Advanced Communication Model with the Voice Control and the Increased Security Level

Serhii Kulibaba[1], Svitlana Popereshnyak[2], Yurii Shcheblanin[1], Oleg Kurchenko[1], and Nataliia Mazur[3]

[1] *Taras Shevchenko National University of Kyiv, 60 Volodymyrska str., Kyiv, 01601, Ukraine*
[2] *State University of Telecommunication, 7 Solomyanska str., Kyiv, 03680, Ukraine*
[3] *Borys Grinchenko Kyiv University, 18/2 Bulvarno-Kudriavska str., Kyiv, 04053, Ukraine*

Abstract

Technologies of voice assistants or assistants are actively used in modern solutions. They are implemented on the basis of software that can perform tasks or provide services based on given voice commands by processing and interpreting human speech. Most communication tools are easy to use, but have typical and limited functionality for commercial use, in particular the implementation of security services. The paper proposes a model of a communication tool with a voice assistant and an increased level of security, which uses an advanced encryption method and a face recognition algorithm to provide security services. To build this model, the internal logic of performing functions and placing objects in systems with the functions of voice assistants or assistants was considered. The proposed solution uses a client-server architecture. As an additional direction of commercial use, the possibility of expanding the capabilities of the communication tool through the integration of auxiliary functionality in the form of a currency exchange module and decentralization of the system is being considered.

Keywords

Software, encryption, face recognition, decentralization, currency exchange.

1. Introduction

Modern information technologies provide users with a wide range of services. Means of communication are actively used for their stable functioning. They can be in the form of a service, application, etc. [1–3].

There is a significant number of similar tools on the market, some of them differ in their purpose, supported services and technological solutions. Developers of communication tools are trying to expand their capabilities in the direction of providing commercial services. However, existing implementations in this direction do not allow to meet the needs of users.

Users prefer means of communication with a convenient, functional and understandable interface, available visual effects, speed of information exchange, reliability of stored information, confidentiality, opportunities for commercial use, etc. [4, 5]. Commercial products and development companies try to implement a number of these requirements, but they do not always include a demonstration of the realized capabilities.

The relevance of the work lies in the high interest of users in the use of communication tools with a voice assistant, an inclusive interface, an increased level of security and opportunities for commercial use.

The novelty of the work is the improvement of existing approaches for recognizing named entities of the program text by taking into account contextual information, directions for increasing the level of security of the communication tool by improving the encryption method and the user's face recognition algorithm, as well as expanding the functional capabilities that are limited in most commercial products are proposed.

CPITS-2022: Cybersecurity Providing in Information and Telecommunication Systems, October 13, 2022, Kyiv, Ukraine
EMAIL: kulibseryyy@gmail.com (S. Kulibaba); spopereshnyak@gmail.com (S. Popereshnyak); sheblanin@ukr.net (Y. Shcheblanin); kurol@ukr.net (O. Kurchenko); n.mazur@kubg.edu.ua (N. Mazur)
ORCID: 0000-0002-7316-1214 (S. Kulibaba); 0000-0002-0531-9809 (S. Popereshnyak); 0000-0002-3231-6750 (Y. Shcheblanin); 0000-0002-3507-2392 (O. Kurchenko); 0000-0001-7671-8287 (N. Mazur)

2. Analysis of Publications, the Status of the Issue and the Statement of the Problem

2.1. Analysis of Research and Publications

Recent research in the fields of machine learning and natural language processing has made it possible to create impressive things that seemed impossible before [6]. Recognition of images and voice commands, recommendation systems, self-driving cars, exoplanet searches, prediction of fraudulent bank transactions, military technology are just a small list of what has become possible thanks to the rapid growth of available information and the development of the aforementioned fields of science. One such technology is voice assistants or assistants, which are software that can perform tasks or provide services for the user based on given voice commands, that is, by processing and interpreting human speech [7].

Voice control is an important emerging feature that is changing the way people live. The voice assistant is commonly used in smartphones and laptops. AI-powered voice assistants are operating systems that can recognize the human voice and respond with integrated voices. In work [8], a voice assistant collects audio data using a microphone, converts it into text, and sends it to users.

AI-powered voice assistants such as Alexa and Siri are increasingly replacing search engines as users actively use them to solve a variety of everyday tasks. Technology providers, as well as marketers, are increasingly working to attract customers to use communication tools with voice assistant functions. The work [9] analyzes the possibilities and options of using communication tools with voice assistants.

