strategy aims to address these concerns while leveraging the company's strengths to create a more dynamic and engaging promotional landscape (Fig. 4).

Table 4: Evaluation of promotional strategy for EA's products.

Strengths	Potential Areas for Improvement
Targeted Approaches. EA tailors its marketing efforts to specific player demographics. Multi-Channel Marketing. EA utilizes a diverse range of channels, including traditional advertising (TV, print), digital marketing (online banners, social media), partnerships (influencers, celebrities), and data-driven targeting. Community Building. By fostering a sense of community through events, forums, and social media engagement, EA builds trust and loyalty among players. Innovation and Creativity. EA's willingness to explore innovative approaches that grab player attention.	Microtransaction Backlash. Criticism surrounding loot boxes and aggressive monetization strategies can alienate players. Finding a balance between generating revenue and maintaining player satisfaction is crucial. Negative Stereotypes. Marketing tactics that perpetuate negative stereotypes regarding gamers (e.g., excessive spending) can damage brand image. It's important to portray gamers in a positive and inclusive light. Transparency and Communication. Clear communication regarding in-game purchases and content updates can help build trust with players. Addressing concerns and offering explanations can alleviate frustration.

Having analysed all the information, the most obvious recommendations will be rebuilding trust with such a huge fanbase, focusing on transparency in monetization (clearly communicate the value proposition of microtransactions and focus on cosmetic or convenience-based items), focus on quality and content (deliver polished experiences with engaging gameplay and a steady stream of meaningful post-launch content) and open communication (actively engage with player communities, address concerns promptly, and showcase responsiveness to feedback).

Tracking the number of new players joining EA games via the freemium model revealed an initial surge, but retention rates showed a drop-off due to limited early-game incentives. So, EA introduced tailored onboarding campaigns and highlighted compelling features of their premium content in advertisements. They also used free trials for premium features, boosting conversion rates.

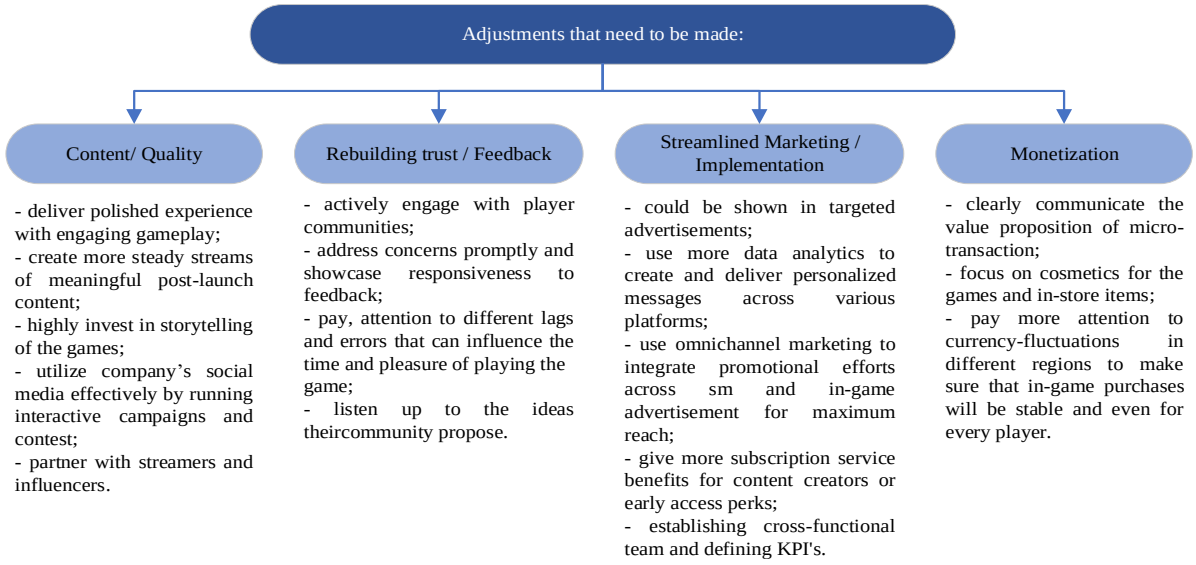


Figure 4: Recommendations for improving promotion strategy of Electronic Arts company.

Revenue from microtransactions (MTX) was high, but a significant portion of feedback from KPIs like Net Promoter Score (NPS) indicated dissatisfaction with perceived "pay-to-win" mechanics. EA shifted the messaging around microtransactions to emphasize cosmetic and convenience upgrades rather than gameplay advantages. They also implemented transparent communication about the value players received for their purchases, increasing player trust.

Metrics from social platforms showed high engagement during release cycles but declining interaction during off-peak periods. EA ramped up community-building initiatives, such as interactive social media campaigns, influencer partnerships, and exclusive behind-the-scenes content, that kept players engaged year-round and fostered a more loyal community.

While flagship franchises like FIFA and Madden consistently generated revenue, player sentiment surveys highlighted franchise fatigue and a desire for innovation. So, EA began investing in experimental projects, broadening their genre portfolio and integrating new technologies like cloud gaming and VR/AR. These efforts aimed to attract new audiences while rejuvenating interest among long-time fans.

So, on our opinion, EA effectively adapted its strategies by addressing player retention, refining microtransaction practices, enhancing community engagement, and investing in innovation to sustain loyalty and attract new audiences. Additionally, the company invested in experimental projects and new technologies to rejuvenate flagship franchises and attract a broader audience.

4 CONCLUSIONS

The year 2020 marked a turning point for the gaming industry. As lockdowns and social distancing measures swept across the globe, people turned to video games for entertainment, connection, and a much-needed escape. This surge in demand propelled the industry to unprecedented heights, fostering a sense of community and shared experiences within virtual worlds. The rise of online multiplayer games further amplified this phenomenon, creating a vibrant digital space where players could connect, collaborate, and forge lasting bonds.

This newfound appreciation for gaming presented a unique opportunity for industry leaders like Electronic Arts. However, capitalizing on this momentum required a keen understanding of evolving player preferences. Gamers today crave not just high-octane graphics and adrenaline-pumping action, but also immersive experiences that foster a sense of belonging. High-quality content, delivered with a focus on player satisfaction, became paramount.

The gaming industry, like any other, thrives on trust. Players invest not just money, but also time and emotional attachment into the games they choose. In recent years, some of EA's marketing strategies may

have inadvertently eroded this trust. Moving forward, prioritizing exceptional game experiences that meet player expectations is essential. Consistent delivery of high-quality content, addressing player feedback with genuine concern, and fostering open communication are all crucial steps towards rebuilding a strong and loyal player base.

Data analytics plays a crucial role in understanding player preferences and tailoring marketing strategies accordingly. By analysing player behaviour within games, companies can gain valuable insights into what resonates with their audience. This data can then be used to personalize marketing campaigns, ensuring that players receive content that is relevant and engaging.

The market itself is brimming with potential, and EA is well-positioned to capitalize on this growth. By focusing on rebuilding trust, delivering high-quality content, fostering vibrant communities, and embracing new technologies, EA can reignite player passion and solidify its position as a leader in the ever-evolving gaming landscape. This shift towards a more player-centric approach will not only benefit EA's bottom line but also create a more positive and engaging experience for gamers worldwide. As the industry continues to flourish, EA has a unique opportunity to shape the future of gaming, fostering a community where players and developers alike can come together to celebrate the joy and power of interactive entertainment.

REFERENCES

[1] A. Hassan, “The Usage of Artifrcial Intelligence in Digital Marketing: A Review”, Applications of Artificial Intelligence in Business, Education and Healthcare, pp. 357–383, 2021, [Online]. Available: https://doi.org/10.1007/978-3-030- 72080-3_20.

[2] A. Epstein, “How pandemic turned everyone into a gamer?” [Online]. Available: https://qz.com/1904276/everyone-is-playing-video-games-during-the-pandemic.

[3] A. Stevanović, S. Mitrović and A. Rajković, “Application of information technologies and the internet in modern business”, Oditor, vol. 8(2), pp. 54-74, 2022, [Online]. Available: https://doi.org/10.5937/Oditor2202054.

[4] Countries by game revenue. [Online]. Available: https://knoema.com/infographics/tqldbq/top-100-countries-by-game-revenues.

[5] D. Singh and S. Pillai, “Implementation of digital marketing strategy in the gaming industry,” 2022 International Conference on Decision Aid Sciences and Applications (DASA), pp. 387–394, 2022.

[6] E. M. Torralba, “Playing games to earn money: The conceptual framework of interaction between gender, learning styles, problematic gaming behavior and success-economic gain motivation of playing games,” SSRN Electronic Journal, 2023.

[7] Electronic Arts: most recent reports. [Online]. Available: https://www.annualreports.com/Company/electronic-arts.

[8] Insights on the theme of how pandemic changed the video game market. [Online]. Available: https://www.bloomberg.com/professional/insights/.

[9] K. Dunakhe and C. Panse, “Impact of digital marketing – a bibliometric review”, International Journal of Innovation Science, vol. 14, no. 3/4, pp. 506-518, 2022, [Online]. Available: https://doi.org/10.1108/IJIS-11-2020-0263.

[10] O. Troisi, G. Maione, M. Grimaldi, F. Loia, “Growth hacking: Insights on data-driven decision-making from three firms”, Industrial Marketing Management, vol. 90, pp. 538-557, 2020, doi: 10.1016/j.indmarman.2019.08.005.

[11] S. Smerichevskyi, O. Polous, I. Mykhalchenko, L. Raicheva, “Formation of marketing competencies in case of startups integration into the intellectualized market space”, Modern Trends in Financial and Innovation Data Processes, Springer Nature, pp. 43-60, 2023, [Online]. Available: https://doi.org/10.1007/978-3-031-54012-7_3.

[12] Statista Research Department. Digital advertising worldwide - statistics & facts. 2023. [Online]. Available: https://www.statista.com/topics/7666/internet-advertising-worldwide/#topicOverview.

[13] Statista. Games – Worldwide. [Online]. Available: https://www.statista.com/outlook/amo/media/games/worldwid.

Development of an Information Management System for it Enterprise's Intellectual Potential

Olena Ivanova[1], Oleg Ivanov [2], Valentyna Ivanova [3] and Alona Kashpruk[4]
[1]*Department of Digital Economy and System Analysis, State University of Trade and Economics, Kyoto Str. 19, 02156 Kyiv, Ukraine*
[2] *Department of Civil Engineering and Professional Education, Poltava State Agrarian University, Skovorody Str. 1/3, 36003 Poltava, Ukraine*
[3]*Department of Economy, Entrepreneurship and Management, Educational and Scientific Institute of Management and Psychology, University of Educational Management, Sichovykh Striltsiv Str. 52-A, 04053 Kyiv, Ukraine*
[4]*Department of Digital Economy and System Analysis, State University of Trade and Economics, Kyoto Str. 19, 02156 Kyiv, Ukraine*
(o.ivanova@knute.edu.ua, oleg.ivanov@pdau.edu.ua, ivanval2177@gmail.com, business2018@i.ua)

Keywords: Intellectual Potential, Intellectual Capital, Knowledge Management System, IT Enterprise Management.

Abstract: The current requirements of the IT market to the speed of decision-making, the high degree of innovation in end products, and the rapid development of technology necessitate the consideration of a system for managing the intellectual potential of employees that would contribute to the formation of competitive advantages, innovation activity support, and intensive knowledge sharing. The purpose of the article is to develop practical approaches to the assessment of intellectual potential and knowledge management and offer organizational changes and platform to implement information management system for IT enterprise's intellectual potential. Considering influence factors and assessment criteria of information potential, the article provides the results of assessment of intellectual potential of an IT enterprise. Assessment of knowledge management system demonstrates the readiness of staff to organize effectively intellectual assets and showed operational areas for improvement. Constancy of qualified personnel, innovative activity, implementation of inventions, professional training and other indicators allowed measuring the effectiveness of the IT enterprise's intellectual potential. Such an assessment will enable the IT enterprise to understand the current state of its intellectual assets and the ways to improve intellectual capital. Based on the results obtained, a prototype of an information system and organizational structure was formed to better manage the intellectual capital of an IT enterprise.

1 INTRODUCTION

Well-organized and properly maintained intellectual potential (IP) management of an IT enterprise allows a company to stay sustainable and productive, create value, and meet organizational goals. The activities of an IT company are inextricably linked to intellectual innovation, which produces information results based on knowledge. It is of great importance that intellectual assets, which contain an enterprise's employee knowledge, skills, training programs or any proprietary information that may provide the company with a competitive advantage, should be organized appropriately. The purpose of the study is to analyze theoretical and practical aspects of IT enterprise's intellectual potential management, and develop the prototype of an information management system of IT enterprise's intellectual potential with a preliminary theoretical substantiation and empirical study of the effectiveness of the existing systems at the enterprise under research.

2 THE CONCEPT OF AN ENTERPRISE'S INTELLECTUAL POTENTIAL

Software enterprises in particular face a huge challenge complicated by the need to align the rapidly evolving technologies with the business objectives. The intellectual potential is the main intangible

resource that contributes to the effective development of the enterprise. Nowadays many Ukrainian and foreign scientists pay a lot of attention to the category of intellectual potential of an enterprise and its structural components.

Defining the intellectual potential of a company (or enterprise), foreign economists Stewart T., Edvinsson L., Swibee K., Petty R., and Guthrie J. identify knowledge as a source of additional competitive advantages. They assumed that the intellectual potential of the company represents the scope of the new knowledge, which it can use to enhance its competitive ability [11]. Shepelenko S.M. determined the intellectual potential as The author notes that intellectual potential includes intangible assets, such as knowledge, human capital, information technology, relationships with customers and partners, which are an important resource for creating value and ensuring the successful development of an enterprise [10].

J. Džino, B. Latinović, and Z. Avramović provided the example of how structural knowledge bases enhance decision-making processes in monitoring on the IT enterprise activity [6]. O. Sobko and S. Stakhurska defined intellectual potential as "..an opportunity to create intellectual added value by setting up the production and commercialization of intellectual-intensive products, which ensures innovative business development" [7]. Many other scientists also identify knowledge, results of intellectual labour, and creative abilities of employees among the defining features of an enterprise's intellectual potential. This is especially true for an IT enterprise, especially a product company, since the success of the application of such IP directly determines the performance and market success of such a business.

Many different approaches to the concept of IP rely on its structure. Thus, the most widespread general approach to the defining intellectual potential is its division into structural, information, human and relationship potential (Table 1).

Each approach reflects the specifics of human resources, corporate resources, and relationships with stakeholders, which are sources of additional knowledge value for the enterprise. The multifaceted characteristic of human potential contributes to the expansion of sources of innovative ideas, the development of various skills and the formation of organizational knowledge.

Yemialyanau A. states that the concept of intellectual potential is similar to the concept of intellectual capital, but if elements of intellectual capital are involved in the activities of the enterprise and bring in income, then the elements of intellectual potential are not fully involved in the enterprise's activities and at a certain point in time do not bring income, but have the ability to bring it in the future [1].

Table 1:.Structure of intellectual potential (based on [2, 3, 7, 9]).

Intellectual potential	Structural potential	Includes officially recorded knowledge, legally protected in patents and trademarks, resources represented by software and formal processes, and the knowledge recorded in instructions, manuals, and other written forms.
	Information potential	Represents the value of information resources specific to the enterprise, including databases, data intelligence systems and management information technologies provided for informed decision-making.
	Human potential	Indicates the level of education, professionalism, qualification, skills, workers' innovation culture, critical thinking, creativity, and self-development, reflected in the company's capacity to remain current, advance and generate value-added assets.
	Relationship potential	Indicates the capacity to establish a direct connection with the market, encompassing relations with suppliers, contractors, and partners, and the strength of customer loyalty as the key determinants of a business's current and future potential for generating revenue and strengthening the brand.

Taking into account the parallels and importance for the enterprise, both concepts ensure the achievement of ultimate goals and the leveraging of intangible and human resources.

There are factors which affect intellectual potential development of the enterprise, both negatively and positively, restrain or push the growth of the enterprise's development. They relate to innovation and intellectual work at the IT enterprise, internal resources, financial opportunities, innovation activity, research & development initiatives, qualitative characteristics of labor resources. Each of the factors play an important role consequently in fostering or impeding IP progress within the domain. We highlight the following factors that will stimulate and constrain the development of IP for the IT enterprise (Table 2):

Table 2: Factors influencing on the development of intellectual potential[1].

Factors influencing on the IP development	
1) stimulate development	2) restrain development
• -available intellectual property rights; • -high profitability; • -high-quality scientific and technical developments; • -qualitative and clear approaches to the IP assessment; • -innovation and intellectualization of output products; • -highly qualified staff.	• -lack of intellectual property rights; • -insufficient investment; • -lack of scientific and technical developments; • -lack of assessment of intellectual potential; • -insufficient level of specialists' experience.

