UDC 004.934.2 **doi: 10.20998/2522-9052.2020.3.02**

Vasyl Trysnyuk[1], Yevhen Nagornyi[1], Kirill Smetanin[2], Igor Humeniuk[2], Tetyana Uvarova[3]

[1] Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv, Ukraine
[2] Korolov Zhytomyr Military Institute, Zhytomyr, Ukraine
[3] National University of Defense of Ukraine named after Ivan Chernyakhovsky, Kyiv, Ukraine

A METHOD FOR USER AUTHENTICATING TO CRITICAL INFRASTRUCTURE OBJECTS BASED ON VOICE MESSAGE IDENTIFICATION

Abstract. The **subject** of the research is the protection of critical infrastructure objects from unauthorized access based on voice message identification. The **purpose of the research** is to consider the main methods and algorithms for identifying an individual by voice, as well as the development of a software implementation of the restricted zone access control system, which allows improving the security of the standard password system by introducing voice authorization technologies and improving the algorithm for selecting voice fragments in audio files. Therefore, the purpose determines **following tasks**: analysis of the advantages and disadvantages of existing methods of information protection to infrastructure objects; determination of the optimal method of identification by voice message; a description of the main features of the implementation and use of Mel-frequency kastelnik coefficients (MFKC). Are used such **methods**: system approach, system analysis, theory of algorithms. The following **results** were obtained: an algorithm was developed for storing users ' voice data for further use in authorization, the voice authorization subsystem; the corresponding software of the computer system has been developed. **Conclusions:** based on the example of applying the biometric method of user authentication to critical infrastructure objects based on voice message identification, this method of user authentication should be used as an additional method of multi-factor authentication by entering an additional password or another method of biometric user authentication.

Keywords: biometric identification; identification by voice; authentication by voice message; information protection; unauthorized access; cyber security.

Introduction

Problem statement. The current stage of technological development of society is closely related to the active use and improvement of information technology (IT), the creation of information space [1]. Under such conditions, ensuring information security and information protection are of paramount importance in all spheres of human life.

With the beginning of the armed aggression of the Russian Federation on the territory of Ukraine, which requires more decisive action to protect this information.

Existing methods of information protection are divided into hardware, software and mixed (ie a combination of the two previous methods). Given the damage that can be caused by unauthorized access to information with limited access, any distortion or destruction, it is important to protect general and special information systems (military, law enforcement, etc.) by entering an access password. Password identification is the simplest in both implementation and use, but it has a number of disadvantages that significantly reduce the level of information security.

Today, biometric technologies for personal identification are becoming increasingly important. Biometrics is a very effective device for face recognition, which on the basis of biometric characteristics makes it possible to identify a person [2].

The latest biometric technologies and systems recognize a person based on the characteristics that are given to a person at birth, ie anatomical features such as fingerprints, facial images, palms, iris, voice or behavioral traits (signature, gait, keyboard handwriting). The main advantages of this method include the fact that personal characteristics are individual to each person, they can not be lost or forgotten, as a password, as well as biometric identifiers are more difficult to forge. Since these features are directly related to the physiology of the user, it can be argued that biometric recognition has a high level of reliability and gives access to systems and facilities only to those users who have the appropriate authority to do so [2,3].

Of course, like any complex system, biometric identification has certain disadvantages, however, it is the most reliable, especially in the case of identification by several biometric identifiers.

Therefore, the purpose of this work is to create a method of biometric authentication of users to critical infrastructure based on voice signals and its software implementation is an urgent scientific task [4].

Recent events on the world stage are accompanied by the process of redistribution of zones of influence in the technological, economic and cyber spaces, the development of information technology (IT), which give rise to new ways of obtaining information. In this regard, ensuring the protection of classified information is an urgent issue and requires states, regardless of their development, to constantly strengthen national security, as well as the ability to counter threats and minimize the risks of real leakage of important information and data.

Analysis of recent research and publications. To date, a number of identification methods based on human biometric characteristics have been developed and implemented.

