DOI: 10.54309/IJICT.2023.11.3.004

Akhmetova D.

International Information Technology University, Almaty, Kazakhstan
Research advisor: Amanzholova S.T.

ENCRYPTION EFFICIENCY OF VARIOUS STEGANOGRAPHIC METHODS

Abstract. At present , great attention is paid to the secure delivery of information and files on the Internet. Thus, the exchange of confidential information through various channels remains inevitable and there is a need for both encryption and hiding the fact of transmission. Word "Steganography" takes roots from the Greek word steganos, which means covered. Steganography is a technique for concealment the transmitting of a secret message. There are several methods to achieve this goal for embedding a secret message in another file, be it a picture, video, audio or text. This article is an attempt to analyze the various steganography techniques depending on the file types and identify the most effective ones.

Keywords: Steganography, encryption, steganographic methods, stegosystem, message hiding.

Introduction

One of the promising areas of information security is formed by modern methods of steganography. Steganographic protection is designed to hide the fact of the presence (transmission) of information. Thus, steganographic encryption methods allow information to be transmitted through different types of files using different algorithms. When transmitting by steganographic methods, a covert channel is organized on the base and inside an open channel using the features of information perception, and for this purpose the following techniques can be used:

- complete concealment of the fact of the existence of a covert communication channel;
- making it difficult to detect, retrieve or modify transmitted hidden messages inside open container messages;
- concealment of hidden information in the protocol [1].

Steganography has many areas of application like secure transfer of important information, documents, passwords [2]. But there is a possibility that the message transmitted by the steganographic method may be susceptible to attacks from the outside. Since the main areas of application of watermarks for data protection include the owner identification, proof of ownership, tracking interactions, data authentication, control of illegal copying, device management, compatibility of different technologies, so the unreliability of steganographic channels leads to the disclosure, and, possibly, further alteration of the transmitted information [3]. Therefore, any stegosystem message must undergo such challenges as:

- non detectability by visual inspection;
- lossy compression –shrink or enlarge a message;
- conversion to another format;
- message delay;
- deleting a part of a message [4];

In computer steganography, various digitized data can be used as containers: raster graphic images, digital sound, digital video, all kinds of digital information carriers, as well as text and other electronic documents. Steganography could be embedded in many types of files, including text, image, network protocol, audio and video. We propose comparison of algorithms according to their file type, such as:

- Line-shift-coding, word-shift coding, Feature coding, language synonym system (for text steganography)
- LSB method, Echo methods, Phase coding (for audio steganography)
- LSB method and Compression (for images)

In this paper we have identified their features, complexity, capacity, detectability, invisibility and presented in discussion section.

The remainder of the paper is organized as:

Section II discusses methods and methodology, gives a review on different stego algorithms' principles of operation.

Section III Results and Discussion presents evaluation of methods, discusses their parameters, advantages and drawbacks.

Section IV concludes the entire paper.

Methods and Methodology

In this section we have presented approaches and algorithms with examples for different types of data.

1.1 Text steganography

We have considered methods for text file types, such as: Line-shift coding, Word-shift coding, Feature coding, Language synonym system

Line-shift coding represents changing the distance between lines of electronic text. It is also called the line spacing method. The maximum and minimum distance between the lines is allocated, which allows to encode the characters "1" and "0", respectively, shown on Fig.1.

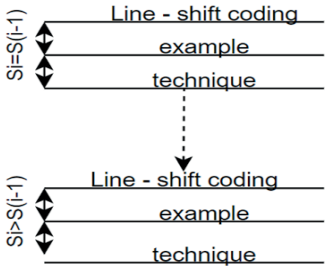

Figure 1 - Line-shift coding

As we can notice this method has low efficiency due to the size in bits of the information to be deposited cannot exceed the number of lines in the container. Therefore, various text editors can convert spacing [5].

Word shift coding is changing the distance between words in one line of the text. The method consists in the fact that information settling is based on changing the distance between the words of the container text. The Fig.2 shows an example of embedding a binary sequence 0101100100 into a text-container. The transition from single space to double encodes ”1”, the transition from double space to single encodes ”0”.