3 ASSESSMENT OF KNOWLEDGE MANAGEMENT

Knowledge management is important for building intellectual capacity and the factors that contribute to its development. Knowledge management software facilitates the organization of information flows, enhances customer support, and fosters the development and preservation of intellectual potential. To assess the efficacy of the installed knowledge base, a survey (n=50) was conducted on a small IT enterprise to determine the qualitative possibilities of forming the knowledge base of the enterprise's IP. The nine parameters were selected based on their relevance to the research object. Depending on the level of familiarity with the knowledge management system (KMS), respondents were divided into three categories (advanced, intermediate, and introductory levels). Considering the experience of using the KMS, respondents used a 5-point Likert scale (strongly agree, agree, undecided, disagree, strongly disagree). The average values of each parameter according to the status of the respondents are depicted (Fig.1).

The results of our study of existing KMS indicated a high quality of knowledge existing in the enterprise (admitted by 52% of respondents), accurate knowledge captured in the enterprise (62% of respondents). According to the average value, the lowest values were given to the parameters "KMS has increased innovation in procedures", "KMS is user friendly", and "KMS provides development of knowledge". This suggests that the KMS is not operating effectively in certain respects and may require enhancement or replacement with a more efficacious alternative.

It is worth noting that the assessment of experienced users seems to be more positive, in particular with regard to the high level of knowledge accuracy, search quality and prospects for the development of the KMS. Nevertheless, experienced users note that the KMS does not have a sufficient impact on innovation and is not sufficiently embedded in their responsibilities. All categories agreed that a knowledge management system is essential to the success of an IT enterprise.

4 INTELLECTUAL POTENTIAL ASSESSMENT

In addition to the analysis of KMS efficiency, it is essential to examine the effectiveness of intellectual potential management within the enterprise. This entails a detailed investigation of intellectual capital and knowledge management systems, which represent two pivotal aspects of intellectual potential. We have applied the method of integrated assessment of intellectual potential by using the methods of expert survey and mathematical statistics.

The system of indicators consists of the following coefficients:

- coefficient of qualified personnel constancy;
- coefficient of innovative activity;
- coefficient of implementation of inventions (offers);
- coefficient of staff coverage by professional training;
- coefficient of engineering, technical and scientific support;
- coefficient of education level of personnel .

The indicators were selected taking into account the contribution of personnel qualitative characteristics, innovation activity and efficiency, scientific support of the company's current activities, trainings and professional training programs, etc. For the IT sector, acquiring new knowledge and honing practical skills is important given the rapid changes in technology and stakeholder requirements.

The survey was conducted among the managers, and the weight of each factor was determined. The consistency of expert opinion on each component indicator of intellectual potential was evaluated using the coefficient of variation. After normalization, the integrated indicator of the enterprise's intellectual potential was determined (Table 3).

[1] developed by authors

[2] based on 4, 8, 11

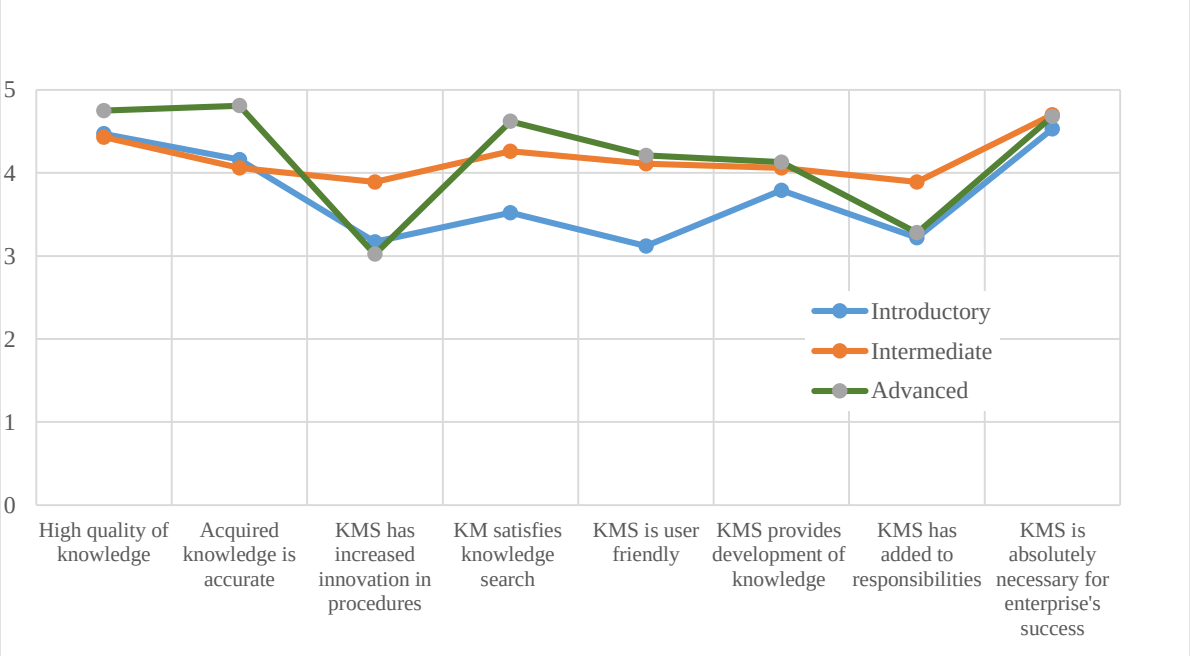


Figure 1: Qualitative estimate of KM system, by acknowledgement level [3].

Table 3: Expert assessment and determination of consistency of experts' opinions on partial indicators of intellectual potential [4].

Indicators	Average (χ_k, ω_k)	Standard deviation (S_k)	Coefficient of variation (V_k)
Coefficient of qualified personnel constancy	0,23	0,0136	5,92
Coefficient of innovative activity	0,16	0,0346	22,06
Coefficient of implementation of inventions (innovation offers)	0,17	0,0242	14,38
Coefficient of staff coverage by professional training	0,20	0,0140	7,07
Coefficient of engineering, technical and scientific support	0,12	0,0199	16,05
Coefficient of education level of personnel	0,12	0,0368	29,69

The coefficient of qualified personnel constancy is considered to be the most significant one (0,23), and the coefficients of engineering, technical and scientific support and educational level of personnel have the minimum specific weight (0,12).

Given the concordance coefficient W=0,721 and checked it by the Pearson's chi-squared test ($\chi 2$=36,057), it is higher than the tabular value of $\chi 2$ (p = 0,05; k = 4 - 1 = 3), which is equal to 11,07, while the multiple rank correlation coefficient exceeds the mark of 0,7. It leads to the conclusion about the reliability of the data obtained by experts, and the feasibility of their further use for analytical purposes. The obtained results showed the increasing the level of motivation of skilled workers and stimulating creative initiatives of staff to improve the intellectual potential of the IT enterprise.

The overall assessment of the IP in terms of its structural indicators was carried out based on a qualitative study in accordance with the structural, human and user components. The number of indicators selected for the respective components was estimated based on the results of the survey, and the corresponding average scale values and ranks were obtained. In the top eight most useful indicators we observe seven indicators that are currently used (Table 4).

[3, 4] developed by authors

Table 4: Estimated results of survey: all indicators [5].

All Indicators	Category	Average Value	Current Use Ranking	Usefulness Ranking
Customer satisfaction	Customer	4,78	4	1
Employee motivation	Human	4,62	7	2
Growth in business or service volume	Customer	4,54	3	3
Leadership skills (managers)	Human	4,41	5	4
Employees' information technology literacy	Human	4,38	8	5
Satisfaction of employees	Human	4,30	2	6
Number of customer complaints	Customer	4,25	9	7
Years of experience	Human	4,20	1	8

Based on the obtained qualitative results of assessing the effectiveness and importance of the factors that determine the IP of the IT-enterprise, we offer the IP management organization intending to reflect the IT business specifics, to follow consistency of goals and resources of IP, to enhance data collection timely, to centralize knowledge management, and to leverage intellectual resources. It is believed that parallel organisational structure aligned to the existing structure will contribute to a better understanding of IP and its optimal use. IP management should be dissolved among senior management positions and be one of the functions of the staff. In addition, building a line management (directorate or department) responsible for IP management will strengthen the processes of identifying, protecting, monitoring, evaluating and promoting IP in the knowledge-intensive and information-rich environment of the enterprise IT according to the structural parts of intellectual potential (Fig. 2).

Although the introduction of such an organisational change may increase the cost of human and organisational resources, such a focus will ensure the important development of significant intellectual potential and future competitiveness and profitability. The implementation of such organisational change should be preceded by the active development of the IT enterprise, which should reach a new qualitative level of activity.

Some aspects of the chosen IP management structure such as double subordination of employees, the complexity of monitoring the implementation of tasks, low flexibility and risk of inconsistency of goals and resources for their implementation can become the main reasons of possible risks occurrence while using this structure.

The specific approach of the IT enterprise to creating better software products is a software process improvement on knowledge-, competencies-, and experience of working practitioners in the enterprise. In addition, the enterprise cannot ensure the transition to innovative production, as well as the implementation of some effective activities without the appropriate level of functional development of intellectual potential. Considering this fact and addressing the issues revealed in the abovementioned qualitative research, the prototypes of functional properties of the IP management platform for IT enterprises are offered (Fig. 3).

The functionality of the IP platform should be adjusted for remote work of the IT enterprise and deployed on Android and iOS operating app systems. Informational organization of the platform includes organization of IP-related indexing information, keeping accurate records of intellectual potential, maintenance and using of project artifacts, collection of various content (article, external links on useful material, videos, trainings, etc.), quality control, flexible educational content according to the learning needs, onboarding information, etc. The single center for the management of IP and establishing control over the platform will ensure its higher efficiency. The possibility of rapid data collection at the functional level and their analysis at the centralized level as one of the key advantages of the approach becomes possible due to the implementation of the platform. The offered functional composition of the platform provides comprehensive framework, which can be adjusted to the specifics of various IT companies and their organizational needs with refined changes during implementation. Future stages include meticulous testing of functions, iterative development, and consequent evaluation to adjust the platform with the business goals and expectations of users.

[5] developed by authors

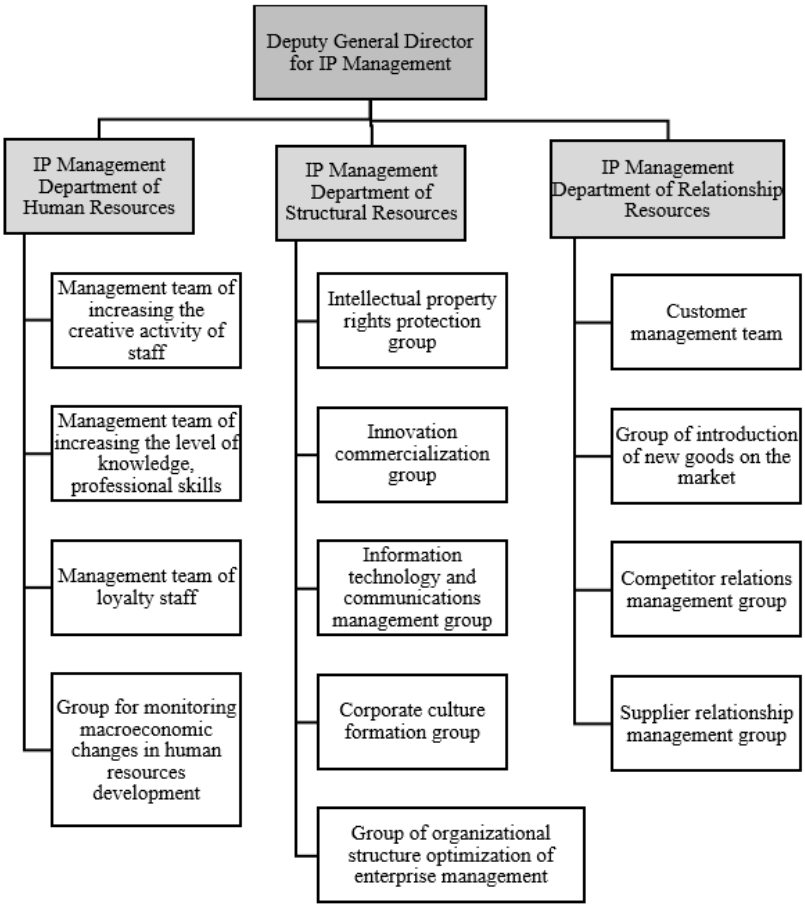


Figure 2: Organizational structure of IP management [6].

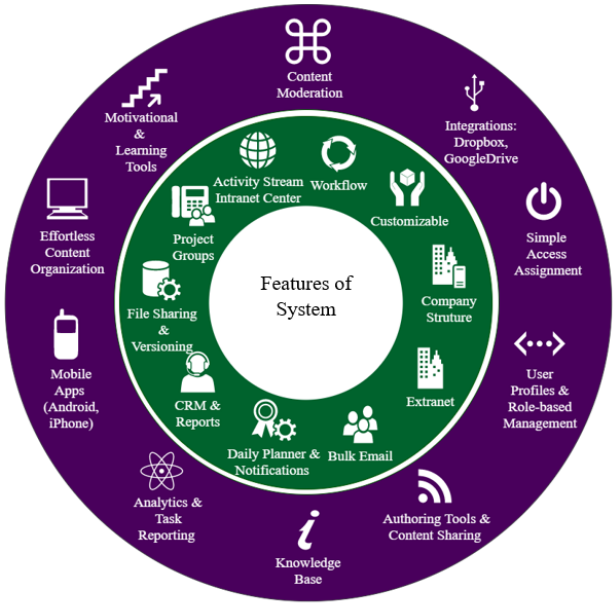


Figure 3: Functional properties of the IP management platform [7].

[6, 7] developed by authors

5 CONCLUSIONS

Because of the theoretical foundations research of the enterprise's intellectual potential issue the general essence, main functions and constituent elements of IP were revealed. The vast majority of scientists define enterprise's intellectual potential as the ability to generate future economic profits from existing intangible resources. The most widespread general approach to defining intellectual potential is its division into human, structural, information and relationship potential. The factors of the enterprise's intellectual potential development were divided into the stimulating and restraining groups. Due to qualitative assessment of knowledge management system, knowledge quality and accuracy were determined as the most prominent indicators of its factual efficiency on the IT enterprise.

In addition, the range of intellectual potential indicators was used to identify increased levels of motivation and creative initiatives as promising drivers of IP improvement in the IT organization. Within the framework of the structural approach to IP, a number of indicators have been identified that will determine the status and efficiency of IP use at the enterprise IT. Taking into account the identified indicators and factors of influence, it was offered to introduce an additional organizational unit and IP management functions into the existing structure. At the same time, the functionality of the information platform is outlined, which serves as the basis for the development and centralized management of IP of an IT enterprise.

REFERENCES

[1] A. Yemialyanau, "Intellectual potential as a factor of effective development of the Enterprise," INDIVIDUAL. SOCIETY. STATE. Proceedings of the International Student and Teacher Scientific and Practical Conference, p. 113, May 2017. doi:10.17770/iss2017.3031.

[2] A. Y. Bulatetskaya, "Risks in the management of corporate intellectual potential," International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 3, pp. 1744–1749, Sep. 2019. doi:10.35940/ijrte.c4596.098319.

[3] A. Z. Gafiyatullina, A.V. Pavlova, M.V. Vesloguzova, "Structure and development of the intellectual potential of the Industrial Enterprise Personnel," Review of European Studies, vol. 7, no. 1, Dec. 2014. doi:10.5539/res.v7n1p117.

[4] C. Devece, D. Palacios, and D. Martinez-Simarro, "Effect of information management capability on Organizational Performance," Service Business, vol. 11, no. 3, pp. 563–580, Jul. 2017. doi:10.1007/s11628-016-0320-7.

[5] G. P. Susila, I. B. Purbawangsa, H. Rahyuda, and N. Artini, "Competitive advantage and intellectual capital as mediators of the influence of corporate governance on firm value," Jurnal Ilmiah Akuntansi, vol. 8, no. 1, pp. 206–224, Aug. 2023. doi: 10.23887/jia.v8i1.59442.

[6] J. Džino, B. Latinović, and Z. Avramović, "Making decisions in monitoring by using decision-making method, knowledge bases and new IT solutions," JITA - Journal of Information Technology and Applications (Banja Luka) - APEIRON, vol. 21, no. 1, Sep. 2021. doi:10.7251/jit2101033d.

[7] O. Sobko and S. Stakhurska, "Development of the intellectual potential of the enterprise in the conditions of modern changes," Economic Analysis, no. 33(2), pp. 194–201, 2023. doi:10.35774/econa2023.02.194.

[8] O. M. Gorelik and Y. I. Koroleva , "Expert Evaluation of Information Potential Efficiency by Business Structures," International Business Management, vol. 10, no. 10, pp. 1939–1942, 2018. doi:10.3923/ibm.2016.1939.1942.

[9] S. Zambon, G. Marzo, L. Girella, M. Abela, and N. D'Albore, "A literature review on the reporting of intangibles", European financial Reporting Advisory Group, report, 2020.

[10] S. M. Shepelenko, "Management of the intellectual potential of enterprises in the context of globalization transformation: Theoretical aspects," Economic Bulletin of Dnipro University of Technology, vol. 85, pp. 187–196, Mar. 2024. doi:10.33271/ebdut/85.187.