In [10], a voice control system based on an artificial intelligence (AI) assistant is proposed. A system of AI assistants used by Google Assistant, a representative artificial intelligence service with an open API, and a conditional automatic launch system were developed. The proposed technology is expected to be applied to various control systems based on voice recognition.

In work [11], a means of communication with a voice assistant for mobile phones, based on the Android operating system, is considered. The tool allows voice control. A native interface is provided for a large continuous speech recognition system based on the Kaldi open source speech recognition toolkit.

Thanks to advances in voice recognition, users can easily control any device in a smart home by simply speaking a voice command. Based on this idea, a new group of smart devices called voice assistants has been developed and released. However, voice itself is not secure and can be attacked in many ways. To protect against voice attacks, you can use the voice detection system [12], which consists in the fact that movements when opening the mouth change the size of the ear canal space, which further changes the air pressure in the ear canals. This approach allows the user's air noise pressure data to be compared and matched to voices to verify and identify the source of the voice.

2.2. Analysis of the State of the Issue in the Applied Field

Currently, there are a significant number of means of communication, but not all of them are reliable and convenient. The proposed model of a means of communication with an increased level of security takes into account the shortcomings of existing solutions and provides an opportunity to realize the function of commercial use.

The proposed model of the means of communication combines a number of solutions that are not available from competitors, and this is what guarantees commercial success [13].

2.3. Analysis of Existing Means of Communication

A number of software communication tools were able to gain success and take a confident position in the market among users. Solutions with a high level of success include the following means of communication:

1. Viber: Viber Media company, Inc [14]. Exchange of messages; calls; creating your own bots; convenient to use.
2. WhatsApp: Facebook company [15]. Exchange of messages, calls; convenient to use.
3. Messenger: Facebook company [16]. Exchange of messages; calls; convenient to use.

4. Telegram: Telegram FZ-LLC company [17]. Exchange of messages; calls; creating your own bots; convenient to use.
5. Instagram: Facebook company [18]. Exchange of messages; calls; display information about yourself, if the profile is not closed; information management; convenient to use.

2.4. Formulation of the Problem

Many means of communication are easy to use. But most of them lack the ability to expand functionality. Not everyone can use their own means without the help of others. Thanks to such modifications as a voice assistant, an increased level of security, the presence of an inclusive interface and the provision of commercial services, the percentage of users of these tools may increase significantly.

3. Study Purpose and Objectives

The purpose of the research is to develop a model of an improved means of communication that provides an increased level of security and commercial opportunities. Consider the feasibility of integrating a voice assistant into similar tools - inclusive. Currently, this technology is added to most commercial products. But this technology is absent in means of communication. Therefore, this modification is considered the newest among other software available on the market.

Adding to the means of communication the possibility of tracking the exchange rate and carrying out its exchange and transfer will expand the possibilities of using this tool. There are analogues of means of communication on the market with the possibility of exchanging one's own funds with other users, but the provision of material assistance is carried out by a minority. Not everyone has the opportunity to personally exchange / transfer their own funds to others. Therefore, a modification in the form of a voice assistant in the communication system will be useful.

The integration of the cryptographic data encryption module and the facial recognition function into the communication medium guarantees an increase in the level of security and reliability.

The set goal is achieved by solving the following tasks:

1. Conducting an analysis of the possibilities of the means of communication present on the market.
2. Development of a model of a means of communication.
3. Improvement of the method of encryption and face recognition.
4. Integration of the commercial capabilities module into a communication tool with support for voice assistant functions.

4. Development of a Model an Improved Means of Communication with a Voice Assistant

4.1. Organization of the Process of Interaction with the System

User interaction with the system is carried out using certain visual means. Thanks to these tools, you can work with the system and reveal its potential. Visual tools in such systems are called widgets [19].

In order for any user to be able to interact with the system, it is necessary to expand the functionality of the system: add authorization, registration, create a dialogue with other users, etc. All or most of these functions should be built into the voice assistant. In Fig. 1 is a use case diagram that demonstrates the main functionality of the system.