To date, there are several approaches to the identification of a person by voice, which are based on the analysis of the structure of the voice signal. The

© Trysnyuk V., NagornyiYe., Smetanin K., Humeniuk I., Uvarova T., 2020

procedure for processing the voice signal is to use short-term analysis, ie the signal is divided into fragments (frames) of a fixed size. Then to each window algorithms of selection of signs are applied. Most popular identification systems use chalk-frequency keptral coefficients (MFCC) or linear prediction coefficients (LPCS) as feature vectors [5]. These methods are based on the selection of vectors of signs of the voice signal, taking into account the peculiarities of sound perception by the human ear. Another method of analyzing voice fragments is the analysis of formant frequencies. The analysis of formant frequencies is one of the oldest methods of identification of a person by voice, and the identification of a person in the formant approach is most often performed on vowel sounds, in which formants can be effectively distinguished [2, 4, 5].

Thus, as can be seen from the analysis of the sources for the implementation of software authentication of users based on human voice biometrics, a sufficient number of methods, but to ensure timely detection of unauthorized access to critical infrastructure.

Setting objectives. To develop a computer control system for critical infrastructure using biometric voice authentication, which will improve standard password systems and solve the problem of confidentiality of user credentials by introducing voice authorization mechanisms.

Presentation of the main material of the research

At the moment of development of information technologies, the result of identification of the person by voice completely depends on input data, mathematical algorithms and computing power. Input data means a sample of a person's voice obtained by recording from a microphone. The quality of such a sample depends on the type of input device (eg professional microphone or mobile phone) and the environment (loud street or quiet room).

Mathematical algorithms are used to compare the obtained voice sample with the samples in the database. Computing power is understood as the speed and quality of processing of biometric features of the user, which depends on the hardware features of the system.

To date, there are several approaches to the identification of a person by voice, which are based on the analysis of the structure of the voice signal. The procedure for processing the voice signal is to use short-term analysis, ie the signal is divided into fragments (frames) of a fixed size. Then to each window algorithms of selection of signs are applied. Most popular identification systems use chalk-frequency keptral coefficients (MFCC) or linear prediction coefficients (LPCS) as feature vectors [5]. These methods are based on the selection of vectors of signs of the voice signal, taking into account the peculiarities of sound perception by the human ear. Another method of analyzing voice fragments is the analysis of formant frequencies. The analysis of formant frequencies is one of the oldest methods of identification of a person by voice, and the identification of a person in the formant approach is most often performed on vowel sounds, in which formants can be effectively distinguished [2, 4, 5]. The method of developing biometric voice authentication at critical infrastructure facilities can be divided into three stages:

- normalization of the input voice signal;
- selection of characteristic features of the voice;
- comparing the received voice message with the message of the reference signal;

An important stage of voice authentication is the preliminary preparation of the input voice signal (Fig. 1) to highlight the characteristics. One of the main methods of normalizing the input voice signal is to remove fragments from the audio file that do not contain voice (fragments of silence or noise).

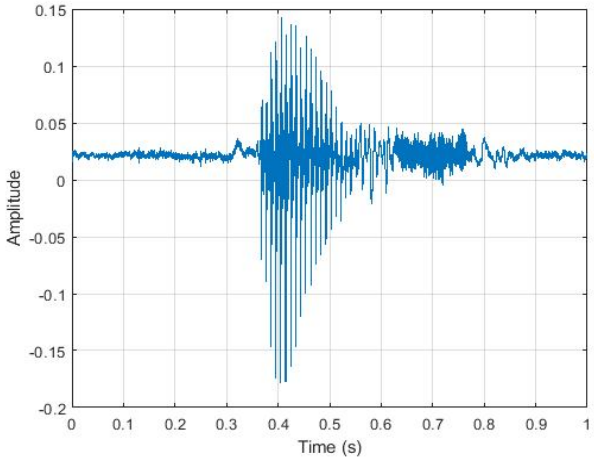

Fig. 1. Audio input signal

To do this, the input audio signal passes through the Voice Activity Detector (VAD), which allows you to select fragments of voice and thus reduce the load on the computer system and increase its speed by reducing the number of unnecessary calculations on areas of audio file that do not contain useful information about the physiological features of the vocal apparatus of the person. An algorithm based on statistical evaluation of the input audio signal was used to select voice fragments. Usually, all audio files recorded with a microphone start with a fragment that does not contain a voice (silence). This is due to the fact that in the process of authorization or registration in the system, the user responds to the beginning of the recording of voice fragments with some delay. This initial fragment of the audio file, lasting 100-200 ms, is the input parameter of the algorithm and contains information about the external environment in which the voice samples are recorded. According to this approach, the values of the audio signal amplitude are considered as random variables that depend on a large number of factors, each of which makes a small contribution, and according to the central limit theorem such random variables (in this case signal amplitude) have a distribution close to normal with parameters. According to the rule of "three sigma", almost all values of a normally distributed random variable lie in the interval $x-3\cdot\delta; x+3\cdot\delta$.