[11] T. Cherkashyna and I. Lisna, "Methodological approach to the determination of efficiency of national intellectual capital use in the knowledge-based economy," Modern Economics, vol. 11, no. 1, pp. 160–166, Oct. 2018. doi:10.31521/modecon.v11(2018)-26.

Study of Digital Skills Using ICT in the Development of the Digital Economy

Oleksandra Tsyra, Yuliia Tochylina and Kateryna Tanashchuk
Department of Public Administration and Digital Economy, State University of Intelligent Technologies and Telecommunications, Kuznechna Str. 1, 65023 Odesa, Ukraine
aleksandra.tsyra@gmail.com, lula_2405@ukr.net, etanaschuk@ukr.net

Keywords: Competences, Competitiveness, Digital Skills, ICT, Human Capital, Knowledge and Abilities, National Economy.

Abstract: The active development of science and technology became the basis of significant changes, which provide for the rapid introduction of ICT in all spheres of economic relations and human life. Such innovations have led to a change in socio-economic activity, which today is defined by digital transformation. The development of the digital economy inevitably affects a person, the demands placed on his abilities, skills and knowledge. The hypothesis regarding the allocation of digital skills in the structure of human capital, the acquisition of which allows the effective use of new electronic services and ICT in work and everyday activities, is becoming dominant in the world. Therefore, we consider it relevant to study the essence and components of digital skills, which can become the basis for comparing their level in different countries. Accordingly, the purpose of this study was to reveal the theoretical and methodological foundations of digital skills and to find out the relationship between the level of digital skills of the population and the competitiveness of the country in the conditions of digitalization of its economy. To achieve the goal, several general scientific and specific methods of cognition were used, including analysis and synthesis, content analysis and the method of classification and analytical grouping, statistical method, economic forecasting. As a result of the study, it is proposed to distinguish the following components in the structure of digital skills: technical and computer skills, communication skills, security skills, media information and data analysis skills, and the ability to solve digital-based problems. It was revealed that the level of development of digital skills depends on the pace of development of the national economy, its competitiveness, the activity of implementation and dissemination of innovative activities. This assumption is confirmed by comparing DSGI (The Digital Skills Gap Index) and GCI (The Global Competitiveness Index).

1 INTRODUCTION

The active development of research and innovation activities led to the rapid introduction of information technologies into all spheres of human life. Such innovations radically change established connections and ways (mechanisms) of activity, and require relevant knowledge necessary for the use of information and communication technologies (ICT) in business, management, everyday life, etc. Gradual digitization and penetration of ICT into the economy are changing, firstly, the classifier of professions; secondly, the requirements put forward in the labour market, which, along with non-cognitive abilities, require new progressive skills that must correspond to modern types of labour activity; thirdly, the organization of everyday life, for example, participation in distance learning or e-government, online payment of utility services and regulation of digital finances, registration in the medical protection system, etc. It's required at least elementary skills in using computer and digital devices, that is, modern digital society needs a digitally competent population. Thus, today almost all aspects of work or daily activities relate to digital skills. Accordingly, the dominant point of view regarding the expediency of selecting certain skills in the structure of human capital, which are beginning to be called digital skills and become a key factor in determining uniqueness and professionalism, is becoming dominant.

Quarantine restrictions, because of Covid-19, the peculiarities of conducting economic activities in the conditions of the deployment of a full-scale war

proved not only the possibility of performing remote work, but also clearly demonstrated the advantages of the spread of electronic methods of communication, electronic trade, e-banking, etc. However, the community's further acceptance of the gains of digitalization and their active development is possible only based on the development of digital skills, since the population must possess at least a minimum amount of digital knowledge and skills that will allow them to use new electronic services.

2 LITERATURE REVIEW

The rapid pace of digital transformation of society and the economy not only leads to the intensification of the development of digital skills of the population, but also leads to their extensive study in the scientific literature, as evidenced by numerous publications. The analysis of the latter allows us to outline four main areas of research in this direction, in particular:

- clarifying the impact of digital skills on employment and income of the population, as well as their relationship with economic development in general;
- analyzing the digital divide between different regions or countries and outlining ways to overcome it;
- developing new training and educational programs that provide for the formation of digital skills and competencies;
- fixing the relevant digital skills in specific professions.

Individual elements of the first area are presented in the scientific works of Lura Rexhepi Mahmutaj and Nora Jusufi [1], which are aimed at studying the importance of digital skills for stimulating innovation. The authors conclude that taking specific measures to form and develop the digital skills of company employees is a strategic step that can increase competitiveness, growth, efficiency and sustainability in a business environment based on the use of digital technologies.

The second direction of research on digital skills is quite comprehensively revealed in the study of Sophie Lythreatis, Sanjay Kumar Singh, Abdul-Nasser El-Kassar [2], who define the digital divide as a significant social problem of our time, which has recently been significantly deepening. The authors identified nine factors that influence the gap and determined that among them, it is education and the formed digital skills that have the greatest impact on digital inequalities of different categories of the population or territories.

Presenting the third direction, the team of authors [3] focuses their research on clarifying the changes inherent in the education system after Covid-19. During 24 trainings on digital technologies, they determined that mastering digital technologies is associated with the digital competencies of scientists. Therefore, higher education institutions should implement a policy to form an innovative climate, which will allow the formation of digital skills of teachers themselves and thereby improve the educational process and promote their acquisition by applicants.

The last direction of studying digital skills is due to the active development of the digital economy, which puts forward radically new requirements for employees. In particular, the study by Sebastian Saniuk, Dagmar Caganova, Anna Saniuk [4] identified key areas of knowledge and skills of workers that are necessary for the implementation of the Industry 4.0 concept. Important attention was also paid to the need to combine enterprises with scientific institutions and technology parks, which will allow the formation of the necessary skills and bringing new ideas to industrial organizations.

Thus, the growing importance of digital skills of the population, firstly, is determined by ensuring effective adaptation to changes in the labor market, increasing labor productivity, developing innovations and introducing new technologies. Secondly, it requires the acquisition of new professional skills and abilities that become an integral element of modern work, which is determined by the rapid involvement of ICT and digitalization processes. Thirdly, it determines the need to continue scientific research in the direction of clarifying new horizons of the use of digital skills, which will allow to outline the directions of involving ICT and digital technologies and devices in professional and everyday activities.

3 UNDERSTANDING DIGITAL SKILLS AND THEIR MAIN COMPONENTS AND LEVELS

The rapid development of ICT and their integration into the economic system determines not only the emergence, but also the constant development of digital skills, which affects the increase in the number of scientific attempts to substantiate their theoretical and methodological foundations. In the scientific community, digital competences, or as they were

originally called, digital literacy, began to be studied at the end of the 20th century. In 1997, Paul Gilster introduced and argued for the first time the concept of "digital literacy", under which he understood the ability to disclose and use information in various formats obtained from a wide range of sources and presented with the help of a computer [5, p. 1]. In the future, the rapid pace of digitization significantly changed not only the components of digital literacy, but also transformed them into digital competencies, which today are actively analyzed at the theoretical level (in particular, there are other names of new competencies, for example, e-skills or 21st century skills) and develop based on dissemination numerous educational programs, courses, trainings, etc.

In a more general sense, digital skills are skills and knowledge that allow the use of technological and digital tools in personal and professional spheres, which involves the ability to search, create, evaluate, and use digital information [6, 7]. In accordance with our beliefs, digital skills should be understood as skills and abilities that are formed throughout a person's life and allow to achieve effective professional and creative self-realization based on the use of digital tools and services, which are necessary for the creation and distribution of digital content, search and protection of information, communication and interaction, assessment of potential risks and formation of a motivational development strategy.

It is worth emphasizing that digital competences and digital skills are not identical concepts, but their interrelation is undeniable, since together they reflect a person's ability to use modern digital technologies in their activities. However, digital skills are a basic and narrower concept, as they reflect specific practical abilities that allow performing tasks involving new devices or technologies. In turn, digital competences should be identified not only with digital skills, but also include a set of knowledge, attitudes and responsibilities that allow for the effective, safe and ethical use of digital technologies in various areas of human life. That is, digital skills are the basis of digital competences or the basis on which broader and more complex components are formed.

The main feature of digital competences is their dynamic nature, because they constantly need to be developed and updated as a reaction to the emergence of new technologies and further digitalization. For example, if previously the main skills of digital skills were associated with the ability to perform basic hardware and software operations, use e-mail and find the necessary information, then in the last five years their transformation and skills that allow success today involve the use artificial intelligence, Big Data, blockchain, cloud computing, Internet of Things, etc. [8]. Therefore, one of the important factors in the development of business activity and the economy as a whole is the introduction of new technologies and digitalization, which, according to researchers, should lead to a change in approximately 44% of the skills of workers over the next five years [9], among which special priority is given creative thinking; analytical skills; technological literacy, which characterizes digital skills; continuous education and the desire for self-education; stress resistance, flexibility.

Table 1: Different approaches to defining the components of digital skills.

Author or Organization	Components of Digital Skills
UK Ministry of Education [10]	Digital foundation skills, communicating, handling information and content, transacting, problem solving, being safe and legal online.
European Commission for Science and Knowledge [11]	Information and data literacy, communication and collaboration, digital content creation, safety, problem solving.
Ester van Laar, Alexander J.A.M. van Deursen and others [12]	Information, communication, collaboration, critical-thinking, creativity, problem-solving.
Bertrand Audrin, Catherine Audrin, Xavier Salamin [9]	Technology use, cybersecurity, content management, communication and collaboration, critical inquiry, responsibility, well-being, identity and development.
Krylov Denys [13]	Information skills, ability to the communication process, security, ability to solve problem situations, ability to create modern content environments.

Table 1 presents different approaches to the selection of structural elements of digital skills, which demonstrates the lack of unanimity on this issue. First, this is due to the recent allocation of such new skills in the structure of human capital, as well as the lack of a unified understanding of their essence and

importance for the development of the economy and the opportunities that ICT discloses for economic activity. The digital skills correspond to the modern development of the economy, which is undergoing changes due to the active implementation of ICT, should form the following elements:

1) technical and computer skills, which are the basis for the formation of digital skills, since they are aimed at creating skills and abilities that allow the effective use of technical and digital and software tools, including computer networks, Cloud services; create modern content environments;
2) the ability to the communication process, which involves the ability to choose the most successful method of communication and the ability to use modern Internet tools necessary for communication in the form of individual or collective communication for the purpose of data exchange in the digital environment (in this aspect, different methods of communication can be distinguished depending on the purpose, for example, video conferences, e-mail, forums, social networks, chats, blogs, etc.);
3) security skills that form an understanding of the need to protect information in a modern digital society and are also designed to create a few measures to ensure the protection of personal data and individual information, guarantee their confidentiality, and manage risks. In this case, the key aspects are the creation of reliable passwords, the use of legal content, compliance with the rules of the culture of behavior on the Internet, the creation of backup copies and the use of modern cloud storage methods;
4) media information skills and data analysis, which form a set of knowledge, abilities and skills that allow, firstly, to find, including using digital tools, the necessary relevant information, to carry out its systematization and critical analysis, with further interpretation, storage and processing; secondly, to interpret static information and other universal indicators;
5) the ability to solve problems based on digital technologies, which is based on making decisions about the involvement of appropriate digital tools and a creative approach to the use of existing technologies and solving technical problems, as well as the identification of digital needs and resources.

Along with highlighting the components of digital skills, it is advisable to highlight their levels, which is presented in Figure 1. This logic is due to the need to acquire various digital skills depending on the performance of the assigned tasks. The basic level of digital skills facilitates everyday life, establishes interaction with others, allows you to receive commercial and financial services, as well as e-government services. The intermediate level also characterizes general digital competences, but, in comparison with the basic level, enables the performance of a significant number of digital tasks in professional activities. Only a small percentage of people in the world have a high level of digital skills, which characterizes the skills of artificial intelligence, Big Data, coding, cybersecurity, Internet of Things [19]. The holistic analysis of digital skills should include an assessment of each of its components at three main levels.

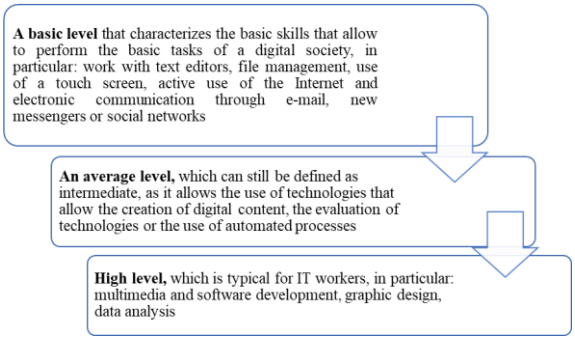


Figure 1: The main levels of Digital Skills.

Therefore, the development of digital skills in the structure of human capital is the only way to preserve the competitiveness of not only a person as a specialist in a certain field, but also the national economy in a changing digital environment.

That is, society is gradually moving to a qualitatively new level of development - digital, which can be considered as the next step of the information technology revolution or the fourth industrial revolution (Industry 4.0). Such changes involve increasing the level and requirements for digital skills and human knowledge, which is the basis for the involvement of broad sections of the population in ICT and digital transformation, as well as the development of social integration. The socio-economic development of the 21st century determines the priority of the economy intangible factors of production, which should be included human and intellectual capital, innovative activity, intellectual properties operating based on knowledge sharing and evaluation. So, knowledge and digital

skills, which today are embodied in human capital, act as the main factor of innovative development. That is, dominant the position of T. Sakaya becomes that we are entering a new stage of civilization, at which the driving force is the values created by knowledge [14].

4 ANALYSIS OF THE DEVELOPMENT OF DIGITAL SKILLS IN UKRAINE AND THE WORLD

We are convinced that the level of development of digital skills depends on the pace of development of the national economy, its competitiveness, the activity of implementing and spreading innovative activities.

The problem of developing basic digital skills today faces all countries of the world, including economic scientists in the EU and the USA. The very innovative directness and focus on the development of knowledge and digital memory ensured the economic advancement of the “Asian tigers” - Japan, South Korea, Taiwan, Hong Kong, Singapore. And the rest of the “Intellectual Island” program in a short period of time achieved great results, which allowed them to take leading positions in many world ratings, and lead to transformation Economical innovations from commodity production to intellectual and creative ones, which to set the pace of world development. The WILEY publishing house in 2021 evaluated DSGI (The Digital Skills Gap Index), which integrally combines three components, namely: Digital Skills Institutions, Digital Responsiveness, Government Support, each of which consists of separate sub-indices [15]. Singapore has the highest level of DSGI – 7.8; the second position is occupied by the USA – 7.5; third place – Finland, which is the best among European countries – 7.5 (Fig. 2).

We also analyzed The Global Competitiveness Index, which reflects the country's ability to ensure a high level of prosperity and depends on how effectively the country uses existing and potential resources. GCI contains 12 sub-indices, which are ranked into four main blocks (Enabling Environment, Human Capital, Markets), each reflects separate elements of competitiveness. According to the GCI, in 2019, the top three were Singapore and the United States [14]. That is, our assumption is confirmed that countries with high growth rates are characterized by the presence of a higher level of digital skills. In the analysis of European countries (Fig. 2), we observe certain deviations, but in general the trend remains: countries with a lower DSGI score have a lower GCI rating. That is, due to the low rates of economic development of Eastern European countries (Turkey, Romania, Ukraine, Moldova) have lower indicators. We believe that there will be a downward trend in the coming years level of DSGI in Ukraine, because the consequences of full-scale will become evident military invasion of Russia. As a result, not only her suffered socio-economic infrastructure, but social ones will also be felt damages, which, first, are determined by the loss of the most valuable resource - man and his capital. In this aspect, the following circumstances can be defined, these are:

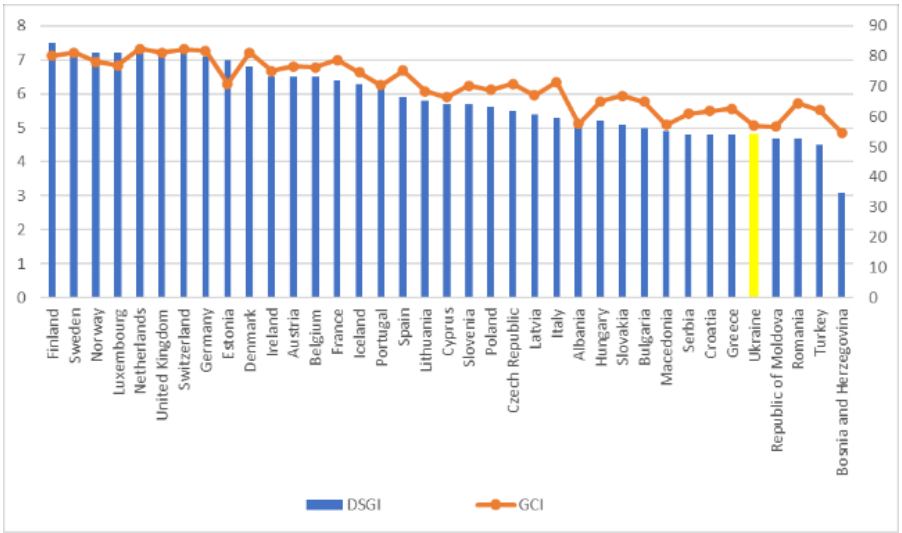


Figure 2: Indicators of DSGI and GCI indices for European countries.