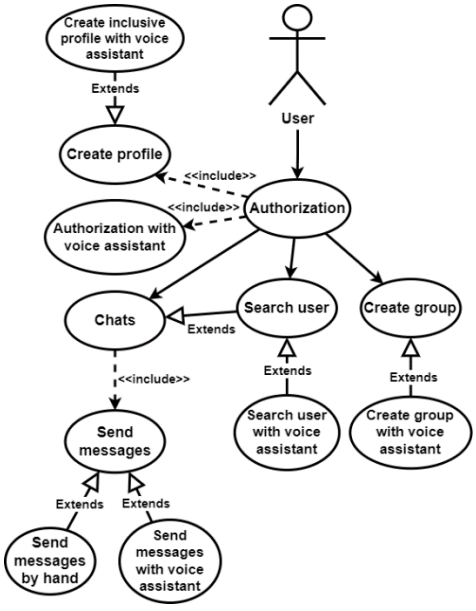

Figure 1: Precedent diagram

Voice assistant. A complete set of algorithms that automatically implements the performance of user functions. This feature will be added and will be available to all users of this system.

4.2. Analysis of the Function of Means of Communication in Relation to Data Saving

To save data and maintain communication with other people through a means of communication, equipment is used, which makes it possible to store and display data in software (Fig. 2). Data is stored through a database [20].

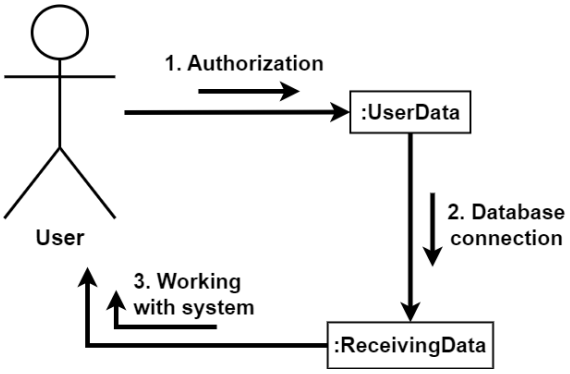

Figure 2: Cooperation diagram of data processing

4.3. Analysis of the Use of Cryptographic Protection in Means of Communication

Cryptographic encryption is the replacement of the data structure with symbols and the creation of a certain key to replace the symbols with the original structure [21]. Encryption is used for three purposes:

1. Confidentiality.
2. Immutability.
3. Confirmation of the source.

Confidentiality. Thanks to cryptographic protection methods, information can be made inaccessible to persons trying to steal this information from the outside.

Immutability. Encrypted information cannot be tampered with during transmission or storage.

Confirmation of the source. Encrypted information has information about the sender.

Data transfer from one user to another occurs as follows:

1. The original text, image or video is encrypted using an encryption algorithm, and the recipient has a special key to decrypt this data.
2. The encrypted message is sent to the recipient.
3. The recipient decrypts the message using a special key.

4.4. Increase the Level of Data Protection

In this work, the Pattern Reverse Subtraction (PRS) encryption method will be used. This method processes the input data by converting the binary value of the current symbol to decimal, where the key is then applied to form the encrypted data.

Encryption will occur as follows:

1. Creating a key.
2. Creating a reverse key.
3. Receiving bytes from input data.
4. Subtraction of the current byte until the condition $0 \leq L$ is met, where L is the length of the key.
5. This number is taken as an index from the reverse. The output number is taken as the index of the value from the key reversal.
6. Adding a number after each operation on the byte, which will reflect the number of differences before obtaining a valid value.

Decryption uses the following formula:

$$\sum_{i=1}^{n} K + L, \qquad (1)$$

where n is the number saved in the encrypted file;
K is symbol index of reverse key;
L is decimal number.

Compression. To achieve a lower amount of equipment physical memory costs, the following approach to data storage will be used (Fig. 3) [22].

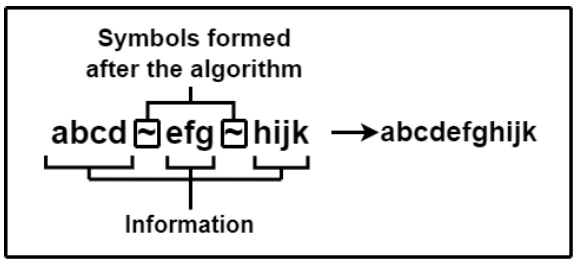

Figure 3: Data compression

Data compression is divided into two types:

1. With losses.
2. Without losses.

With losses. Removing unnecessary bits of data where after decompressing the compressed file you can get the raw data.