In the next step, all signal amplitude values for which this rule is not met are considered as voice, and

all others are considered as fragments of silence or noise. This approach has its drawbacks due to the correct selection of voice fragments in audio files. According to this approach, fragments of the voice begin where the amplitude values are outside the interval $x-3\cdot\delta; x+3\cdot\delta$ [6]. However, in an experimental research, it was found that audio files often have areas with very small deviations, or their number is small (5-10 amplitude values, which is equivalent to a duration of approximately 0.000625 s at a sampling rate of 8000 Hz). Such parts of the audio file cannot be considered as a fragment of the voice, because the average duration of the letters in a word is tens of milliseconds, words - hundreds of milliseconds, or even a few seconds. However, in an experimental research, it was found that audio files often have areas with very small deviations, or their number is small (5-10 amplitude values, which is equivalent to a duration of approximately 0.000625 s at a sampling rate of 8000 Hz). Such parts of the audio file cannot be considered as a fragment of the voice, because the average duration of the letters in a word is tens of milliseconds, words - hundreds of milliseconds, or even a few seconds. For each frame, a temporary array of boolean values is created, which contains only the value "true" or "1" if the "three sigma" rules are met, and "false" or "0" if such a condition is not met. The next step is to calculate the probability of occurrence of the element with the value "true" P1 and the probability of occurrence of the value "false" P0. Probabilities are calculated by finding the ratio of the number of occurrence of a value to the total number of values in the array (the length of the array, or the number of amplitude values in the fragment). If the value of P1 is less than some threshold value, it is considered that this fragment contains a voice, and if not, then silence. The value of y was determined by experimental research of the results of the algorithm for the selection of fragments of voice in audio files and the results of the experiments were taken to be equal α to 0.65. The parameter α can be interpreted as follows: if 65% of the values of the amplitude of the audio signal in the fragment is outside the interval $x-3\cdot\delta; x+3\cdot\delta$, the system decides that the fragment contains a voice, and otherwise - the fragment contains pauses (silence). After signal normalization, it is necessary to identify the features that characterize the features of the voice apparatus of a particular user. In the field of digital signal processing (DSP) and, in particular, speech recognition and voice identification, the most active applications are the so-called Mel-frequency cepstral coefficients (MFCC). The main idea of the MFCC method is to approximate as much as possible the information coming to the input of the system to the information coming to the auditory analyzer of the human brain. The voice signal is initially an array of amplitude values obtained by sampling the output analog signal with a certain frequency Fs using an analog-to-digital converter (ADC) of the sound card. The next step is to divide the input signal into frames, usually lasting 25-30 ms [4]. The frames overlap each other by 25-70%. Frame overlaps are used to compensate for information loss at the beginning and end of each frame that occurs as a result of applying a window function in the next step of the algorithm. In most digital processing tasks, it is not possible to examine the signal at an infinite interval. Limiting the analysis interval is equivalent to the product of the output signal on the window function [4, 5, 7]. The window function is used in order to avoid unnatural breaks in the voice fragments of the audio file and, accordingly, distortions in the spectrum of the audio signal [6]. Multiplying the output signal by the value of the window function allows you to reduce the amplitude value at both ends of the current frame and thus prevent a sharp change in the values at the endpoints. The point is that the fast Fourier transform algorithm [6] assumes that the signal is continuous and periodic, and in this case the signal is divided into frames of fixed length and a window function is used to avoid distortions in the spectrum. Hemming's window is used as a window function, as it is most often used in the tasks of speech recognition and identification of a person by voice [3, 5]:

$$w_n = 0,54 - 0,46 \cdot \cos\left(2 \cdot \pi \frac{n}{N-1}\right), n = 0,...N-1, \quad (1)$$

where N is the length of the window.

In Fig. 2 shows the input signal and its spectrogram. The horizontal axis of the spectrogram shows the time in seconds, the vertical axis - the frequency, and the color of each point of the image is determined by the value of the amplitude at a certain frequency at a particular time t. Each person when pronouncing sounds, letters is characterized by its own set (combination) of frequencies and due to this the human ear is able to distinguish one sound from another and, in particular, the voices of different people. In order to select this set of frequencies for each frame of the input audio signal, chalk frequency analysis is used. According to this method, the obtained representation of the signal in the frequency domain is divided into bands using a jar (comb) of triangular filters. The filter limits are calculated on the Mel scale, which is the result of studies of the ability of the human ear to the perception of sounds at different frequencies [4, 5].