1) killed on the battlefield;
2) forced migrants who went abroad to in search of safe living conditions;
3) lost opportunities for development, because education and self-study during the war are not priority types activity.

But we believe that the Ukrainian government needs to pay attention to the development and distribution among the broad segments of the digital population skills, because they make it possible to develop the economy on the latest foundations, which will be able to provide a breakthrough to the post-industrial society.

Therefore, the world community understands the benefits that the economy and society receive from the development of digital skills. So, EDSC fully supports the goals of the European Program for the Development of Digital Skills, which provide for the achievement of the following level:

- by 2025, 70% of the population aged 16 to 74 should acquire a basic level of digital skills;
- by 2030, 80% of the population aged 16 to 74 should acquire a basic level of digital skills, which is the goal of the European Digital Decade [17].

Thus, digital skills not only change the usual way of working, but also create the basis for sustainable economic growth. After all, the gradual integration of modern complex technologies and Artificial Intelligence into production processes is becoming evident. Accordingly, the establishment of effective coordination involves the use of digital tools, the use of which requires at least an average level of digital skills. Accordingly, the use of digital skills becomes a prerequisite for achieving success, which determines the outline of factors that encourage the development of new competencies, namely:

1) the opening of new opportunities, because having at least an average level of digital skills is a guarantee of success in a changing external environment that is evolving quite quickly from the mass use of personal computers to Artificial Intelligence;
2) the ability for professional growth and constant self-improvement, since the competitiveness of a specialist is determined by his skills in using ICT and other digital technologies, which develop at a fairly fast pace and therefore require their constant study;
3) creativity and the ability to innovate, allowing to implement new ideas, to encourage innovations;
4) increase in productivity and efficiency, which is achieved on the basis of the use of new equipment and technologies, which, in addition, allow to simplify production processes and save time for performing other tasks;
5) achievements of remote employment;
6) use of educational and informational online resources;
7) development of new types of entrepreneurships, in particular Internet trade;
8) the ability to learn new things and adapt to new requirements of the time.

In general, the emergence of the digital economy demonstrated not only a shortage of highly skilled workers with new skills, but also an insufficient level of digital skills among the general population. In particular, the Ministry of Digital Transformation of Ukraine periodically analyzes the acquisition of digital skills among the adult population based on four levels: 1) no skills, which characterizes the lack of digital skills in such four areas as information, communication, solving life problems and creating digital content; 2) low skills, which imply a lack of digital skills in one of the four areas of competence; 3) basic skills – the level of mastery of digital skills in all four areas at a level no lower than "average"; 4) above basic skills – the level of mastery of digital skills in all four areas at a level no lower than "above average". The disadvantage of such a study is that the proposed methodology differs somewhat from global standards, which makes it impossible to compare Ukrainian and Western assessments. However, the conducted research demonstrated positive changes, which predict a reduction in the share of the population that does not have any skills in using digital technologies, which in 2023 was only 7.2% as opposed to 15.1% in 2019 [18]. In terms of age, of course, the largest percentage of the population that does not have digital skills belongs to the age group of 60-70 years and is 17.4%. This is understandable, because the older generation is quite skeptical about innovations and their active use. It is necessary to note the positive trend in the development of above basic skills, as the share of the population with this level of digital skills has increased by almost 50% over the past four years (Fig. 3).

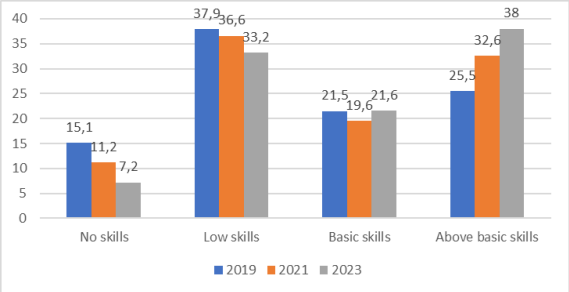


Figure 3: Dynamics of Digital Skills development among the adult population of Ukraine.

With the help of a trend line based on linear dependence, we forecast the further development of digital skills in Ukraine (see Fig. 4). Using the linear dependence corresponding to the equation y=3.125x-6283.6, in 2024 we should expect an increase in above basic skills to 41.4%, in 2025 – 44.5%, which can be a positive gain for the Ukrainian economy, which must go through the stage of recovery due to the destruction caused by the development of a full-scale war. The gradual growth of digital skills is significant, because such reconstruction is better to be carried out based on digital transformation.

5 CONCLUSIONS

Among the main factors affecting the level of development of Digital Skills in Ukraine are the level of education, type of employment and financial situation. We believe that in order to eliminate the problems of Digital Skills development in Ukraine, it is advisable to form a National Strategy for the Development of Digital Skills, which should be developed for the short-term perspective, because new skills are not static, so they need to be constantly improved, as a result of the emergence of new technologies and their impact on Digital Economy and Digital Society. In addition, it is desirable to use European methods and approaches to conduct more comprehensive and thorough analyzes of the assessment of the Ukrainian level of Digital Skills, which will allow a comprehensive comparison of the achieved level of new digital skills of the population of different countries.

Therefore, the development of digital skills is the basis for accelerating the process of digitalization of society and the national economy, gradually they are becoming a mandatory attribute of modernity. Accordingly, the spread of digital skills among broad segments of the population is the key to the active implementation of ICT and a consistent transition to a higher level of development. We believe that the prospects for further scientific research in this aspect should include the development of methods or a scale for assessing digital skills. However, it is worth emphasizing that such a calculation must necessarily focus on the field of work of the owner of digital skills.

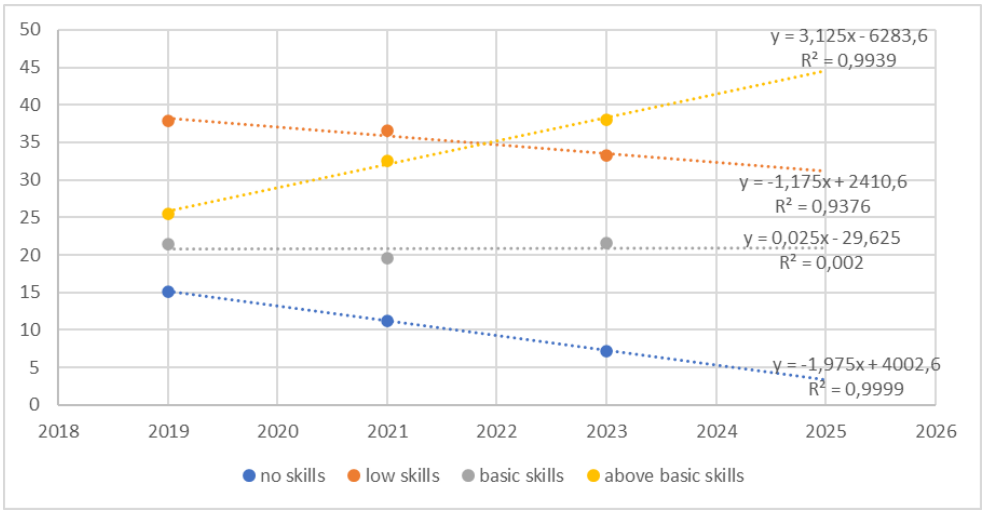


Figure 4: Forecasting the development of digital competences in Ukraine.

REFERENCES

[1] L. R. Mahmutaj and N. Jusufi, "The Importance of Digital Skills in Firms' Innovation: The Case of Western Balkans," Journal of Technology Management & Innovation, vol. 18, no. 3, pp. 90–102, 2023. doi: 10.4067/S0718-27242023000300098.

[2] S. Lythreatis, S. K. Singh, and A.-N. El-Kassar, "The digital divide: A review and future research agenda," Technological Forecasting and Social Change, vol. 175, 2022. doi: 10.1016/j.techfore.2021.121359.

[3] P. M. L. Ng, P. K. Chow, P. Wong, and W. M. B. Luk, "The impact of digital technology training on developing academics' digital competence in higher education context," Education + Training, vol. 66, no. 9, pp. 1276–1292, 2024. doi: 10.1108/ET-10-2023-0417.

[4] S. Saniuk, D. Caganova, and A. Saniuk, "Knowledge and Skills of Industrial Employees and Managerial Staff for the Industry 4.0 Implementation," Mobile Networks and Applications, vol. 28, pp. 220–230, 2023. doi: 10.1007/s11036-021-01788-4.

[5] P. Gilster, Digital Literacy. New York: John Wiley, 1997, 292 p.

[6] "Digital Skills," [Online]. Available: https://www.iseazy.com/glossary/digital-skills/#:~:text=Digital%20skills%20are%20skills%20and,productivity%2C%20creativity%20and%20professional%20resilience.

[7] "What Are Digital Skills?" [Online]. Available: https://digitalskills.unlv.edu/digital-marketing/what-are-digital-skills/.

[8] "Digital Skills Toolkit," [Online]. Available: https://www.itu.int/en/ITU-D/Digital-Inclusion/Documents/ITU%20Digital%20Skills%20Toolkit.pdf.

[9] B. Audrin, C. Audrin, and X. Salamin, "Digital skills at work – Conceptual development and empirical validation of a measurement scale," Technological Forecasting and Social Change, vol. 202, May 2024. doi: 10.1016/j.techfore.2024.123279.

[10] Department for Education, Essential Digital Skills Framework. GOV.UK, 2019. [Online]. Available: https://www.gov.uk/government/publications/essential-digital-skills-framework/essential-digital-skills-framework.

[11] A. Ferrari and Y. Punie, "DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe," Publications Office of the European Union Luxembourg. [Online]. Available: http://digcomp.org.pl/wp-content/uploads/2016/07/DIGCOMP-1.0-2013.pdf.

[12] E. van Laar, A. J. A. M. van Deursen, J. A. G. M. van Dijk, and J. de Haan, "21st-century digital skills instrument aimed at working professionals: Conceptual development and empirical validation," Telematics and Informatics, vol. 35, no. 8, pp. 2184–2200, 2018.

[13] D. V. Krylov, "Transformation of the market in the minds of digitalization of the national economy of Ukraine," Economic Synergy, vol. 2, no. 12, pp. 21–32, 2024. doi: 10.53920/ES-2024-2-2.

[14] T. Sakava and G. Fields, The Knowledge-Value Revolution or a History of the Future. Kodansha USA Inc., First Edition, 1991, 379 p.

[15] "The Digital Skills Gap Index," [Online]. Available: https://dsgi.wiley.com/global-rankings/.

[16] The Global Competitiveness Report, 2019. [Online]. Available: https://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf.

[17] "Digital Education Action Plan – Action 9," [Online]. Available: https://education.ec.europa.eu/focus-topics/digital-education/action-plan/action-9.

[18] "Research on digital literacy in Ukraine," Ministry of Digital Transformation in Ukraine. [Online]. Available: https://osvita.diia.gov.ua/uploads/1/8800-ua_cifrova_gramotnist_naselenna_ukraini_2023.pdf.

[19] Digital Skills Insights, 2020. [Online]. Available: https://www.itu.int/dms_pub/itu-d/opb/phcb/D-PHCB-CAP_BLD.03-2020-PDF-E.pdf.

Methodical Approach to the Startup Ecosystem Formation: Foreign Experience and Opportunities for Ukraine

Viktoriia Khaustova, Mykola Kyzym, Olena Reshetnyak, Mykyta Khaustov, and Nataliia Trushkina
Research Center for Industrial Problems of Development of the National Academy of Sciences of Ukraine,
Inzhenernyi Lane 1a, 61165 Kharkiv, Ukraine
{v.khaust, m.kyzym, khaustov.mkt}@gmail.com, {reshetele, nata_tru}@ukr.net

Keywords: Startup, Project, Company, Organization, Life Cycle, Management, Scaling, Strategic Development, Innovation, Innovative Activity, Ecosystem, Institutional Support, Methodical Approach, Model, Entrepreneurship, Partnership, University.

Abstract: In modern conditions, startups play a key role in ensuring the sustainable development of national economies around the world. In view of this, the purpose of this study is to substantiate a methodological approach to creating a startup ecosystem in Ukraine based on the generalization of best global practices. To achieve the stated goal of the study, the following scientific approaches and methods were used: analysis and synthesis; system approach; bibliometric analysis; content analysis; economic and statistical analysis; rating evaluation; expert assessments; comparative analysis; system-logical analysis; cluster analysis; monographic method; graphic method. The article clarifies the essence of the concept of a "startup ecosystem" as a network association of organizational structures of various forms of ownership that provide support for the creation of innovative startup projects at different stages of their life cycle with subsequent scaling. As a result of the study, it was proven that the development of startups in Ukraine involves the creation of an effective ecosystem that would provide their support at all stages of the life cycle (from the creation of an innovative project to its implementation in the form of a company). To achieve this, the composition of the elements of the startup ecosystem was formed according to the main stages of their life cycle and a methodical approach to the formation of a startup ecosystem in Ukraine was developed, which is characterized by a specific set of institutions in the country, the complex combination of whose efforts will provide support for the formation of startups (at all stages of their life cycle) and will contribute to the activation of innovative entrepreneurship. The practical value of scientific developments lies in the fact that the proposed methodological approach has been implemented in the activities of the Association of Industrialists and Entrepreneurs of the Kharkiv Region.

1 INTRODUCTION

A modern trend in innovative and entrepreneurial activities of the countries of the world is the creation and scaling of startups, which become the driving force of their economic development based on the creation of new jobs, attracting investments, stimulating innovations, expanding market segments and improving the quality of life. Startups are the so-called "engine" of technologies and innovations, an "amplifier" for modernization, digitalization [1-3] and sustainable development [4] of the countries of the world.

It can be noted that the startup movement is developing due to modernization and creativity, which have great potential to change the global economy. Today, there are 150 million startups in the world, of which 50 million new startups are launched every year. On average, 137 thousand startups appear every day [5].

In 2021, the volume of financing for various startups in the world amounted to more than 600 billion USD. At the same time, the income of some international startups as innovative ideas and companies has increased many times: Apple's assets exceed 350 billion dollars, Google – 130 billion, Facebook – 51 billion dollars. In addition, the development of startups allows solving employment problems in countries around the world. In 2022, in countries that are members of the Organization for Economic Cooperation and Development, startups

provided about 50% of all new jobs. Governments of countries around the world and various organizations and institutions support the development of startups, realizing their importance and potential benefits for the economy.

Therefore, at present, a startup should be considered as a "kind of catalyst" for economic growth at the local, national, and global levels. They can help find financial resources, technologies and specialists that will bring the country's economy to a fundamentally new level of development. For the formation and functioning of a full-fledged and effective ecosystem of startups, it is necessary to reform the relevant institutions, implement the concept in the field of IP culture, and officially make the development of startups an important component of the economic strategy.

The above testifies to the relevance of the selected research topic and the need for further development within the outlined issues.

2 LITERATURE REVIEW

Startups are the object of research by many foreign and domestic scientists, who focus on various aspects of this problem: essence and characteristic features, varieties and their features, life cycle models, success factors and advantages, sources of funding, support ecosystem, etc.

Among the works devoted to various aspects of the creation and development of startups, it is appropriate to highlight the research of leading foreign (S. Blank [6], M. Bliemel et al. [7], S. Breschi et al. [8], J. Gans [9], M. Henry [10], A. Skala [11]) and Ukrainian (N. Bielikova [12], O. Dymchenko et al. [13], L. Frolova et al. [14], O. Havrysh et al. [15], I. Hubarieva [16], N. Kulyk [17], N. Podolchak et al. [18], I. Savin [19], O. Trofymenko and O. Ilyash [20], M. Tymoshenko [21]) scientists.

Based on the significant scientific achievements of scientists in this field, it should be noted that the dynamism of the development of the startup industry, the specificity and scale of challenges, both global and national, as well as the increasingly significant role of startups in ensuring innovative entrepreneurship and economic growth of the countries of the world require deepening of research related to the analysis of the specifics of the development of startups at various stages (creation, development) and stages of the life cycle, approaches to identifying its stages and determining effective managerial influences, improving the general management cycle for the creation and scaling of startups, determining directions for improving the ecosystem of startups taking into account the specific features of certain countries of the world.

Separately, it is worth noting the research of scientists, which is devoted to the theoretical and applied aspects of the formation of startup ecosystem models. It should be emphasized that the concept of the startup ecosystem has recently been widely used in the context of innovation and entrepreneurship [22-25]. The understanding of the term "startup ecosystem" is quite broad, and although there is no single, universally accepted definition, the term is used in relation to a certain geographic region (for example, Silicon Valley) with a high density of startups and entrepreneurs [23]. It should be noted that the scope of the ecosystem can vary from a collection of several startups to a region or country, but the use of this term in relation to individual cities is the most common. For example, The Global Startup Ecosystem Ranking report defines a startup ecosystem as "a city or geographic area (with an approximate radius of 100 km) that uses shared resources" [26].