Without losses. Compression of information without replacing the symbols of the input data, where after decompression the output data will be obtained.

When studying this encryption method, it was decided to use a type of losses compression for a large amount of data. The study showed that the use of this type of compression did not lead to the detection of errors and incorrectness of the original data.

4.5. Algorithm of Recognition of Characteristic Points on the Face

One of the methods of maintaining confidentiality and preserving the reliability of the program is authentication using characteristic points on the face-face recognition (Fig. 4).

Face recognition is a means of recognizing characteristic points of a person's face and its verification. Characteristic points are determined on the face, and make it possible to recognize the corresponding person [23].

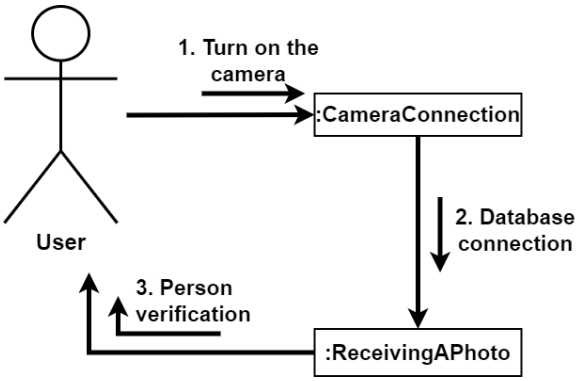

Figure 4: Cooperation diagram of verification person

This technology is needed to increase the level of data security of registered users.

4.6. Performance Criteria

Functional completeness is performing the main functions of effective management and providing a convenient interface for the user [24].

The functional completeness is as follows:

1. Authorization. Inclusive authorization.
2. Registration. Inclusive registration.
3. Creating a chat for interviews with the possibility of using a voice assistant.
4. Creation of group chats for interviews with the possibility of using a voice assistant.
5. Exchange of messages with the possibility of using a voice assistant.
6. Sending photos and videos.
7. Viewing videos and photos using a voice assistant.
8. Sending voice messages.
9. Converting voice messages into text.
10. Changing profile settings.

Portability is a property that indicates how easy it is to transfer / install the system to a certain device. Currently, software development tools have the ability to create an executable program for various operating and mobile systems [25].

Reliability and fault tolerance. For example, the system will have a number of functional capabilities, where in the future it will be possible to make corrections in the processing of data and clients, so reliability and fault tolerance are at a sufficient level [26].

Security. In any case, the data will be transferred over the network between the server devices. Therefore, the proposed PRS data encryption method with E2E (end-to-end) data transmission is used [27].

For authorization in the system, it will be possible to create additional authentication—facial recognition.

Extensions is the possibility of adding/modifying the software product. The following diagram (Fig. 5) shows the structure of the future software product.

Productivity. Thanks to well-written software, you can achieve a high level of productivity. To achieve high performance, you need to have equipment that will ensure fast processing of input and output data, in particular voice requests from users [28].