Conversion to chalk scale is carried out by the formula:

$$B(f) = 1127 \cdot \ln\left(1 + \frac{f}{700}\right), \quad (2)$$

It is believed that the information carried by the low-frequency components of the voice signal is more important than the high-frequency components and therefore the chalk scale is linear up to 1 kHz and logarithmic above 1 kHz, ie at low frequencies, filters are applied linearly, while at high frequencies - logarithmically. These filters are unevenly located on the frequency axis, so such filters have more in the spectrum with low frequencies (up to 1 kHz) and less in the high frequency range (over 1 kHz) [2, 6, 8]. The filters are applied to the squares of the modules of the Fourier transform coefficients, and the obtained values are logarithmic:

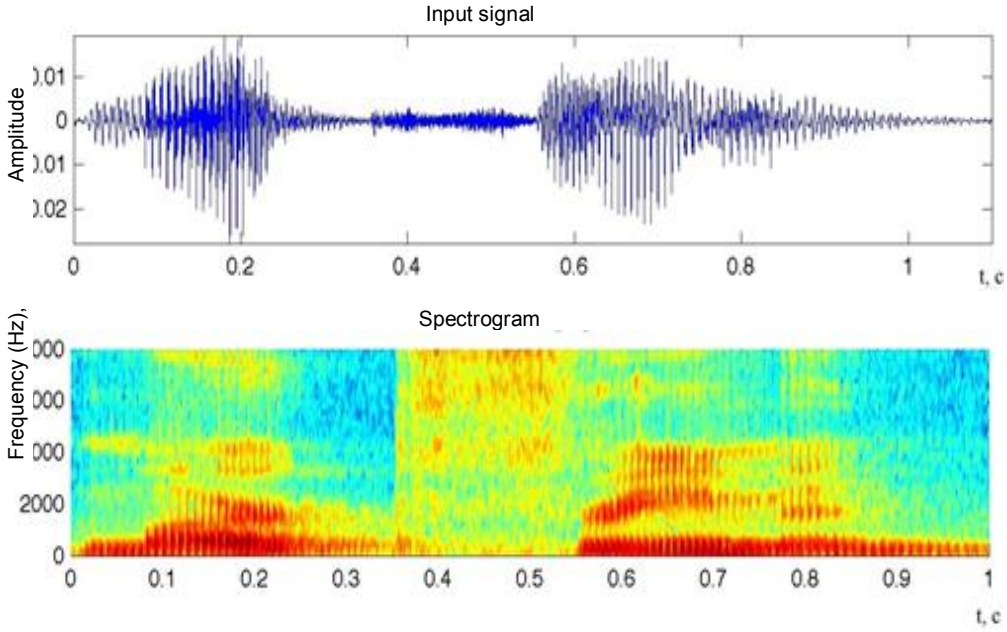

Fig. 2. Spectrogram of the input signal

$$e_m = ln\left(\sum_{k=0}^{N} |X_k|^2 \cdot H_{m,k}\right), m = 0,..., N_{FB-1}, \quad (3)$$

where N_{FB} is the number of filters (most often about 24 filters are used [2]), $H_{m,k}$ - weights of the obtained filters.

The approach partially eliminates the noise components in the frequency domain, as the most important frequencies of the human voice are in the range from 70 Hz to 3400 Hz. The last stage of the process of selection of voice features is the use of discrete cosine transformation (DCT, Discrete Cosine Transform), which will result in a set of chalk-frequency keptral coefficients (MFCC), which will be elements of the vectors of facial features:

$$C_i = \sum_{m=0}^{N_{FB}-1} e_m \cdot cos\left(\frac{\pi \cdot i(m+0,5)}{N_{FB}}\right), \quad i = 0,..., N_{MFCC},$$

where e_m - logarithmic values of Fourier transform coefficients, N_{MFCC} - number of coefficients (size of feature vectors). Keptral coefficients can also be obtained using the IFFT algorithm (inverse fast Fourier transform), but in this case the DCT algorithm is used, which is more efficient because it does not use work with complex numbers [7, 8, 9]. As a result, for each fragment of the original voice signal we obtain a finite set of chalk frequency coefficients of $C_i = (C_1; C_2; ...C_N)$ the kepter, which contains N elements and is a vector of characteristic features of the voice of a particular user.