As in natural ecosystems, a key characteristic of a startup ecosystem is the interdependence (or "common fate") of the various organisms within the ecosystem. In other words, the startup ecosystem represents the relationship between its individual actors or groups of actors, which distinguishes ecosystems from other concepts, such as clusters, organizations (for example, meetings, and hackathons) [27]. Recently, many authors, including C. Mason and R. Brown, emphasize the significant role of entrepreneurs in the startup ecosystem, defining it as a set of interconnected entrepreneurial entities (both potential and existing), entrepreneurial organizations (for example, firms, venture capitalists, business angels, banks), institutions (e.g. universities, public sector institutions, financial institutions) and entrepreneurial processes (e.g. the rate of opening new businesses in the region, the number of growing firms, the level of venture business, the number of businesses, that are scalable, the level of entrepreneurial ambitions, etc.), which formally and informally unite with the aim of developing the local entrepreneurial environment [22].

Startup Commons defines a startup ecosystem as a network consisting of individuals, startups at different stages of development, and different types of organizations cooperating as a system to create new startup companies [28].

Depending on the territorial specification, scientists define and study the ecosystem of startups at the level of a specific city, country or at the global level [29]. The startup ecosystem encompasses a variety of organizations that interact to bring significant growth to the market through growth of startup companies, returns to investors, and benefits to end customers. Interrelationships between components create a complex network of relationships between startups, organizations, and the external environment.

Taking into account the above and previous research results, it is proposed to consider the startup ecosystem as a network association of organizational structures of various forms of ownership that provide support for the creation of innovative startup projects at various stages of their life cycle with subsequent scaling [25].

In the scientific literature, there are various models of startup ecosystems. One of them is proposed by H. Caleb [30], who puts entrepreneurs in the spotlight because their leadership is considered key to the success of the startup economy. It also includes such elements as the state, universities, mentors, service companies, corporations, investors and public initiatives that support entrepreneurship in the startup ecosystem model. Each of these elements is considered as a comprehensive tool, the support of which forms a successful ecosystem of startups.

B. Spigel [31] considers organizational and social factors as the key internal components of the model of startup ecosystems. These factors include general entrepreneurial culture, success stories, human talent, investment capital, social networks and mentorship. The proposed model also includes material factors such as: government institutions and universities, service companies, as well as infrastructure and local markets with geographic significance. These internal factors act as components that influence processes in the startup ecosystem.

Scientists are also considering approaches to defining the ecosystem of startups based on the triple and quadruple helix model (The Triple and Quadruple Helix Model) [32-37]. According to these approaches, the formation of innovations in startups is a result achieved not by one institutional player, but the result of interaction between all participants of the ecosystem as a source of new organizational plans [32, 33]. Thus, at its core, the triple helix model describes the institutional framework for innovation development (including startup-based), which is created by a triadic network between the academic community, industry, and government [32-35]. The main relationships between the three institutional factors in the triple helix model of startup ecosystems are as follows [34]: (1) technology transfer, (2) cooperation and conflict resolution, (3) shared leadership, (4) replacement of certain functions, and (5) chain. In particular, technology transfer is a key element of the triple helix, especially in industries with high innovativeness of startups [36, 37].

In recent years, researchers have proposed some variations of the triple helix model to integrate additional variables that affect collaboration between NGOs, academia, and business. In particular, a four-link spiral model was proposed to describe the startup ecosystem, which, in addition to the three key institutional bodies, also takes into account the role of culture, civil society, and mass media [38]. The model of four spirals determines the importance of the environmental component in the development of startup ecosystems [39]. Among the interactions that take place in the wider social context, the processes of production of innovations and new knowledge in modern economies are the main ones. In addition, special attention is paid to institutional planning and changes, i.e., systemic processes of updating and reforming the institutional management base of innovative entrepreneurship [38].

Startup ecosystems include entities participating in the four-link spiral model, whose actions determine the business path and success of startups, especially for those characterized by a high degree of technological innovation and provision of services at the global level, needing relationships with by other subjects of the ecosystem [40]. The Aspen Institute [41] analyses the existing framework of the entrepreneurial ecosystem and identifies the following main determinants of the ecosystem that affect entrepreneurial success: education and scientific research, human resources, government intervention, networking and support, financing.

C. Ziakis et al. [42] propose to include the following factors influencing the sustainability of startups in the startup ecosystem model: education

and scientific research, human capital, funding, government, business support and communication, entrepreneurial culture, and incentives for startups. The startup ecosystem can be considered through the analysis of its structural components (business angels, accelerators and incubators, mentors, state support, venture capital, investors, universities, research and development) and their role in the development of startups [43]. In general, this ecosystem is an open system aimed at creating and constantly improving startups. This system integrates startups with the entities that ensure their existence, and all of them function in a certain environment, being interconnected by a network of dynamic relationships.

In view of this, the purpose of this study is to further develop theoretical and methodological support and develop a methodical approach to creating an ecosystem of startups in Ukraine based on the generalization of the best global practices.

3 METHODOLOGY

The basis for the study was research data from international organizations, rating agencies, information portals; statistical reporting of the World Bank; data from international scientometric databases Scopus and Web of Science; Laws of Ukraine; by-laws and regulatory legal acts of the Cabinet of Ministers of Ukraine; materials of the State Statistics Service of Ukraine; scientific publications on the selected topic and related ones; developments from the project activities of the Research Center for Industrial Problems of Development of the National Academy of Sciences of Ukraine, information materials and reports of domestic companies, Internet resources and results of the authors' own research.

To achieve the set goal of the study, the following scientific approaches and methods were used: grouping, deduction and abstraction; analysis and synthesis; systematic approach; morphological analysis; bibliometric analysis; content analysis; economic and statistical analysis; rating evaluation; integral evaluation; expert assessments; comparative analysis; system-logical analysis; cluster analysis; monographic method; graphic method.

In order to study the world experience in assessing the development of startups and their support infrastructure, the following indices were considered: Global Startup Ecosystem Index; Most startup friendly countries in the world; Global Startup Ecosystem Report (GSER); Startup Index of Nations, Cities; Startup Ranking.

As statistical analysis shows, the leader in the number of startups at the end of 2022 is the USA. According to experts, there are currently 75,056 startups in the country due to a favourable business climate for their development and a developed ecosystem. The USA has 5.7 times more startup teams than India, which ranks second with 13,125 active startups [44].

According to experts [45], in 2017, more than 6 thousand startups were operating in India, and in 2022 – 80 thousand. It is worth paying tribute to the state policy of India, the implementation of which contributed to ensuring low taxes for startup developers and creating institutional conditions for full development. And therefore, at least 5% of the country's GDP is provided by startups.

According to the Ministry of Science and Technology of India [46], in 2014-2023 the number of startups in the country increased by 257 times or from 350 to 90 thousand. At the same time, 100 startups appeared in the field of space technologies. Currently, the number of biotechnology startups has increased from 50 to 6 thousand startups. According to the Startup India website, there were 96,753 startups registered with DPIIT (The Department for Promotion of Industry and Internal Trade is a central government department under the Ministry of Commerce and Industry in India), as well as 196 government-recognized accelerators and 1,093 incubators.

In 2022, the Government of India decided to invest INR 1,711 crore in startups under the Startup Fund of Funds (FFS) scheme. In addition, an additional INR 267.50 crore has been allocated for incubators under the Startup Seed Fund (SISFS) scheme. A number of other schemes have been introduced to boost the local startup ecosystem, such as the Credit Guarantee Scheme for Startups (CGSS), a three-month income tax exemption.

An example of best practice is the experience of Turkey, where startup development was systematically ignored until 2010. However, when the government of the country realized the role of this direction and its development prospects, startups raised a total of about 100 million dollars. Getir was launched as a startup in 2015, after which it managed to raise over 400 million dollars, and later increased its value to 2.6 billion dollars. Currently,

Getir, a startup in the field of restaurant delivery, has over 32,000 employees and operates in the markets of Turkey, the USA, the Netherlands, Germany and the UK.

The number of startups in the top 10 countries is shown in Table 1.

Table 1: Number of startups in leading countries as of the end of 2022[1].

Country	Number of startups, units
United States	75,056
India	13,125
United Kingdom	6,220
Canada	3,303
Indonesia	2,347
Germany	2,295
Austria	2,262
France	1,567
Spain	1,407
Brazil	1,116

According to Startup Ranking [48], the leading countries in terms of the number of startups founded in 2023 were the United States (3,687 startups), followed by India (990) and the United Kingdom (555). The analysis shows that the majority of startups are in the areas of financial technology, healthcare, and artificial intelligence (Table 2).

Table 2: Distribution of startups in the world based on their industry of implementation and the corresponding percentage of the total number, 2022[2].

Implementation sector	Share, %
FinTech	7.6
Life Sciences and Healthcare	6.8
Artificial Intelligence	5.0
Gaming	4.7
Advertising Technology	3.3
Educational Technology	2.8
Green Technology or Eco-Friendly Practices and Technologies	2.1
Blockchain	1.5
Robotics	1.3
Cybersecurity	0.7
Agricultural technologies	0.6

Let us consider the leading countries of the world in the context of the creation of startups and their support ecosystems, as well as their socio-economic development. The level of creation of startups and their support ecosystems in the countries of the world is determined by such indicators as: The Index of the Most Startup-Friendly Countries in the World, the Global Startup Ecosystem Index and the number of startups per 1,000 population of the country [48; 50-54]. These indicators were used to conduct a cluster analysis of 55 countries of the world using the Statistica 8.0 package.

According to the analysis [25], three clusters of countries of the world were identified by the level of creation of startups and their support ecosystems (Table 3).

Table 3: Clustering of countries around the world by the level of startup creation and their support ecosystems[3].

Cluster	Cluster characteristics	Countries
First	Average level of values according to the indicators of the Index of the most startup-friendly countries in the world, the Global Startup Ecosystem Index and the number of startups per 1 thousand population	18 countries, including Canada, Luxembourg, Israel, Belgium, Sweden, Denmark, the Netherlands, Australia, etc.
Second	High level of values according to the indicators studied	3 countries: US, Singapore, Estonia
Third	The lowest values of indicators among the studied countries according to the Index of the most startup-friendly countries in the world, the Global Startup Ecosystem Index and the number of startups per 1 thousand population	33 countries, including India, Germany, Poland, South Korea, Japan, the Czech Republic, Turkey, Ukraine, etc.

Let us analyse in more detail the startup ecosystems of some countries of the world, in particular, the leading countries and some representatives of the cluster, which are potential leaders in the development of startups.

Thus, the first country in the group of leaders is the United States, which occupies a leading position in the development of startup ecosystems, having 7 cities (regions) in the ranking of the 20 largest ecosystems for the period 2019–2023. The United States has formed an advanced environment for the development of startups, which presents opportunities for breakthrough technological creativity [55]. The national startup ecosystem of the United States is used by foreign entrepreneurs, determining that the country has the best conditions for scaling and developing a global company - the American version of free market capitalism is

[1] compiled based on [47]

[2] compiled based on [49]

[3] compiled based on [25; 48; 50-54]

focused on profits and high risk, which also provides startups with opportunities, financing and ecosystem support. All this has allowed creating a startup ecosystem, which is mainly based on the private sector without excessive intervention from the public sector. This market-driven approach to startup development has led to the creation of some of the most successful startups in the world.

Singapore is a recognized leader among Asian countries for starting a business, including venture capital. Singapore is a model for innovation, a country of choice for registering startups operating in Asia, due to its financial stability, approach to business and tax policies. Because Singapore has a small market and population, the growth of its ecosystem depends on scaling abroad [56]. Singapore's economy is attractive to large business investors, including a significant number of venture capitalists who are willing to fund international startups. Singapore's ecosystem is characterized by a growing number of accelerators and support networks for startups. For example, Enterprise Singapore has programs to stimulate the development of startups, allowing them to access financial assistance and business loans in their early stages of existence. Initiatives such as Startup SG are aimed at promoting Singapore's startup ecosystem. The country also has a strong digital infrastructure and many sources of investment. In addition, the country's universities are involved in the development of startups not only by preparing a highly skilled workforce for the research and development sector, but also by connecting programs to startups and encouraging entrepreneurship on campus. The Singaporean public sector implements many tools to support the startup ecosystem: it provides significant support to young businesses, provides a favourable innovation policy (dividends and capital gains are not subject to tax), and there are significant tax preferences for registered resident companies [57].

Estonia, a very small country with a population of 1.3 million, has the highest IT development potential in Europe – about 10% of the country's working-age population is employed in the IT sector, which accounts for about 7% of the country's GDP and 14% of its exports. Tallinn, the capital of Estonia, has been compared to the Silicon Valley of Europe, as it has the highest number of startups per capita [58]. One of the most important milestones in the history of the Estonian ecosystem was the success of Skype, an application developed mainly in Estonia. The founders of Skype used this windfall to support new successful Estonian startups, such as Skycam, teleport, and SpaceApe [59]. According to Statistics Estonia, one in 56 employed Estonians was involved in startups, and the top 20 startups created 59% of jobs in the sector in 2022, demonstrating the country's need to attract talent from abroad [59]. The Work in Estonia program launched by the Estonian government, aimed at attracting new IT professionals as well as talent in the natural and exact sciences, has doubled the international talent pool in the past five years. According to the White Paper Startup Estonia 2021–2027 [58], the country continues to support its startup ecosystem and has set a number of goals aimed at growing Estonia's startup and technology sector to 15% of the country's GDP by 2025.

As the main government initiative for the Estonian ecosystem, Startup Estonia works on policy development, as well as promoting and strengthening the ecosystem. In addition, thanks to events such as Latitude59, the country continues to attract international attention and investment in its ecosystem every year. Estonia's success in developing startups and creating one of the most effective systems of support for entrepreneurship development is based on a complex combination of the following key factors: an active and progressive government that promotes the development of innovative businesses, simplification of business processes; an effective and dynamically developing startup ecosystem (venture funds, incubators, accelerators, universities and other development institutions), ensuring full access to the European market and markets in other regions of the world [58].

If we consider Ukraine, it should be noted that in 2019, 510 million dollars was invested in Ukrainian startups and IT companies, of which the share of foreign capital was 80% of the received investments. According to Startup Ranking [48], the country ranked 37th among 137 countries in the world in 2023. In other international rankings dedicated to the development of the startup movement and its support ecosystem, Ukraine ranks as follows: in terms of the number of startups per 1,000 population of the country – 45th place in 2022, in terms of the index of the most startup-friendly countries in the world – 43rd place, in terms of the global startup ecosystem index – 49th place in 2023.

According to The global startup ecosystem Report 2023, Ukraine improved its results compared

to the previous 2022, rising one position to 49th place in the world. In Europe, Ukraine maintained its position and is in 30th place, ahead of Hungary and Serbia with an overall score gap of less than 5% [26]. As noted in the report [45], Ukraine is on the path of recovery, and most of its ecosystems have risen in the 2023 ranking compared to 2022.

The report emphasizes that due to the ongoing war in Ukraine, it is difficult to predict the consequences for startup development and the speed of recovery of Ukraine's physical and economic infrastructure. In addition, it is noted that, despite the fact that Ukraine experienced economic difficulties for several years even before the war, the country has created a startup ecosystem that is scalable and global [26].

In recent years, despite the unfavourable political and economic situation in Ukraine, the startup industry has grown by an average of 20–30% per year. The number of internationally recognized Ukrainian startups also increased, and the amount of their funding increased. The number of international startup deals in 2022 remained at the previous year's level and amounted to 33, but the amount of their funding increased more than 2 times or from 40.7 to 87.6 million dollars. [26, 60] The startup ecosystem of Ukraine was in the process of active formation and was characterized by positive dynamics of its development before the start of the full-scale war in 2022. But 2022–2023 showed that the development of startups slowed down due to the difficult economic situation of the country.

It is worth noting that one of the important problems for Ukrainian startups is the lack of financial resources for their development. Startups have the opportunity to attract funds from several sources, which depend on what stage of the life cycle they are at. Thus, in the early stages (pre-seed), entrepreneurs typically use their own funds and/or borrow from friends and family (bootstrapping) before seeking financial assistance from angel investors, accelerators, incubators, etc. [61]. During the seed stage, startups typically raise capital from angel investors and venture capital firms to support their minimum viable product (MVP) and develop product-market fit. In many startup ecosystems, venture capital firms and government agencies may also typically collaborate to provide startups with seed funding. If successful at the seed stage, startups move on to Series A and B financing rounds, which are primarily provided by venture capital firms, to raise significant financial resources [61, 62].

Thus, startups are a potentially important sector of the economy and a fundamental asset for the future growth of Ukraine. Startups have become a modern challenge in the context of the intensification of the development of the innovative and digital economy.

4 RESEARCH RESULTS

Based on the analysis of the startup ecosystems of the countries of the world that are leaders and potential leaders in this field, and their development trends, it was determined that the construction of startup ecosystems is based on a new approach, which involves considering the ecosystem of startup development as a dynamic self-regulating and evolutionary system with a high level of uncertainty [44]. This approach requires a new type of cognitive skills for cooperation between different types of organizations, such as private, public, public and municipal.