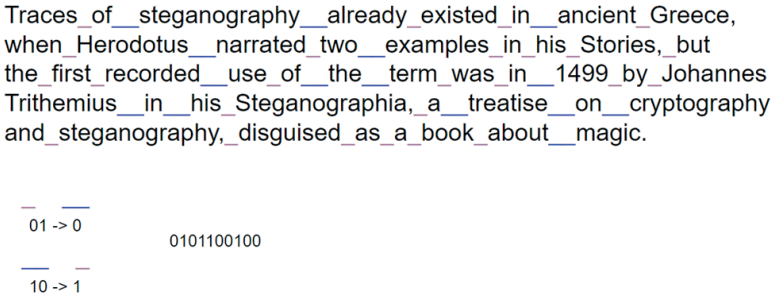

Figure 2 - Word shift coding

Feature coding is making specific changes to fonts. This method consists in changing the spelling of individual letters of the standard font used. Thus, the letter “A” can be modified by changing its font from “Times New Roman” to “Georgia”. In this case, you can encode the stego message so that the modified letter will mean ”1”, and the unmodified letter - ”0”. The result of embedding the secret message ”1” into the text-container “A”, when using the feature coding method and text processor MS Office Word is shown on Fig. 3.

Figure 3 - Feature coding

One of the most popular methods is a method based on the synonym system of the language used to write electronic text. Studies conducted for the case of the English language showed that the average number of synonyms in one subset of synonyms is 2.56. The minimum number of synonyms in one set of synonyms is 2, and the maximum is 13. As an example, let us give the set of synonyms S0: “propensity”, “predilection”, “penchant”, “proximity”. In the given set of synonyms, each word has a single identical semantic meaning, which allows each word to be encoded with its own unique code, for example, “propensity” - 00, “predilection” - 01, “penchant” - 10, “proximity” – 11 [6].

1.2 Audio steganography

For audio files the following algorithms are considered: LSB, Echo Hiding, Phase Coding.

LSB means least significant bit, is based on embedding a bit from messages in the least significant position (8th bite) of the audio cover in a deterministic method, as shown on Fig. 4.

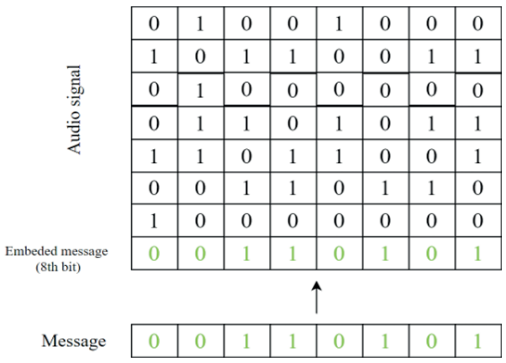

Figure 4 - LSB method

Sampling frequency - the frequency with which the signal is converted from analog to digital. Time sampling means that the signal is represented by a number of its samples (samples) taken at regular intervals, Measured in Hz [7]. Bit depth (sound depth) is the number of bits of digital information for encoding each sample, which means with what accuracy the input signal is measured. The greater the bit depth, the smaller the error of each individual conversion of the amplitude of the signal into a number. With the smallest bit depth possible, there are only two options for measuring audio fidelity: 0 for silence and 1 for full volume. For a bit depth of 8, 2^8 = 256 different values can be obtained. Accordingly, for 8 kHz sampled audio with bit depth equals 8, 8 kbps data of secret message could be hidden. Therefore, LSB method is quite simple in its implementation and has a great hiding capacity and can be used with other cryptographic or steganographic cyphers. Despite these advantages of this method, without additional ciphers and tools, the LSB can be easily decrypted and decomposed into the original container and message [8].

Echo Hiding embeds a secret message into segments of audio signals using a short echo is a repetition of the original audio signal. The embedded echo depends on the following three parameters: initial amplitude, offset (delay), and decay rate. For the intervening echo to remain unnoticed, there are several ways to hide it. First, the time delay between the original signal and the echo should not exceed 1ms. Secondly, the amplitude and decay rate must be below the level of audibility of the human ear. The echo embedding scheme shown on Fig. 6.

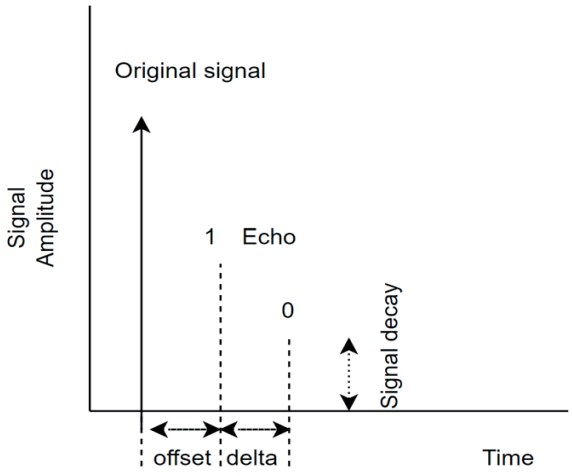

Figure 6 - Echo hiding method

To embed the echo, two impulses are used - the first copies the original signal, the second - embeds the echo. In this method, when encrypting “1”, time equal to offset is used. To encrypt “zero” = offset + delta. This method is resistant to the addition of noise. However, the disadvantage of this method is that the echo size is limited [9].