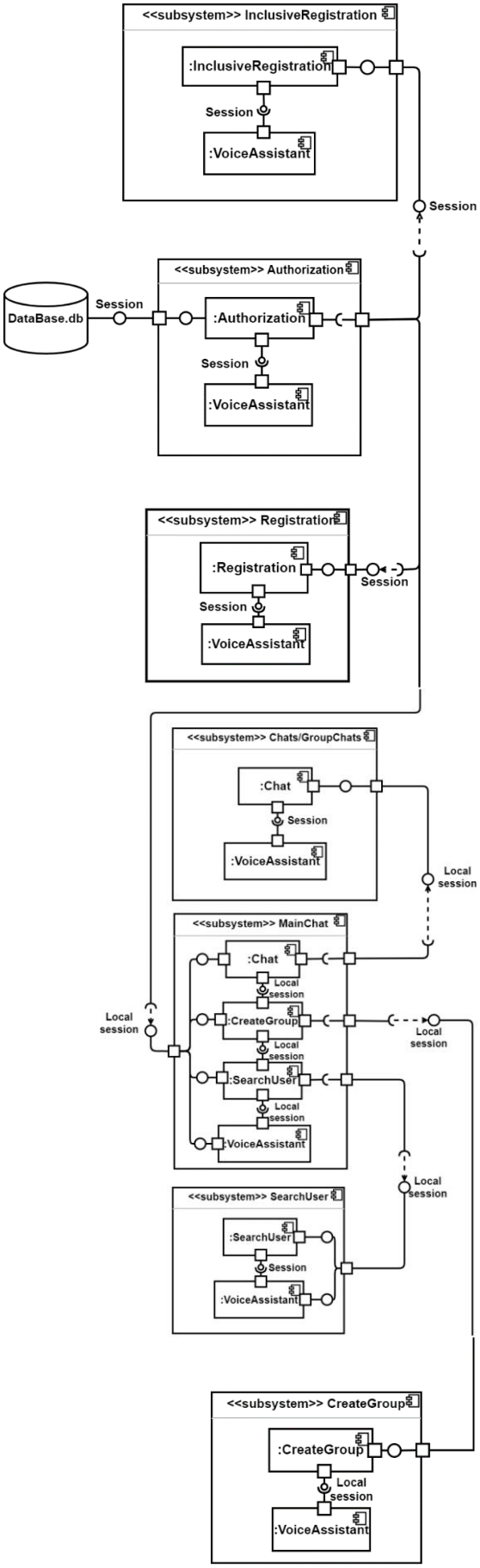

Figure 5: Components diagram

5. Improving the Means of Communication by Adding Commercial Opportunities

5.1. Cryptocurrency and Other Currencies Exchange Module

The expansion of the audience of users is planned due to the addition of opportunities for conducting financial transactions. This extension will make the software product more successful. Due to the minimal increase in the commission for the transfer of currencies, there will be an increased profit for the corresponding period of time. Finance departments are interested in cooperating with such tools, because their number of users will be able to expand and increase the level of profit. Cooperation will take place through the API (*Application Programming Interface*) of financial departments for the interaction of their services with their own developed software and system users [29].

An account will be created for the financial support of the relevant users, where funds will be received for the percentage of the transfer of funds from other users and sent to the relevant rehabilitation institutions.

5.2. Decentralization

Decentralization is the process of distributing people, rules, performing calculations, data stored from a single location or central governing body [30]. The internal structure of the decentralized system consists of blocks and chains of a single profile, which are connected to each other and have the name blockchain. The meaning of the structure in the system is as follows: blocks are general information, and a chain is a connection between blocks. The content of the information may depend on the conditions specified in the system software: time of sending funds, message, comment, etc.

When applying the blockchain technology, it is necessary to define and distribute rules between users and computing devices, achieving consensus, in particular, connecting other users who must verify the authenticity of data due to stored copies [31, 32].

This improvement will be useful for people who do not want to lose their personal information

and use the application without entering information about themselves [33].

5.3. Analysis of the Success of Improvements

The implementation of *points 5.1 and 5.2* will make it possible to raise the level of commercial attractiveness of the improved means of communication. Currency exchange in a closed system can be interesting for a large number of users.

The following formula can be used to obtain the probability of success of improving the means of communication:

$$C = A \times B \times D, \quad (2)$$

where C is the probability that the changes will be successful;
A is dissatisfaction with the existing situation;
B is a clear formulation of the goal of the changes;
D is specific first steps to achieve goals.

6. Conclusion

The paper examines the functioning scheme of means of communication called "messengers." To build a model of an improved means of communication with a voice assistant, the internal logic of performing functions and placing objects in similar means was considered. The work proposed improved principles of increasing the level of security of products of this type by improving the encryption method and the face recognition algorithm, as well as expanding the functionality that is missing in most commercial products. The tool is built using a client-server architecture. This security model and method provide an opportunity to protect data at a sufficient level. To provide convenience to users, a decision was considered to add a voice assistant to the system. This modification will be useful for all users of the system. The work proposed the possibility of expanding the system with the help of auxiliary functionality in the form of a currency exchange module, which is intended for a certain category of users interested in this system.

The system is built on the basis of a micro service approach, which ensures scalability and low resource consumption when adding new functionality or replacing existing implementations. The created voice assistant for integrated development environments allows to increase the productivity of developers, and also provides an opportunity to support the standards of conventions of inclusive interfaces.

Based on the results of the research, the logic of the communication tools was determined. Thanks to this, you can design a software product that will be different from similar products on the market.