Voice recognition differs from other systems in that in this case the subject of recognition is a process, not a static image as in the case of fingerprint recognition or iris.

Therefore, most often the voice sample is not represented as a single feature vector, but as a sequence of feature vectors, each of which describes the characteristics of a small area of the voice signal [5, 7, 9]. Therefore, most often the voice sample is not represented as a single feature vector, but as a sequence of feature vectors, each of which describes the characteristics of a small area of the voice signal [5, 7, 9]. The sequence of vectors obtained after the signal processing step is used to build a model of a person's voice. The main parameter used to identify the user is a measure of the similarity of the two sound fragments (input sample and sample in the database) [9, 10, 11]. Fig. 3 shows a block diagram of user authentication to critical infrastructure [12].

In authorization mode, the user tries to log in by presenting the ID as a voice message. The system analyzes this sample, compares it with the reference sample presented from the voice message database and tries to identify the person by voice. If the person is identified, the system decides to grant access. [13-14]

Conclusions and prospects for further research

The article presents the results of solving a topical scientific problem, which was to develop a method of user authentication to critical infrastructure based on the identification of a person by voice message and its software implementation. During the analysis of technological development of IT it is established that one of the potentially possible methods of reliable protection of information and OID from NSD is biometric identification and / or authentication of users, in particular on the basis of voice message of the user. The peculiarity of this method is the possibility of its implementation in mobile applications, which is especially relevant in today's informatization. One of the most effective ways to increase the security of the system is the use of multifactor authentication mechanisms, which is based on the simultaneous use of several authentication factors (knowledge of the secret, possession of the object, physical characteristics), which significantly increases the security of the system.

Therefore, this user authentication system should be used as an additional method of multifactor authentication by entering an additional password or another method of biometric user authentication.

Prospects for further research are to improve the proposed method by integrating into the system of biometric authentication by voice another method of biometric authentication (eg, facial geometry).