The experience of the world's leading countries shows that it is reasonable to consider the policy of the development of startup entrepreneurship separately from the general policy of supporting entrepreneurship in the country. Successful development of the innovative private sector requires strategic collaborations within the ecosystem. By itself, the public sector cannot act as a driver of economic growth, but it can provide systemic support for the development of local start-up entrepreneurial initiatives. The public sector can take the lead in creating long-term platforms and subsystems of local startup ecosystems. This includes fostering cooperation and sustainable exchange of information between all ecosystem participants, as well as formulating a development vision for all ecosystem participants.

Thus, the countries of the world have accumulated a lot of experience in forming and supporting the development of the startup ecosystem. The experience of successful startup ecosystems should be implemented in domestic practice, but taking into account its features, existing strengths and weaknesses.

In accordance with the above, it is advisable to develop a methodical approach to the formation of the startup ecosystem in Ukraine, which involves its formation according to the main stages of the life

cycle of startups [25]. The essence of the methodological approach is to create a network of relevant innovation structures and institutions, whose activities are aimed at ensuring the effective implementation of a startup at the main stages of its life cycle. In addition, it is proposed to involve third-party organizations into the startup ecosystem, cooperation with which is carried out within the framework of concluded memoranda and agreements. The following general scheme of a methodical approach to the formation of the startup ecosystem in Ukraine is proposed, taking into account the best global practices:

The first stage. Improvement of the organizational structure of the university. In order to establish many startups, it is necessary to form a high entrepreneurial culture in society, in this regard; the role of the education system in the creation, development and success of startups is multifaceted. For the development of entrepreneurship culture, it is important to include appropriate content at all levels of education [63]. But thanks to higher education, a person receives the most modern knowledge and skills in a narrower academic or professional field, which, among other things, allow him to independently solve the most complex problems in an innovative way [64, 65]. The introduction of multidisciplinary content on startup entrepreneurship in higher education is of great importance for future high-tech startups, as the highly educated population has the greatest potential to create new startups.

Universities play an important role in the creation, development and success of startups [66]. Their contribution consists in training and imparting knowledge, supporting an innovative environment and providing students with opportunities to develop entrepreneurial skills [66]. Universities should become not only a place for training specialists in various specialities, which can potentially be useful for startups, but also places where scientific research and development of new technologies are carried out, innovation development is ensured, providing access to laboratory facilities, financial support and expert knowledge [67]. Many of today's successful startups emerge from academic laboratories and research groups. Many European entrepreneurial universities have startup incubators and accelerators where students and graduates can develop their ideas and get support from experts and investors. Universities can create an environment that promotes the development of entrepreneurial spirit among students [68]. This can be achieved through the organization of entrepreneurial clubs, innovation forums, hackathons and other events that encourage students to think about creating their own businesses. In addition, successful entrepreneurs and experts can be involved in mentoring [68].

Universities can collaborate with local enterprises and business communities, ensuring integration between the learning process and real business needs [68], and assists students and graduates in commercializing their innovative developments, for example, justifying the feasibility of startup ideas, and drawing up business plans, providing assistance in patenting, licensing and other forms of intellectual property protection [67].

Thus, at the first stage, for the formation of an effective ecosystem of startups, it is necessary to ensure the development of Ukrainian universities as entrepreneurial and innovative, which is of key importance for the creation, development and success of startups.

One of the key elements of the startup ecosystem should be the startup center and the startup school. The startup center provides the following functions: organization of presentations, crash tests, brainstorming; conducting educational seminars, trainings, business games with the participation of entrepreneurs; support for participation in all-Ukrainian and international competitions; development of partnership and mentoring with leading startup projects, startup hubs, investment funds. A startup school is opening within the startup center. The goal of the startup school is to promote creative thinking, professional development and the acquisition of skills necessary to generate business ideas and create your own startups. The startup school interacts: with university departments to implement the educational process and participate in the formation of an innovative ecosystem of startups; with the heads of faculties (institutes) of higher education institutions for the purpose of preparing start-up projects of the appropriate focus and their expert evaluation; with other structural divisions of the university for the quality performance of the tasks and functions assigned to the startup school.

As important structural divisions of a startup school, it is advisable to organize:

- co-working (from English Co-working – “to work together”) – to create a workspace where startup teams can form, work together on the

development of their projects, exchange ideas and experience, discuss business ideas, etc.;

- startup and business idea contests – regular holding of startup project contests, their examination and selection make it possible to identify promising business ideas and directions for their improvement, draw attention to the startup movement, attract stakeholders, including potential investors;
- startup festival – an event aimed at spreading awareness of a wide range of stakeholders of the startup movement in startups at their initial stage of existence, finally forming startup teams, attracting the attention of investors;
- bank of startup ideas – formation of a database of startup ideas for their further improvement or implementation, formation of business, government, and public proposals for their implementation.

The startup center, the startup school and all its structural elements should comprehensively support the first stage of the startup life cycle regarding the formation of startup teams, the development of the startup idea and the creation of the MVP.

One of the key structural elements of the startup ecosystem is a Science Park, which is created at the initiative of a higher education institution in accordance with the current legislation, in particular the Law of Ukraine "On Science Parks", and with the involvement of representatives of the business sector by pooling the founders' contributions for the purpose of developing scientific and technical and innovative activities for the commercialization of the results of scientific research and their implementation on domestic and international sales markets.

In accordance with the formation of ecosystem support for startups at various stages of their life cycle, it is recommended to include the following structural elements in the Science Park: business incubator, technological platform, patent office, startup accelerator, Technology Park, legal clinic, marketing research centre, fund endowment.

The second stage of the methodical approach to the formation of the startup ecosystem in Ukraine involves the involvement of third-party organizations in the startup ecosystem. At this stage, agreements and memorandums of support are concluded with startup accelerators; memoranda are signed with venture funds; memorandums and agreements are signed with international financial organizations and development funds.

Key actions to ensure the life cycle of a startup from the side of the ecosystem at the appropriate stages include [25]:

- the first stage (initial): team formation (formation of creative thinking, activation of entrepreneurial activity); business idea generation (assessment of socio-economic development problems, determination of business and society needs, assistance in justifying business ideas, examination of startup ideas);
- the second stage (seeding): creation of MVP (estimation of necessary resources for creation of MVP, resource and technical assistance in creation of MVP); building business models (assessment of potential sales markets, definition of consumer segments, key values of the startup's product, sales channels, relations with consumers, key activities, partners, costs and revenues); creation of a prototype (assessment and assistance in providing resources, organization of creation of a prototype);
- the third stage (executive): development of a business plan (assessment and expert consulting on the justification and development of the main sections of the business plan); strategy development (assessment of the competitive environment, justification and development of a startup strategy);
- the fourth stage (final): protection of intellectual property (assessment of the technology market and assistance in the preparation of protective documents for the protection of intellectual property rights); search for financing (assessing the market for financial resources, attracting resources for financing startups).

Thus, an ecosystem of startups can be formed, which provides comprehensive support for startups throughout their life cycle [25]. In addition, when building an ecosystem of startups, it is necessary to take into account the complexity of the problem of finding a startup idea. The product that the startup offers must be in demand, take into account the needs of end users. That is why it is necessary to pay considerable attention to the integration of education, science, business and government during the formation of a startup idea. Considering the system of interrelationships between education, science and business within the ecosystem of startups, it should be noted that the development of a

business idea, as well as the innovation process as a whole within the university, is built based on the activities of subjects that generate new knowledge (scientists, teams of scientists, students, graduate students, departments, laboratories, divisions) and entities that commercialize developments (departments of scientific and startup activities, technology transfer centers, etc.).

At the stage of developing an idea and conducting scientific research, the elements of the managed subsystem, expressed through the structural elements of subdivisions and departments, should be involved: institutes, faculties, scientific laboratories, scientific and educational centers, startup centers (schools), etc. At the stage of implementing the results of scientific activity into practice, engineering centers, specialized research institutes and centers, etc. are involved [66].

The proposed methodical approach to the formation of an ecosystem for the formation of startup companies in Ukraine was tested and implemented at O. M. Beketov National University of Urban Economy in Kharkiv, and the organizational structure of its startup ecosystem was developed:

- Science Park in the form of a limited liability company;
- a startup center that provides for the following functions: organization of presentations, crash tests, brainstorming; conducting educational seminars, trainings, business games with the participation of entrepreneurs; support for participation in all-Ukrainian and international competitions; development of partnership and mentoring with leading startup projects, startup hubs, investment funds; including a start-up school;
- a business incubator, which provides for the creation of conditions conducive to the creation and development of small innovative enterprises, ready for the effective functioning of the market both in the university incubator and after leaving it;
- the business environment as a customer of the final innovative product, in particular, clusters of the city of Kharkiv and other regions of Ukraine.

Therefore, a methodical approach to the formation of the startup ecosystem in Ukraine around the university can form the basis for the development of innovative entrepreneurship and contribute to the development of the startup movement in the country and its socio-economic development in particular.

5 CONCLUSIONS

As a result of the conducted research, the following scientific results of a theoretical and applied nature were obtained:

1) It was determined that the startup ecosystem is a network association of organizational structures of various forms of ownership that provide support for the creation of innovative startup projects at various stages of their life cycle with subsequent scaling.
2) On the basis of the study of theoretical and practical aspects of the creation of startup ecosystems in the countries of the world, their features and models in different countries, as well as the advantages they give to startups, are determined.
3) According to the results of the research of the experience of the leading countries of the world in the context of the development of startups and ecosystems of their support by conducting a cluster analysis based on data from 55 countries of the world according to indicators: the Index of the most startup-friendly countries in the world, the Global Index of startup ecosystems and the number of startups per 1,000 population for 2021, three clusters of countries with low, medium and high values of these indicators are highlighted. The leading countries included in the cluster with high indicator values are the USA, Singapore and Estonia. Ukraine is included in the cluster with the lowest indicator values. According to the results of the study of 55 countries of the world according to the following indicators: GDP per capita, number of startups per 1,000 population and the Global Innovation Index and the conducted cluster analysis, three groups of countries with low, medium and high values of these indicators were identified. The USA, Singapore and Estonia were also identified as the leading countries included in the cluster with high indicator values. Ukraine is also included in the cluster of countries with the lowest indicator values.
4) The peculiarities of the startup ecosystems of the leading countries and potential leaders in this field were studied and the characteristic features of their startup ecosystems were

summarized according to the following characteristics: state regulation, financing of startups, training (academic support), information support, infrastructural support.

5) A methodical approach to the formation of the startup ecosystem in Ukraine has been developed, which is characterized by a specific set of the country's institutions, a complex combination of efforts of which will provide support for the formation of startups (at all stages of their life cycle) and promote the activation of innovative entrepreneurship.

Prospects for further research are theoretical and methodological substantiation of possible strategies for the development of the startup ecosystem in the conditions of the post-war reconstruction of the national economy of Ukraine, as well as the development of practical recommendations for the use of public-private partnerships and crowdfunding as effective mechanisms for financing startups.

REFERENCES

[1] V. Ye. Khaustova, O. I. Reshetnyak, and M. M. Khaustov, "Promising Areas of IT Development in the World," The Problems of Economy, no. 1, pp. 3-19, 2022. [Online]. Available: https://doi.org/10.32983/2222-0712-2022-1-3-19.

[2] M. O. Kyzym, V. Ye. Khaustova, and N. V. Trushkina, "Network Economy: Evolution of Development, Prerequisites for the Formation of the Conception, Conceptual Approaches to the Definition," Business Inform, no. 11, pp. 40-51, 2022. [Online]. Available: https://doi.org/10.32983/2222-4459-2022-11-40-51.

[3] A. Kwilinski, "The Relationship between Sustainable Development and Digital Transformation: Bibliometric Analysis," Virtual Econ., vol. 6, no. 3, pp. 56-69, 2023. [Online]. Available: https://doi.org/10.34021/ve.2023.06.03(4).

[4] X. Ma, I. Gryshova, O. Reshetnyak, V. Khaustova, M. Shcherbata, M. Khaustov, and D. Bobrovnyk, "Assessment of the Impact of Scientific and Technical Activities on the Economic Growth of World Countries," Sustainability, vol. 14, no. 21, 14350, 2022. [Online]. Available: https://doi.org/10.3390/su142114350.

[5] A. Rafique, "The Role of Startups in Economic Prosperity of Developing Countries," Sentient Solutions, 2017. [Online]. Available: https://sentientso.com/the-role-of-startups-in-economic-prosperity-of-developing-countries/.

[6] S. Blank and B. Dorf, The Startup Owner's Manual, 1st ed. Willey, NY, USA, 2020. [Online]. Available: https://www.perlego.com/book/1425741/the-startup-owners-manual-the-stepbystep-guide-for-building-a-great-company-pdf.

[7] M. Bliemel, R. Flores, S. de Klerk, and M. P. Miles, "Accelerators as Start-up Infrastructure for Entrepreneurial Clusters," Entrep. Region. Dev., vol. 31, no. 1-2, pp. 133-149, 2019. [Online]. Available: https://doi.org/10.1080/08985626.2018.1537152.

[8] S. J. Breschi, C. Lassebie, and C. Menon, A Portrait of Innovative Startups across Countries, OECD Science, Technology and Industry Working Papers, no. 2018/02, France: OECD Publishing, 2018.

[9] J. Gans, L. Scott Erin, and S. Stern, "Strategy for Start-ups," Harv. Bus. Rev., no. 5, 2018. [Online]. Available: https://hbr.org/2018/05/doentrepreneurs-need-a-strategy#strategy-for-start-ups.

[10] M. Henry et al., "A Typology of Circular Start-ups: Analysis of 128 Circular Business Models," J. Clean. Prod., vol. 245, 118528, 2020.

[11] A. Skala, Digital Startups in Transition Economy. Cham, Switzerland: Springer International Publishing, 2019.

[12] N. Bielikova and I. Hubarieva, "Features of the Formation of Adaptive Startup Management Systems in an Instable Environment," The Problems of Economy, no. 3, pp. 168-173, 2023. [Online]. Available: https://doi.org/10.32983/2222-0712-2023-3-168-173.

[13] O. Dymchenko et al., "Modeling the Influence of Startup Ecosystem Components: Entrepreneurial Aspect," Scientific Horizons, vol. 25, no. 11, pp. 131-140, 2022. [Online]. Available: https://doi.org/10.48077/scihor.25(11).2022.131-140.

[14] L. Frolova, K. Zhadko, O. Ilyash et al., "Model for Opportunities Assessment to Increase the Enterprise Innovation Activity," Business: Theory and Practice, vol. 22, no. 1, pp. 1-11, 2021.

[15] O. A. Havrysh et al., Management of Startup Projects. Kyiv: KPI named after Igor Sikorskyi, Polytechnic Publishing House, 2019.

[16] I. Hubarieva, "Startup Ecosystems in Ukraine: Problems, Trends, Prospects," Economy and Society, no. 55, 2023. [Online]. Available: https://doi.org/10.32782/2524-0072/2023-55-95.

[17] N. M. Kulyk, "Implementation and Development of Startups in the Ukrainian Economy," Efficient Economy, no. 4, 2020. [Online]. Available: https://doi.org/10.32702/2307-2105-2020.4.94.

[18] N. Yu. Podolchak, V. Ya. Karkovska, and Ya. V. Levytska, "Ecosystems of Startup Development as a Principle of Implementation of State Innovation Policy," Investments: Practice and Experience, no. 8, pp. 82-87, 2021. [Online]. Available: https://doi.org/10.32702/23066814.2021.8.82.

[19] I. Savin, K. Chukavina, and A. Pushkarev, "Topic-Based Classification and Identification of Global Trends for Startup Companies," Small Bus Econ., vol. 60, pp. 659-689, 2023. [Online]. Available: https://doi.org/10.1007/s11187-022-00609-6.

[20] O. Trofymenko and O. Ilyash, "A Comparative Analysis of the Development of Startups in the World Economy," The Problems of Economy, no. 3, pp. 25-30, 2023. [Online]. Available: https://doi.org/10.32983/2222-0712-2023-3-25-30.

[21] M. V. Tymoshenko, "Innovative startups as a factor of economic development: best global practice and experience of Ukraine," Sci. Notes Lviv Univ. Bus. Law, iss. 33, 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6642172.

[22] C. Mason and R. Brown, Entrepreneurial Ecosystems and Growth Oriented Entrepreneurship. OECD, 2014.

[23] B. L. Herrmann et al., The Global Startup Ecosystem Ranking 2015. Tech. Rep. Compass, 2015.

[24] M. Kyzym et al., "Cluster Analysis Usage as Prerequisite for Implementing Strategies of Countries Startup Ecosystems Development," Lect. Notes Netw. Syst., pp. 290-301, 2023.

[25] M. M. Khaustov, Startups: Creation and Scaling. Kharkiv: FOP Liburkina L.M., 2023. [Online]. Available: https://ndc-ipr.org/publications/?page=1&filter=pmonografiyip.

[26] The Global Startup Ecosystem Report 2023 (GSER 2023). [Online]. Available: https://startupgenome.com/report/gser2023.

[27] International best practices on supporting startup ecosystem. [Online]. Available: https://www.mekongbiz.org/wp-content/uploads/2017/02/International-best-practices-on-supporting-startup-ecosystem.pdf.

[28] Startup Commons. [Online]. Available: https://www.startupcommons.org.