The phase encoding method differs from the previous methods in that it uses phase shifts to embed the message, as shown on Fig. 7.

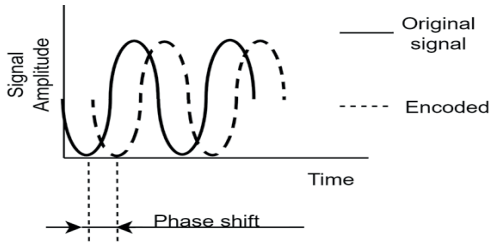

Figure 7 - Phase coding method

This method consists of the following steps:

1. The initial audio file is divided into a header and content.
2. The content of the original audio signal is divided into a number of segments equal to the length of the secret message.
3. Each segment is subjected to a Fourier transform to determine the phases of the signals.
4. The converted segments are phase-shifted according to the secret message bit, as shown on Fig. 8.

$$0 = \text{old phase} + \frac{\pi}{2}$$

$$1 = \text{old phase} - \frac{\pi}{2}$$

Figure 8 - Phase shifting according to the value of secret bit

5. The encrypted segments are subjected to an inverse Fourier transform and connected to the original header [10, 11]. This process is shown on Fig. 9.

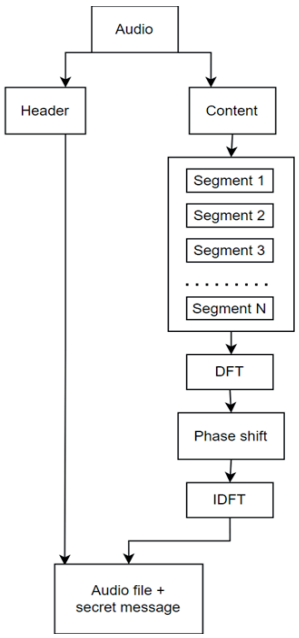

Figure 9 - Phase coding process

1.3 Image steganography

Digital images can be called a matrix, which is a multidimensional array of numbers - pixels, each of which is represented as bits and corresponds to a certain color, i.e. saturation of light. Each pixel is described using three primary colors - red, green and blue. In decimal, the range from 0 to 255 is used to describe each of these three colors, which is 8 bits in binary. Image steganography is divided into two groups: Transform domain and Image domain [12]. Image techniques, also known as spatial domain, embed the secret message directly into the pixel intensity by changing its value, while the transform, also known as frequency, uses manipulations such as compression, various mathematical transformations, and algorithms. In this section we have considered two algorithms for image steganography:

- JPEG Compression (Transform Domain)
- Least Significant Bit (Image Domain)

When executing this algorithm, compressing a DCT JPEG, the first step is to convert the color space from RGB to YUV. The YUV color space is mainly used for photo processing, reducing the color bandwidth to accommodate human perception. "Y" means brightness in grayscale (Luminance, Luma), "U" and "V" are chroma, concentration [13, 14].

Compared to RGB video signal transmission, its biggest advantage is that it requires small bandwidth, while RGB requires three independent video signals to be transmitted simultaneously. Another important step is to change the range of pixel values from -128 to 127 instead of 0 to 255, which is the standard range for 8-bit images. This compression process divides the already converted YUV image into blocks of 8 by 8 pixels, each of which is transformed using the DCT (Discrete Cosine Transform) algorithm, consisting of Fourier transforms. After lossy compression, the secret bits are already built into a new matrix of pixel values . This is done before applying the Huffman code to further lossless compression of the DCT coefficients [15]. The JPEG compression result can be seen in Fig. 10.