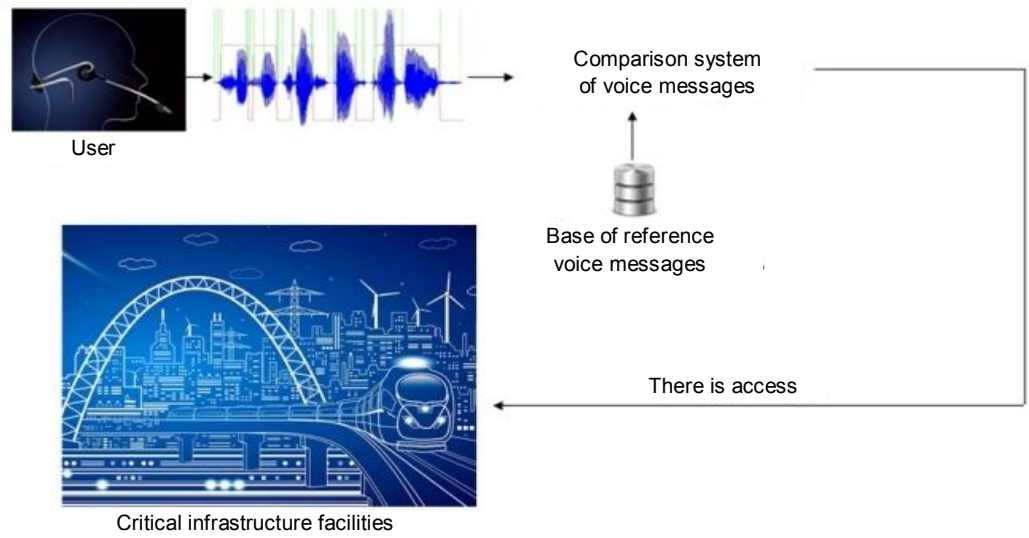

Fig. 3. Block diagram of user authentication to critical infrastructure

REFERENCES

1. (2015), *The concept of creating a national system of identification of citizens of Ukraine, foreigners and stateless persons*, Order of the Cabinet of Ministers of Ukraine № 1428-2015-r dated 23.12.2015, available at: https://zakon.rada.gov.ua/laws/main/1428-2015-%D1%80.
2. Bugaenko, H. and Gorbenko, I. (2012), "Analysis of three biometric methods of identity authentication", *Applied Radio Electronics*, Vol. 11, No. 2, pp. 262-266.
3. Tsarev, R. and Lemekha, T. (2016), *Biometric technologies*, textbook, ONAZ, Odessa, 140 p.
4. Moroz, A. (2011), "Biometric technologies of human identification. System overview", *Mathematical machines and systems*, No. 1, pp. 39-45.
5. (2020), *Chalk-frequency keptral coefficients*, available at: http://poisk-ru.ru/s62453t2.html.
6. Bidyuk, P. and Bondarchuk, V. (2009), "Modern methods of biometric identification", *Legal, regulatory and metrological support of information security in Ukraine*, No. 1 (18), pp. 137-146.
7. Lalit R., Bahl, Peter F., Brown, Peter V., de Souza and Robert L., Mercer (1987), "Speech Recognition With Continuous Parameter Hidden Markov Models", *Proceedings of the International Conference on Acoustics. Speech and Signal Processing*, Vol. 2, Issues 3-4, September–December 1987, IEEE, New York, pp. 219-234.
8. Viola, P. and Jones, M. (2001), "Rapid object detection using a boosted cascade of simple features", *IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 511-518.
9. Smetanin, K., Humeniuk I. and Nekrylov A. (2020), "Protection from unauthorized access to information and telecommunication systems through biometric user identification", *Proceedings of the International Scientific and Technical Conference "PERSPECTIVES OF TELECOM*, available at: http://conferenc.its.kpi.ua/proc/article/view/200915.
10. Campbell, J.P. (1997), "Speaker Recognition: A Tutorial", *Proceedings of the IEEE*, Vol. 85, No. 9, pp. 1437-1462.
11. Ing-Jr, Ding, Chih-Ta, Yen and Yen-Ming, Hsu (2013), "Development so Machine Learning Schemes for 135 Dynamic Time-Wrapping-Based Speech Recognition", *Mathematical Problems in Engineering*, Vol. 2, pp. 35-42.
12. Voronov A.A. (1983), *Methods of automatic speech recognition*, Vol. 1, MIR, Moscow, 328 p.
13. Trysnyuk, V., Demydenko, O., Smetanin, K. and Zozulia, A. (2020), "Improvement of the complex evaluation method of vital activity risks", *European Association of Geoscientists & Engineers, Geoinformatics: Theoretical and Applied Aspects 2020*, May 2020, Vol. 2020, pp. 1-5, DOI: https://doi.org/10.3997/2214-4609.2020geo071.
14. Smetanin, K., Litvinenko, A. and Kostenko, D. (2019), "Biometric authentication of users based on voice signals in cyberspace", *Problems of theory and practice of information confrontation in the conditions of hybrid wars, scientific-practical. Conf.*, Oct. 24–25. 2019, ZhVI, Zhytomyr, pp. 205-207.

Received (Надійшла) 11.05.2020
Accepted for publication (Прийнята до друку) 01.07.2020

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Триснюк Василь Миколайович – доктор технічних наук, старший науковий співробітник, завідувач відділу досліджень навколишнього середовища, Інститут телекомунікацій і глобального Інформаційного простору НАН України, Київ, Україна;

Vasyl Trysnyuk – Doctor of Technical Sciences, Senior Researcher, Head of the environmental research department, Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv, Ukraine;
e-mail: trysnyuk@ukr.net; ORCID ID: https://orcid.org/0000-0001-9920-4879.

Нагорний Євген Ігорович – аспірант, Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна;
Yevhen Nagornyi – postgraduate student, Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv, Ukraine;
e-mail: rhbz777@gmail.com, ORCID ID: https://orcid.org/0000-0001-5902-9295.

Сметанін Кирило Володимирович – кандидат технічних наук, викладач кафедри захисту інформації та кібербезпеки, Житомирський військовий інститут імені С. П. Корольова, Житомир, Україна;
Kirill Smetanin – PhD (Technical), lecturer at the Department of Information Security and Cybersecurity, Korolov Zhytomyr Military Institute, Zhytomyr, Ukraine;
e-mail: kiry221982@gmail.com; ORCID ID: https://orcid.org/0000-0002-6062-550X.

Гуменюк Ігор Володимирович – кандидат технічних наук, старший викладач кафедри захисту інформації та кібербезпеки, Житомирський військовий інститут імені С. П. Корольова, Житомир, Україна;
Igor Humeniuk – PhD (Technical), Senior lecturer at the Department of Information Security and Cybersecurity, Korolov Zhytomyr Military Institute, Zhytomyr, Ukraine;
e-mail: ig_gum@ukr.net; ORCID ID: https://orcid.org/0000-0001-5853-3238.