So, as the system operation scheme was considered in detail, these systems can be expanded not only externally, but also internally. Several examples have been given of how the functionality of the product can be extended.

7. References

[1] Y. Nakamura, et al., Design and Evaluation of In-Situ Resource Provisioning Method for Regional IoT Services, 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), 2018, pp. 1–2. doi: 10.1109/IWQoS.2018.8624127.

[2] O. Iosifova, et al., Techniques Comparison for Natural Language Processing, in Proceedings of the Modern Machine Learning Technologies and Data Science Workshop, vol. 2631, 2020, pp. 57–67.

[3] I. Iosifov, O. Iosifova, V. Sokolov, Sentence Segmentation from Unformatted Text Using Language Modeling and Sequence Labeling Approaches, in Proceedings of the 2020 IEEE International Scientific and Practical Conference Problems of Infocommunications. Science and Technology, 2020, pp. 335–337. doi: 10.1109/picst51311.2020.9468084.

[4] Z. Shi, Y. Liang, X. Wang, Analysis of Demand Side Energy IoT Communication Channel Requirements of Integrated Stations, Equipment, and Users, in 7th Asia Conference on Power and Electrical Engineering, 2022, pp. 793–797. doi: 10.1109/acpee53904.2022.9784018.

[5] V. Buriachok, V. Sokolov, P. Skladannyi, Security Rating Metrics for Distributed Wireless Systems, in Workshop of the 8th International Conference on "Mathematics. Information Technologies. Education:" Modern Machine Learning Technologies and Data Science (MoMLeT and DS), vol. 2386, 2019, pp. 222–233.

[6] A. S. Subramanian, et al., Far-Field Location Guided Target Speech Extraction Using End-to-End Speech Recognition Objectives,

ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 7299–7303. doi: 10.1109/ICASSP40776.2020.9053692.

[7] C. Fan, et al., Gated Recurrent Fusion With Joint Training Framework for Robust End-to-End Speech Recognition, in: IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, 2021, pp. 198–209. doi: 10.1109/TASLP.2020.3039600.

[8] S. Subhash, et al., Artificial Intelligence-based Voice Assistant. 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4). 2020. pp. 593–596. doi: 10.1109/WorldS450073.2020.9210344.

[9] S. Malodia, et al., Why Do People Use Artificial Intelligence (AI)-Enabled Voice Assistants, in IEEE Transactions on Engineering Management, 2021. doi: 10.1109/TEM.2021.3117884.

[10] T.-K. Kim, Short Research on Voice Control System Based on Artificial Intelligence Assistant, 2020 International Conference on Electronics, Information, and Communication (ICEIC), 2020, pp. 1–2. doi: 10.1109/ICEIC49074.2020.9051160.

[11] B. Popović, et al., Voice assistant application for the Serbian language, 2015 23rd Telecommunications Forum Telfor (TELFOR), 2015, pp. 858–861. doi: 10.1109/TELFOR.2015.7377600.

[12] J. Shang, J. Wu, Voice Liveness Detection for Voice Assistants using Ear Canal Pressure, 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), 2020, pp. 693–701. doi: 10.1109/MASS50613.2020.00089.

[13] A. Asaul, et al., The Latest Information Systems in the Enterprise Management and Trends in their Development, 2019 9th International Conference on Advanced Computer Information Technologies (ACIT), 2019, pp. 409–412. doi: 10.1109/ACITT.2019.8779874.

[14] A. Vasilaras, et al., Android Device Incident Response: Viber Analysis, 2022 IEEE International Conference on Cyber Security and Resilience (CSR), 2022, pp. 138–142. doi: 10.1109/CSR54599.2022.9850300.

[15] R. Wahaz, et al., Is WhatsApp Plus Malicious? A Review Using Static Analysis, 2021 6th International Workshop on Big Data and Information Security (IWBIS), 2021, pp. 91–96. doi: 10.1109/IWBIS53353.2021.9631860.

[16] Y. Mei, et al., Turbine: Facebook's Service Management Platform for Stream Processing, 2020 IEEE 36th International Conference on Data Engineering (ICDE), 2020, pp. 1591–1602. doi: 10.1109/ICDE48307.2020.00141.