[29] I. V. Lytvyn, "Formation and development of startup ecosystems," Sci. Bull. Uzhhorod Natl. Univ., iss. 41, pp. 61-65, 2022. [Online]. Available: https://doi.org/10.32782/2413-9971/2022-41-11.

[30] H. Caleb, "Savannah's startup ecosystem," Jan. 27, 2014. [Online]. Available: http://www.hiimcaleb.me/savannahs-startup-ecosystem/.

[31] B. Spigel, "The relational organization of entrepreneurial ecosystems," Entrep. Theory Pract., vol. 41, iss. 1, pp. 49-72, 2015.

[32] M. Ranga and H. Etzkowitz, "Triple Helix Systems: An Analytical Framework for Innovation Policy and Practice in the Knowledge Society," in Entrep. Knowl. Exch., New York: Routledge, 2015.

[33] H. Etzkowitz, "Innovation in Innovation: The Triple Helix of University-Industry-Government Relations," Soc. Sci. Inf., vol. 42, pp. 293-337, 2003.

[34] H. Etzkowitz and L. Leydesdorff, "The Triple Helix–University-Industry-Government Relations: A Laboratory for Knowledge Based Economic Development," EASST Rev., vol. 14, pp. 14-19, 1995.

[35] D. Sarpong et al., "Organizing Practices of University, Industry and Government That Facilitate (or Impede) the Transition to a Hybrid Triple Helix Model of Innovation," Technol. Forecast. Soc. Chang., vol. 123, pp. 142-152, 2017.

[36] L. M. Ranga, J. Miedem, and R. Jorna, "Enhancing the Innovative Capacity of Small Firms through Triple Helix Interactions: Challenges and Opportunities," Technol. Anal. Strateg. Manag., vol. 20, pp. 697-716, 2008.

[37] E. G. Carayannis and D. F. Campbell, Knowledge Creation, Diffusion, and Use in Innovation Networks and Knowledge Clusters: A Comparative Systems Approach Across the United States, Europe, and Asia. Santa Barbara, CA, USA: Greenwood Publishing Group, 2006.

[38] A. Marcovich and T. Shinn, "From the Triple Helix to a Quadruple Helix? The Case of Dip-Pen Nanolithography," vol. 49, pp. 175-190, 2011.

[39] G. S. Nyman, "University-Business-Government Collaboration: From Institutes to Platforms and Ecosystems," Triple Helix, vol. 2, pp. 1-20, 2015.

[40] J. M. Unger et al., "Human Capital and Entrepreneurial Success: A Meta-Analytical Review," J. Bus. Ventur., vol. 26, pp. 341-358, 2011.

[41] Aspen Network of Development Entrepreneurs, Entrepreneurial Ecosystem Diagnostic Toolkit. Washington DC: Aspen Institute, 2013.

[42] C. Ziakis, M. Vlachopoulou, and K. Petridis, "Start-Up Ecosystem (StUpEco): A Conceptual Framework and Empirical Research," J. Open Innov. Technol. Mark. Complex, vol. 8, 35, 2022. [Online]. Available: https://doi.org/10.3390/joitmc8010035.

[43] M. O. Kravchenko and Ye. A. Hazaryan, "Peculiarities of formation of startup ecosystem in Ukraine," East. Eur. Econ. Bus. Manag., iss. 5(22), pp. 90-97, 2019.

[44] Startups disrupting industries and changing the world– and doing it all at scale. [Online]. Available: https://news.microsoft.com/en-cee/2022/09/20/startups-disrupting-industries-and-changing-the-world-and-doing-it-all-at-scale/.

[45] A. Bagrecha, "India's Startup Stats," Jordensky, 2022. [Online]. Available: https://www.jordensky.com/blog/unlocking-indias-startup-stats-your-data-delight#:~:text.

[46] C. Thathoo, "Number Of Indian Startups Up 300X Since 2014: MoS Jitendra Singh," Inc42, Apr. 10, 2023. [Online]. Available: https://inc42.com/buzz/number-of-indian-startups-up-300x-since-2014-mos-jitendra-singh/.

[47] A. Minaev, "Startup Statistics (2023): 35 Facts and Trends You Must Know," First Site Guide, May 30, 2024. [Online]. Available: https://firstsiteguide.com/startup-stats/.

[48] Startup Ranking. [Online]. Available: https://www.startupranking.com.

[49] J. Wise, "How Many Startups Are Created Each Year? (Worldwide Stats)," EarthWeb, Aug. 10, 2023. [Online]. Available: https://earthweb.com/how-many-startups-are-created-each-year/.

[50] Global Startup Ecosystem Index 2022. [Online]. Available: https://startupgenome.com/reports/gser-fintechedition.

[51] "Most Startup Friendly Countries in the World 2021," CEOWorld Magazine, Apr. 26, 2021. [Online]. Available: https://ceoworld.biz/2021/04/26/most-startupfriendly-countries-in-the-world-2021/.

[52] "Discover the World's Best Startup Ecosystems," StartupBlink. [Online]. Available: https://www.startupblink.com.

[53] "Global Innovation Index (GII)," World Intellectual Property Organization (WIPO). [Online]. Available: https://www.wipo.int/global_innovation_index/en/.

[54] "World Development Indicators," The World Bank. [Online]. Available: https://databank.worldbank.org/source/world-development-indicators.

[55] "The Startup Ecosystem of United States," StartupBlink. [Online]. Available: https://www.startupblink.com/startup-ecosystem/united-states.

[56] "The Startup Ecosystem of Singapore," StartupBlink. [Online]. Available: https://www.startupblink.com/startup-ecosystem/singapore.

[57] N. Pangarkar and P. Vandenberg, Singapore's Ecosystem for Technology Startups and Lessons for Its Neighbors: Country Report No. 2, Asian Development Bank, Jun. 2022. [Online]. Available: https://www.adb.org/sites/default/files/publication/804956/singapore-ecosystem-technology-startups.pdf.

[58] S.-K. Saluveer and M. Truu, Startup Estonia: White Paper 2021–2027, 2020.

[59] "The Startup Ecosystem of Estonia," StartupBlink. [Online]. Available: https://www.startupblink.com/startup-ecosystem/estonia.

[60] The Global Startup Ecosystem Report 2022, Startup Genome. [Online]. Available: https://startupgenome.com/report/gser2022.

[61] Future of the Thai Startup and Venture Capital Ecosystem, Deloitte Consulting Limited, 2023. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/sg/Documents/human-capital/sea-hc-venture-capital-ecosystem-thai-feb2023.pdf.

[62] R. Baldridge, "What Is A Startup? The Ultimate Guide," Forbes, Jun. 3, 2024. [Online]. Available: https://www.forbes.com/advisor/business/what-is-a-startup/.

[63] A. K. Dwivedi, Educating Youth to Launch and Sustain Business Start-Ups, SSRN, 2016. [Online]. Available: https://doi.org/10.2139/ssrn.2858755.

[64] D. Leo-Paul et al., "The Impact of Entrepreneurial Education on Technology-Based Enterprises Development: The Mediating Role of Motivation," Administrative Sciences, no. 11, p. 105, 2021. [Online]. Available: https://doi.org/10.3390/admsci11040105.

[65] V. Barba-Sánchez and C. Atienza-Sahuquillo, "Entrepreneurial Intention Among Engineering Students: The Role of Entrepreneurship Education," European Research on Management and Business Economics, vol. 24, no. 1, pp. 53-61, 2018. [Online]. Available: https://doi.org/10.1016/j.iedeen.2017.04.001.

[66] M. O. Kyzym, V. Ye. Khaustova, O. I. Reshetnyak et al., Integration of Education and Science in Ukraine on the Basis of the Future University Model, Kharkiv: FOP L. M. Liburkina, 2021.

[67] A. Deshpande and S. Guthrie, Entrepreneurial-University Ecosystem: An Overview of Key Concepts, RAND Europe, 2019. [Online]. Available: https://www.rand.org/content/dam/rand/pubs/working_papers/WR1300/WR1304/RAND_WR1304.pdf.

[68] S. Enginsoy, "Role of Universities in Startup Ecosystems," StartupBlink. [Online]. Available: https://www.startupblink.com/blog/role-of-universities-in-startup-ecosystems/.

Assessment of the Impact of Social Responsibility of Integrated Business Forms on Regional Development

Iryna Stankevych[1], Inna Yatskevych[1], Hanna Sakun[1], Nataliya Shyriaieva[1], Oksana Vasylenko[2] and Eduard Siemens[2]

[1] *Department of management and marketing, Odessa State Academy of Civil Engineering and Architecture, Didrikhsona Str. 4, 65029 Odesa, Ukraine*
[2]*Departments of Electrical Engineering, Mechanical Engineering and Engineering Management, Anhalt University of Applied Sciences, Bernburger Str. 57, 06366 Köthen, Germany*
arnika@ukr.net, innav2018@gmail.com, sakun.anna@ukr.net, mop@odaba.edu.ua, {oksana.vasylenko, eduard.siemens}@hs-anhalt.de

Keywords: Assessment, Construction Industry, Factors, Integration Form, Participants of Partnership Relations, Region, Socio-Economic Development, Social Responsibility.

Abstract: Research has shown that the current entrepreneurial environment promotes the formation and growth of integrated interaction forms. These integrated relationships foster alignment among key stakeholders, including partners, consumers, employees, society, and government authorities, emphasizing the importance of social responsibility. A particularly pressing issue is the implementation of social responsibility through integrated forms as a community, necessitating further research to evaluate their impact on regional socio-economic development. Using statistical data and regression analysis, researchers developed a model to assess how regional employment and unemployment rates, alongside the inflation index, influence the effect of integration forms on socio-economic development. Normative fluctuations for unemployment (2%) and employment (7,5%) in the region were identified. The regression analysis resulted in the calculation of an indicator (*IF_avg)* - to measure the contribution of integrated forms to socio-economic development. This indicator enables the assessment of whether such forms positively impact regional development. Testing this approach on the construction company "Gefest" in the Odessa region demonstrated a positive impact, with a calculated result of 1,26%. While these findings provide initial insights, they are not without limitations. Future improvements could incorporate additional indicators such as investment volume, innovative activity levels, and labor productivity to refine the model and enhance its accuracy in evaluating the influence of integration forms.

1 INTRODUCTION

Trends in the entrepreneurial environment contribute to changes in the formation, management, and organization of interactions within partnerships. The specific conditions of the Ukrainian environment determine the particular responsibility of business leaders toward their partners, employees, consumers of products (and users of services), and society as a whole – this is social responsibility.

The social responsibility of integrated forms of business involves addressing social conflicts that arise in the process of interaction with other partners, without leading to negative consequences such as increased unemployment or worsening poverty. It also includes the preservation of ecology and various other aspects of responsibility, which entail not only the implementation of ideas but also the realization of innovative initiatives, interaction with partners, and other forms of accountability.

The role of integrated forms in modern society goes beyond merely creating jobs and generating profits; it requires setting new objectives. The more civilized the business environment in a country, the greater its impact on social life within cities, regions, and the nation as a whole. Consumers of products and society evaluate the activities of integrated forms not just based on production and financial outcomes but also on how

they conduct their operations: whether they care about their employees and how well they align with the interests of other market participants, residents of the area, and society as a whole. Therefore, the question of social responsibility among participants in partnership relations within integrated forms of cooperation and the assessment of the level of impact on the socio-economic development of the region becomes increasingly relevant.

2 LITERATURE REVIEW

Among the leading works dedicated to the assessment of social responsibility among participants in partnership relations within integrated forms of cooperation, the following scholars' contributions should be highlighted [1... 5].

In [1], it is stated that social responsibility is an ethical focus for individuals and companies whereby they seek to take action and be accountable for practices that benefit society. Social responsibility means that individuals and companies must act in the best interests of their environment and society as a whole.

The concept of CSR in [2] is defined as a business's commitment to responsibly managing its operations' social, environmental, and economic effects in line with public expectations.

According to [3], corporate social responsibility is a mechanism for businesses to assess the impact they have on society and put responsible, ethical policies in place to support individuals, the local community, the marketplace, and the environment.

In [4], CSR is described as a business model that ensures companies operate in a manner that is ethical and beneficial for society at large. This involves taking into account the social, economic, and environmental impacts of business operations. By adopting CSR, companies voluntarily commit to contributing positively to societal goals.

The discussion in [5] revolves around CSR implementation, which can involve a 'built-in' and 'bolt-on' approach. The former is strategic, incorporating socially responsible behaviours into companies' operations, processes, and decision-making.

It should be noted that social responsibility is studied in some works as an element of moral duty [6; 7], a company's development strategy that guarantees the effective formation of sustainable development of the company [8-11], the formation of corporate culture [12] and partnerships [13]. Today, it is clear that the business environment has a better chance of success than its competitors focused solely on economic gain through the formation and development of integrated forms. However, the participants of integrated forms of cooperation face new characteristics of social responsibility.

3 METHODOLOGY

The entrepreneurial environment within integrated forms of cooperation fosters the development of new components that characterize social responsibility in integrated forms of cooperation (IF), which expand its role for participants in partnership relationships within integration (PRI). These components include commitments to employees, consumers of products (services, works), partners, and society. At the same time, the very existence of PRI depends on the external environment. The components of this environment include consumers, suppliers, local communities, media, and others. In other words, this is the public environment that actively influences the achievement of the PRI's goals. Therefore, PRI must balance purely economic objectives with the economic and social interests of the environment. This situation has become the basis for determining the main characteristics of IF: its role for all participants in these relationships; factors that influence the formation and development of IF, and the assessment of the level of IF's impact on the socio-economic development of the region. As a result, we will examine the identified components of IF in more detail.

An inevitable consequence of the formation and development of integration cooperation is the emergence of social responsibility. This responsibility is primarily placed on the PRI, as it seeks to create comfortable conditions for interaction for all participants. The authors have determined that each party involved in these relations, forming its own social responsibility in accordance with its own goals, is responsible for the social responsibility of all participants (Fig. 1).

From Figure 1, it is evident that participants in integrated forms of cooperation aim to create interconnected socially responsible actions that

consider the interests and outcomes of all parties involved: the entrepreneurial environment, employees, executive authorities, consumers, and society during the functioning and development of these relationships. Here are some key aspects of social responsibility in IF:

- Relations between PRI and Employees: The aspiration to create proper working conditions, ensuring safety and a decent standard of living.
- Relations between PRI and Consumers: The provision of high-quality products (services, works).
- Relations between PRI and Executive Authorities: Contributing to the resolution of issues related to reducing unemployment and other social problems at both state and local levels.
- Relations between PRI and Society: Promoting sustainable economic development at various levels.

Therefore, we will analyze the factors that influence the formation and development of IF using the example of the Odessa region in the construction sector, with the aim of forecasting the level of influence of IF on the socio-economic development of the region.

The justification of the factors influencing the formation and development of IF allows for the reflection of existing trends in the development of integration forms, the identification of problematic aspects, and subsequently, the adoption of managerial decisions aimed at the development of enterprises in the construction sector, thereby improving the socio-economic situation of the Odessa region as a component of a united country.

A significant number of researchers have dedicated studies to examining the relationship between various factors based on statistical reliability regarding the strength of correlation, its adjustment accounting for the selection of functionally dependent parameters, and selecting basic technical parameters for deriving a normative formula, which can be implemented through regression analysis.

Today, according to the authors, the use of regression analysis allows for the establishment of a theoretical expression of the relationship between characteristics, that is, the form of this relationship, through the construction of a regression equation. The results of this analysis help identify priority directions and, based on the primary factors, enable forecasting, planning their development, and making managerial decisions. Regression analysis is closely related to correlation analysis, which studies the direction and strength of the connection between independent variables, as well as the form of dependency between them, embodied in the regression function [14-16].

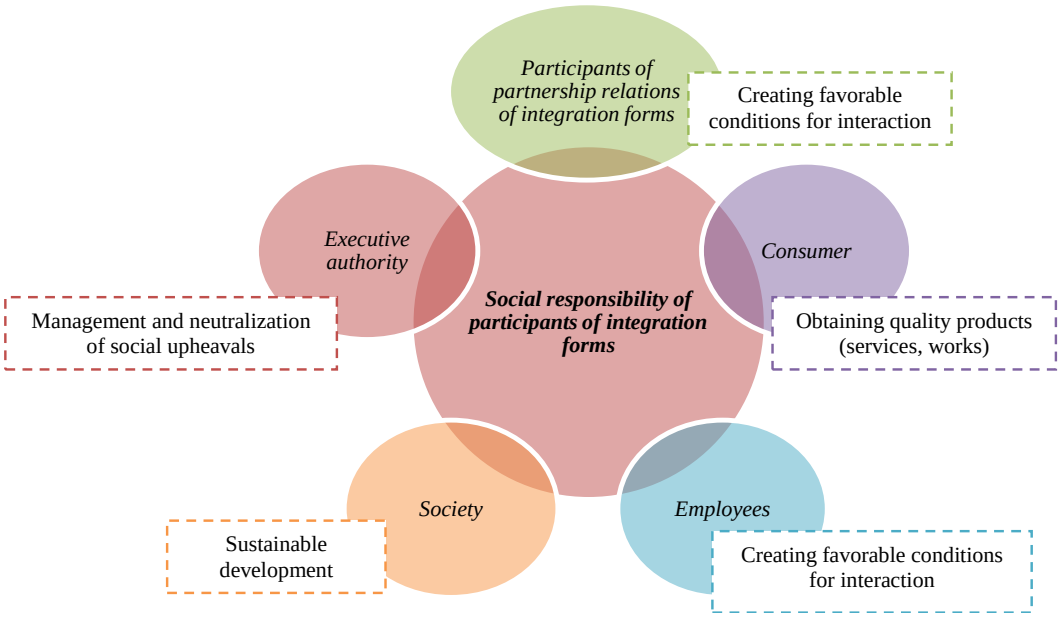


Figure 1: Social responsibility in integrated forms of cooperation.