Уварова Тетяна Володимирівна – науковий співробітник, Національний університет оборони України імені Івана Черняховського, Київ, Україна;
Tetyana Uvarova – Researcher, National University of Defense of Ukraine named after Ivan Chernyakhovsky, Kyiv, Ukraine;
e-mail: rhbz777@gmail.com, ORCID ID: https://orcid.org/0000-0002-5807-2563.

Метод автентифікації користувачів до об'єктів критичної інфраструктури на основі ідентифікації особи за голосовим повідомленням

В. М. Триснюк, Є. І. Нагорний, К. В. Сметанін, І. В. Гуменюк, Т. В. Уварова

Анотація. Предметом дослідження статті є захист об'єктів критичної інфраструктури на основі ідентифікації особи за голосовим повідомленням від несанкціонованого доступу. **Мета роботи** розглянути основні методи і алгоритми ідентифікації особи за голосом, а також розроблення програмної реалізації системи контролю доступу до забороненої зони, яка дозволяє покращити захищеність стандартної парольної системи шляхом запровадження механізмів голосової авторизації. Удосконалення алгоритму виділення фрагментів голосу в аудіофайлах. В статті вирішуються наступні **завдання:** аналіз переваг та недоліків існуючих методів захисту інформації до об'єктів критичної інфраструктури; визначення оптимального методу ідентифікації особи за голосовим повідомленням; характеристика основних особливостей впровадження і використання мел-частотних кепстральних коефіцієнтів (MFCC); Використовуються такі **методи**: системний підхід, системний аналіз, теорія алгоритмів. Отримано наступні **результати:** розроблено алгоритм збереження голосових даних користувачів з метою їх подальшого використання під час авторизації, підсистему голосової авторизації; розроблено відповідне програмне забезпечення комп'ютерної системи. **Висновки:** на прикладі застосування біометричного методу автентифікації користувачів до об'єктів критичної інфраструктури на основі ідентифікації особи за голосовим повідомленням, саме цей метод автентифікації користувачів слід застосовувати як додатковий спосіб багатофакторної автентифікації шляхом введення додаткового пароля або ще одного метода біометричної автентифікації користувачів.

Ключові слова: біометрична ідентифікація; ідентифікація особи за голосом; автентифікація особи за голосовим повідомленням; захист інформації; несанкціонований доступ; кібербезпека.

Метод аутентификации пользователей к объектам критической инфраструктуры на основе идентификации личности по голосовому сообщению

В. Н. Триснюк, Е. И. Нагорный, К. В. Сметанин, И. В. Гуменюк, Т. В. Уварова

Аннотация. Предметом исследования статьи является защита объектов критической инфраструктуры на основе идентификации личности, по голосовому сообщению, от несанкционированного доступа. **Цель работы -** рассмотреть основные методы и алгоритмы идентификации личности по голосу, а также разработка программной реализации системы контроля доступа к запрещенной зоны, которая позволяет улучшить защищенность стандартной парольной системы путем введения механизмов голосовой авторизации. Усовершенствование алгоритма выделения фрагментов голоса в аудиофайлах. В статье решаются следующие **задачи:** анализ преимуществ и недостатков существующих методов защиты информации к объектам инфраструктуры; определение оптимального метода идентификации личности по голосовому сообщению; характеристика основных особенностей внедрения и использования мел-частотных кепстральных коэффициентов (MFCC). Используются такие **методы:** системный подход, системный анализ, теория алгоритмов. Получены следующие **результаты:** разработан алгоритм сохранения голосовых данных пользователей с целью их дальнейшего использования при авторизации, подсистеме голосовой авторизации; разработано соответствующее программное обеспечение компьютерной системы. **Выводы:** на примере применение биометрического метода проверки подлинности пользователей к объектам критической инфраструктуры на основе идентификации личности, по голосовому сообщению, именно этот метод аутентификации пользователей следует применять как дополнительный способ многофакторной аутентификации путем ввода дополнительного пароля или еще одного метода биометрической аутентификации пользователей.

Ключевые слова: биометрическая идентификация; идентификация лица по голосу; аутентификация пользователя по голосовому сообщению; защита информации; несанкционированный доступ; кибербезопасность.