[17] C. Huda, F. A. Bachtiar, A. A. Supianto, Reporting Sleepy Driver into Channel Telegram via Telegram Bot, 2019 International Conference on Sustainable Information Engineering and Technology (SIET), 2019, pp. 251–256. doi: 10.1109/SIET48054.2019.8986000.

[18] W. Uriawan, et al., Pearson Correlation Method and Web Scraping for Analysis of Islamic Content on Instagram Videos, 2020 6th International Conference on Wireless and Telematics (ICWT), 2020, pp. 1–6. doi: 10.1109/ICWT50448.2020.9243626.

[19] E. Zhang, S. Peng, Y. Zhai, Design and Application Development of the Camps Navigation System Based on ArcGIS Runtime SDK for Android: Taking the Yunnan Normal University as an Example, 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2019, pp. 1262–1266. doi: 10.1109/IAEAC47372.2019.8997730.

[20] P. Seda, et al., Performance testing of NoSQL and RDBMS for storing big data in e-applications, 2018 3rd International Conference on Intelligent Green Building and Smart Grid (IGBSG), 2018, pp. 1–4. doi: 10.1109/IGBSG.2018.8393559.

[21] S. Kulibaba, O. Kurchenko, Cryptographic Method of Pattern Reverse Multiplication Data Encryption, Cyber Security: Education, Science, Technology, vol. 3, iss. 15, 2022, pp. 216–223.

[22] S. Yamagiwa, R. Morita, K. Marumo, Bank Select Method for Reducing Symbol Search Operations on Stream-Based Lossless Data Compression, 2019 Data Compression Conference (DCC), 2019, pp. 611. doi: 10.1109/DCC.2019.00123.

[23] R. He, et al., Adversarial Cross-Spectral Face Completion for NIR-VIS Face Recognition, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 5, 2020, pp. 1025–1037. doi: 10.1109/TPAMI.2019.2961900.

[24] A. Kolodenkova, E. Khalikova, S. Vereshchagina, Data Fusion and Industrial Equipment Diagnostics Based on Information Technology, 2019 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), 2019, pp. 1–5. doi: 10.1109/FarEastCon.2019.8934322.

[25] T. Deakin, et al., Performance Portability across Diverse Computer Architectures, 2019 IEEE/ACM International Workshop on Performance, Portability and Productivity in HPC (P3HPC), 2019, pp. 1–13. doi: 10.1109/P3HPC49587.2019.00006.

[26] S. Talwani, I. Chana, Fault tolerance techniques for scientific applications in cloud, 2017 2nd International Conference on Telecommunication and Networks (TEL-NET), 2017, pp. 1–5. doi: 10.1109/TEL-NET.2017.8343578.

[27] Y. Kai, H. Qiang, M. Yixuan, Construction of Network Security Perception System Using Elman Neural Network, 2021 2nd International Conference on Computer Communication and Network Security (CCNS), 2021, pp. 187–190. doi: 10.1109/CCNS53852.2021.00042.

[28] P. Y. Tilak, et al., A platform for enhancing application developer productivity using microservices and micro-frontends, 2020 IEEE-HYDCON, 2020, pp. 1–4. doi: 10.1109/HYDCON48903.2020.9242913.

[29] IEEE Standard for Learning Technology — ECMAScript Application—Programming Interface for Content to Runtime Services Communication—Redline, in IEEE Std. 1484.11.2-2020, 2021, pp. 1–60.

[30] E. Işık, M. Birim, E. Karaarslan, Chainex Decentralized Application Development & Test Workbench, 2021 15th Turkish National Software Engineering Symposium (UYMS), 2021, pp. 1–4. doi: 10.1109/UYMS54260.2021.9659637.

[31] J. Jayabalan, et al., A Study on Distributed Consensus Protocols and Algorithms: The Backbone of Blockchain Networks, 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1–10. doi: 10.1109/ICCCI50826.2021.9402318.

[32] T. Salman, R. Jain, L. Gupta, A Reputation Management Framework for Knowledge-Based and Probabilistic Blockchains, 2019 IEEE International Conference on Blockchain, 2019, pp. 520–527. doi: 10.1109/Blockchain.2019.00078.

[33] Y. Shu, Y. J. Gu, J. Chen, Dynamic Authentication with Sensory Information for the Access Control Systems, in IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 2, 2014, pp. 427–436. doi: 10.1109/TPDS.2013.153.