4 RESULTS AND DISCUSSION

Particular attention is drawn to evaluating the social responsibility of integrated forms of cooperation concerning their interactions with society. This is linked to expenditures on social activities, which are regarded not merely as simple costs but as vital assets that should be properly utilized. Currently, the human factor significantly influences the competitiveness and effectiveness of a company's operations, primarily through interactions between the company's employees and the consumers of its products.

Integrated forms of cooperation encompass complex economic interactions, where integrated enterprises jointly share revenues, bear responsibility for losses and risks, manage all stages of production and product realization collaboratively, and respond swiftly to changes in the country's political situation, among other factors.

Considering integrated forms as a distinct economic behavior style, their functioning is based on pursuing new business development opportunities, focusing on innovation, effectively utilizing limited resources from various sources, and delivering high-quality products. These aspects are crucial for achieving economic and social development at the regional level. Integrated forms play a significant role in the region's socio-economic development by investing in regional programs, creating jobs, and enhancing product accessibility for society.

When assessing the contribution of integrated forms to the socio-economic development of the region, one potential indicator is the employment rate of employees within the enterprise. Thus, when determining the level of contribution of integrated forms to the socio-economic development of the region, the ratio of employed workers to the total number of employees in the industry (*E)* is considered.

To determine the contribution of integrated forms *(IF)* to the socio-economic development of the region *(IF_avg)*, it is recommended to consider the unemployment rate *(U)* and inflation rate *(I)*. This recommendation is justified for several reasons. First, an increase in unemployment reduces the number of consumers and, consequently, the consumption of goods. Second, a decrease in employment within the industry complicates consumer access to products due to a reduction in the number of enterprises or their branches (Table 1) [17-19].

Table 1: Indicators influencing the level of contribution of integrated forms to the socio-economic development of the region selected for research.

Year	Employed in Odessa region, thousand people	Employed in the construction industry, thousand people	Thousand people Unemployment in Odessa region, %	Inflation index, %
2014	1009,4	24,0	7,0	124,9
2015	1016,2	41,0	6,7	143,3
2016	100,6	40,2	6,9	112,4
2017	986,6	39,5	7,4	113,7
2018	1001,9	42,6	6,6	109,
2019	1020,1	45,1	6,1	104,1
2020	1028,4	45,6	6,7	105,0
2021	1052,1	46,4	7,2	110,0
2022	*	*	*	126,6
2023	*	*	*	105,1

* Information is unavailable due to the state of war in Ukraine.

Table 1 presents data on employment in the construction industry and the unemployment rate in the Odessa region. The data in Table 1 allows for the determination of the dependence of *IF_avg* and the normative values for establishing the boundaries of the impact of unemployment and inflation.

Based on the data in Table 1, the following model has been obtained (1):

$$IF\_avg = f(E; U; I), \quad (1)$$

where E is the employment level in the industry in the region, %; *U* – is the unemployment rate in the region, %; *I* - is the inflation index, %.

The authors recommend normative values for determining the boundaries of the impact of unemployment and inflation based on statistical research:

- the normative value for the fluctuation of the unemployment percentage in the region (KU) is 2%;
- the normative value for the fluctuation of the employment percentage in the industry in the region (KE) is 7,5%.

At the same time, the level of inflation significantly affects the effectiveness of integrated forms, as it reflects changes in the overall price level of goods and services purchased by the population for consumption.

The inflation index can have three conditions: $I > 100\%$, $I < 100\%$, and $I = 100\%$.

Considering the dynamics of changes in the inflation index, it is important to analyze the impact of IF on the socio-economic development of the region. To achieve this goal, we will transform the inflation index to ensure stable results within certain values. In practical applications, we can use the mathematical function arctg, with the formula $y = \text{arctg}(x)$.

Thus, when the argument of the function (x) changes from 0 to ∞, the value of the function (y) changes in the range from 0 to $\frac{\pi}{2}$. As a result, we note that for results when I = 100%, the contribution of the multiplier associated with the inflation index in determining the level of impact of IF on the socio-economic development of the region will be minimal. For this:

1) We integrate the value of the inflation index from percentage to a relative unit: I/100. Therefore, when I = 100%, the following actions occur: I = 100 / 100 = 1.
2) The value of arctg when I = 100% equals $\frac{\pi}{4}$, and as a result, these actions take the form of: $\text{arctg}\frac{I}{100}$.
3) The contribution of the multiplier, which characterizes the inflation index in the context of determining the impact of IF on the socio-economic development of the region when I = 100%, will decrease (if the value of this multiplier is -1). As can be observed, the implementation of the second point is already occurring. As a result, the value of this multiplier equals $\frac{\pi}{4}$, thus we will normalize it as a coefficient, which will equal $\frac{4}{\pi}$.

As a result, we obtain: the inflation index can range from $0 < I < \infty$, with its transformed range from 0 to 1 when determining the level of impact of the Integrated Forms (IF) on the socio-economic development of the region. If *I > 100 %* – the inflation index decreases the overall value of IF_avg; if *I < 100 %* – the inflation index increases the overall value of IF_avg. To fulfill this condition, we carry out the actions described previously. We integrate the multiplier from $\frac{4}{\pi}\text{arctg}\left(\frac{1}{100}\right)$ one form to another $\frac{1}{\frac{4}{\pi}\text{arctg}\left(\frac{1}{100}\right)}$, with the results provided in Table 2.

Table 2: Calculations of the Integrated Inflation Index.

Inflation index (*I*), %	$f(I)=\frac{1}{\frac{4}{\pi}\text{arctg}\left(\frac{1}{100}\right)}$
110	0,942876
99	1,006439
100	1

Based on the conducted observations, we can determine the level of impact of the IF on the socio-economic development of the region, which is based on employment levels, unemployment rates, and the inflation index. Consequently, the level of impact of IF on the socio-economic development of the region is defined by the following expression:

$$IF_{\_avg} = \left(\left(\frac{K_U}{U}\right)\cdot\left(\frac{E`}{K_E}\right)\cdot\left[\frac{1}{\frac{4}{\pi}\text{arctg}\left(\frac{I}{100}\right)}\right]\right)\cdot 100, \quad (2)$$

where:

- IF_avg – the level of impact of the integrated forms on the socio-economic development of the region;
- KU – normative value for the fluctuation of the unemployment percentage in the region – 2%;
- KE – normative value for the fluctuation of the employment percentage in the industry in the region – 7,5%;
- U – unemployment percentage in the region (%);
- I – inflation index in the country (%).
- E' – the level of employed workers in the enterprise relative to the total number of employees in the industry in the region (%).

Thus, a significant increase in the indicator *(IF_avg)* occurs with a marked increase in the percentage of employees in the enterprise compared to the total number of employed in the industry *(E')* and a significant decrease in the unemployment percentage *(U)* and inflation index *(I)*. Therefore, if the values of K_E are less than the normative level, it indicates the potential for development in the industry.

At the same time, if *IF_avg > 1*, it characterizes the presence of the contribution of *IF* to the socio-economic development of the region, and the higher the value, the more significant its role through employment. If *IF_avg* < 1, it indicates the absence of its contribution to the region.

The reliability of using the evaluation of the level of impact of IF on the socio-economic

development of the Odessa region is confirmed by the corresponding calculations of the construction company “Gefest” which started its activities in the Ukrainian market in 1997 in Odessa. Today, “Gefest” is associated with quality, stability, elegant taste, and rapid development [20, 21].

According to the statistical data of the Odessa region and the financial and consolidated reports of the construction company "Gefest", we note that in 2021: The unemployment rate in the Odessa region was 7,2%; the inflation index was 110%; the number of employed workers in the construction company "Gefest" was 620; the level of employed workers in the construction company "Gefest" relative to the total number of employees in the construction industry in the Odessa region was 1,33%.

The calculation of the recommended approach to determining the level of impact of the activities of the construction company "Gefest" on the socio-economic development of the Odessa region for 2021 is as follows:

$$IF_{-\text{avg}} = \left(\left(\frac{2}{7{,}2}\right)\cdot\left(\frac{1{,}33}{7{,}5}\right)\cdot\left[\frac{1}{\frac{4}{\pi}\operatorname{arctg}\left(\frac{110}{100}\right)}\right]\right)\cdot 100 =$$

$$=(0{,}27\cdot 0{,}1775\cdot 0{,}942876)\cdot 100=4{,}75\%$$

In the context of this study, it is determined that in these relations there is a synergistic effect and a positive result (1,26 %) of its contribution to the development of Odessa.

When determining the efficiency of IF functioning, it should be noted that intensive interaction of the PRI is necessary. Thus, intensification of the production of each PRI is essential to counteract unfavorable trends in economic development, characterized by rising inflation, unstable economic and political situation, etc. The intensive activity of the PRI in the formation and development of IF will allow to counteract these negative phenomena and give the development of relations the proper dynamism. At the same time, there are comprehensive opportunities for the PRI 's creativity to develop these relations.

5 CONCLUSIONS

It is important to note that for an accurate and comprehensive assessment of the contribution of integration forms to the socio-economic development of a region, it is necessary to consider not only basic factors such as employment levels, unemployment rates, or the inflation index but also a significantly broader range of indicators. Among these factors, the following stand out:

- Investment volume – analyzing the volume of domestic and foreign investments attracted to the region can indicate the attractiveness of integration forms for businesses and their ability to create new opportunities for economic development;
- Level of innovative activity – assessing the implementation of new technologies, developments, patent activity, and other indicators of innovation in the region can demonstrate how integration forms contribute to modernization and the competitiveness of the regional economy;
- Labor productivity – analyzing the efficiency of labor resource utilization, including productivity at both the enterprise and regional levels, can serve as an important indicator of the impact of integration forms on development;
- Environmental activity – integration forms that take into account environmental aspects (reducing emissions, implementing eco-friendly technologies, waste management) ensure not only economic but also social benefits, improving the quality of life for the population.
- Infrastructure development – integration forms can facilitate the development of transportation, industrial, educational, and other types of infrastructure, which directly influence the socio-economic growth of the region;
- Level of public involvement – the participation of public organizations, residents, and other stakeholders in decision-making processes related to the implementation of integration forms serves as an indicator of their impact on the democratization of regional development processes.

A comprehensive analysis of these factors will provide a more objective and multidimensional assessment of the effectiveness of integration forms. This will contribute to the creation of strategies that not only consider economic impacts but also incorporate social, environmental, and innovative aspects. Such an approach will help ensure the sustainable development of the region, taking into account its specific characteristics, needs, and potential.

ACKNOWLEDGMENTS

We acknowledge support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Open Access Publishing Fund of Anhalt University of Applied Sciences.

REFERENCES

[1] Ganti, "Social Responsibility in Business: Meaning, Types, Examples, and Criticism," Investopedia, March 6, 2024, [Online], Available: https://www.investopedia.com/terms/s/socialresponsibility.asp.

[2] "Integrating corporate social responsibility into business strategy," LinkedIn, [Online], Available: https://www.linkedin.com/pulse/integrating-corporate-social-responsibility-business-strategy-zckre.

[3] P. Robinson, "The Ultimate Guide To Integrating Corporate Social Responsibility," Meltwater Blog, November 27, 2020, [Online], Available: https://www.meltwater.com/en/blog/successfully-integrating-corporate-social-responsibility-in-marketing.

[4] R. N. Sharma, "Integrating Corporate Social Responsibility in Business Strategy," Medium, August 8, 2023, [Online], Available: https://medium.com/@rajendro/integrating-corporate-social-responsibility-in-business-strategy-19d48f6f97eb.

[5] E. D. Rinawiyanti, C. Huang, and S. As-Saber, "The integration of social responsibility into business operation: Case study of Indonesian manufacturing industry," IOP Conf. Ser. Mater. Sci. Eng., vol. 703, pp. 1-8, August 2019, doi: 10.1088/1757-899x/703/1/012016, [Online], Available: https://iopscience.iop.org/article/10.1088/1757-899X/703/1/012016.

[6] T. Sendlhofer, "Decoupling from moral responsibility for CSR: Employees' visionary procrastination at a SME," J. Bus. Ethics, vol. 167, pp. 361-378, 2020, [Online], Available: https://link.springer.com/article/10.1007/s10551-019-04174-z.

[7] Q. H. Vuong et al., "Identifying the moral-practical gaps in corporate social responsibility missions of Vietnamese firms: An event-based analysis of sustainability feasibility," Corp. Soc. Responsib. Environ. Manag., vol. 28, pp. 30-41, 2020, [Online], Available: https://www.researchgate.net/publication/345344588_Identifying_the_moral-practical_gaps_in_corporate_social_responsibility_missions_of_Vietnamese_firms_an_event-based_analysis_of_sustainability_feasibility.

[8] E. Jastrzębska and P. Legutko-Kobus, "Implementation of sustainable development by cities and businesses in Poland: Evolution of the approach," Stud. Ecol. Bioethicae, vol. 20, no. 1, pp. 53-66, 2022, [Online], Available: https://czasopisma.uksw.edu.pl/index.php/seb/article/view/9823.

[9] A. Patuelli, J. O. Carungu, and N. Lattanzi, "Drivers and nuances of sustainable development goals: Transcending corporate social responsibility in family firms," J. Clean. Prod., vol. 373, pp. 1-15, 2022, [Online], Available: https://www.researchgate.net/publication/363141073_Drivers_and_nuances_of_sustainable_development_goals_Transcending_corporate_social_responsibility_in_family_firms.

[10] J. J. Tarí et al., "A taxonomy of quality standard adoption: Its relationship with quality management and performance in tourism organizations in Spain," J. Tour. Serv., vol. 21, pp. 22-37, 2020, doi: 10.29036/jots.v11i21.151.

[11] D. Streimikiene and R. R. Ahmed, "The integration of corporate social responsibility and marketing concepts as a business strategy: Evidence from SEM-based multivariate and Toda-Yamamoto causality models," Oeconomia Copernic., vol. 12, no. 1, pp. 125-157, 2021, [Online], Available: https://www.researchgate.net/publication/350853704_The_integration_of_corporate_social_responsibility_and_marketing_concepts_as_a_business_strategy_Evidence_from_SEM-based_multivariate_and_Toda-Yamamoto_causality_models/.

[12] C. C. Chen et al., "The role of corporate social responsibility and corporate image in times of crisis: The mediating role of customer trust," Int. J. Environ. Res. Public Health, vol. 18, no. 16, 2021, [Online], Available: https://www.researchgate.net/publication/353708002_The_Role_of_Corporate_Social_Responsibility_and_Corporate_Image_in_Times_of_Crisis_The_Mediating_Role_of_Customer_Trust.

[13] E. Stawicka, "Sustainable development in the digital age of entrepreneurship," Sustainability, vol. 13, no. 8, 2021, [Online], Available: https://www.researchgate.net/publication/350902473_Sustainable_Development_in_the_Digital_Age_of_Entrepreneurship.

[14] O. Vasylenko and I. Sencha, "Matematychno-statystychni metody analizu v prykladnykh doslidzhenniakh: navchal′nyy posibnyk," 2012, pp. 1-168, [Online], Available: https://duikt.edu.ua/uploads/l_377_27629033.pdf.

[15] "Correlation analysis," QuestionPro, [Online], Available: https://www.questionpro.com/features/correlation-analysis.html.

[16] "Correlation analysis," Adobe Business, [Online], Available: https://business.adobe.com/blog/basics/correlation-analysis#:~:text=Quick%20definition%.

[17] "Zaynyate naselennya za vydamy ekonomichnoyi diyal′nosti po rehionakh," Derzhavna sluzhba statystyky Ukrayiny, [Online], Available: https://ukrstat.gov.ua/operativ/operativ2013/rp/zn_ed_reg/zn_ed_reg_u/arch_zn_ed_u.htm.

[18] "Indeks inflyatsiyi v Ukrayini," TOV "MinfinMedia", [Online], Available: https://index.minfin.com.ua/ua/economy/index/inflation/.

[19] "Sotsial′no-ekonomichne stanovyshche Odeskoyi oblasti," Holovne upravlinnya statystyky v Odeskoyi oblasti, [Online], Available: https://od.ukrstat.gov.ua/arh/smi/arh_smi_2020.htm.

[20] Construction Company "Gefest," [Online], Available: https://gefest.ua.

[21] G. Otlyvanska et al., "Measuring the economic value of investment activities: A case study of Ukrainian telecommunications companies," Proc. Int. Conf. Appl. Innov. IT, vol. 12, pp. 149-157, 2024. doi: 10.25673.