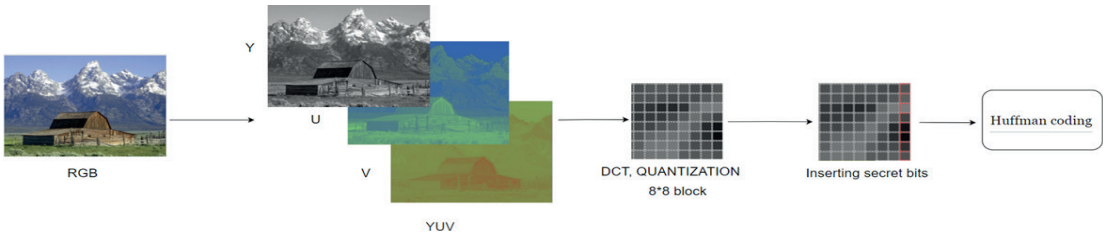

Figure 10 - JPEG compression

The advantage of steganographic techniques in the JPEG compression algorithm is that the changes that have been made to the original image are invisible to the human eye. JPEG compression also has lossless compression - the lossy compression part consists of Discrete Cosine Transform and quantization, and the lossless compression part consists of Huffman coding, which is done after the secret bits are embedded. That is why it is difficult for attackers to reveal hidden data [16].

LSB is a very common, easy-to-implement, capacious method of injecting and extracting secret information. It refers to covering an image using a spatial domain. In this algorithm, the least significant bit (the rightmost bit, or 8th bit) is replaced by the secret message bit. LSB encryption for images is shown on Fig. 11.

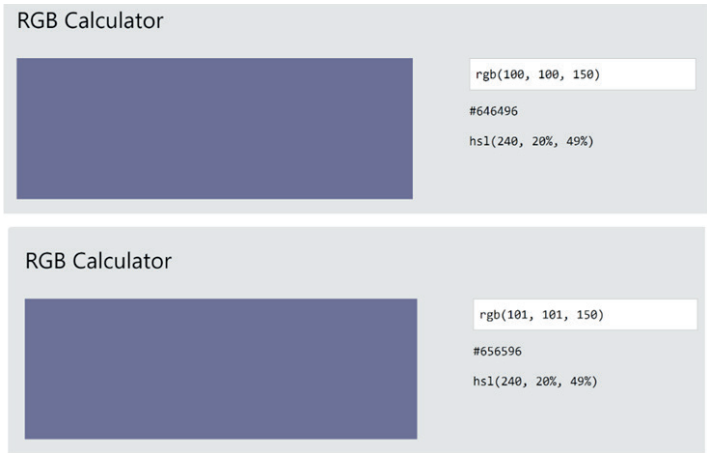

Figure 11 - LSB technique

As previously described, each pixel in an image is made up of red, green, and blue. The color from the picture of our example in binary :(1100100, 1100100, 10010110). Let's say we want to add secret bits whose sequence = 110. Then our example will look like: (1100101, 1100101, 10010110). To the human eye, the change remains imperceptible.

Results and Discussion

This section will present the results obtained by analyzing the algorithms by file type from the previous section. Here we have considered such parameters as invisibility, capacity, detectability, complexity, advantages, and disadvantages to analyze effectiveness of different methods.

Hiding capacity is the total number of bits of the secret message that the image (stegocontainer) can hold. This value must be high to deliver a significant number of secret bits.

Invisibility is appeared here from the perspective of a passable inline container. Those how noticeable are the external changes in the encoded file in comparison with the original, before embedding the secret message.

Detectability is responsible for how quickly and easily an attacker can uncover a secret message upon learning of its existence. How easy he can recognize the algorithm and decode it.

Complexity - how difficult the algorithm is in implementation, the complexity of entering a secret message.

1.1 Text steganography methods evaluation

We propose an augmented characteristic for text steganography methods in Table 1.

In methods such as Line-shift coding, Word-shift coding, Feature coding, the algorithm may be completely useless since many text editors have built-in algorithms for formatting text data, removing extra spaces and making equal line spacing. Therefore, these methods are considered unreliable. And the Synonym system of the language method needs to compose and search for synonyms, replace them in the text and also agree with the recipient about how many and which bits are responsible for each of the synonyms. However, a hacker can declassify such a message only if he learns about the exchange of synonym codes.

Table 1 – «Comparison of text steganography techiques»

Method	Invisibility	Capacity	Detectability	Complexity	Advantages	Disadvantages
Line-shift coding	Low	Low	High	Low	Easy to implement	Text editor can format this method
Word-shift coding	Low	High	High	Low	Easy to implement	Text editor can format this method, Visible to everyone
Feature coding	Low	High	High	Low	Easy to implement	Text editor can format this method, Visible to everyone
Synonym system of the language	High	Medium	Low	High	Very well suited for sending a secret message, because cannot be detected without knowing that there is a message	It will take time to prepare and select synonyms, The recipient and the sender agree in advance on the code of each synonym.

1.2 Audio steganography methods evaluation

Characteristics for audio steganography methods we can see in Table 2.

The LSB method is universal and is suitable not only for audio files, but also for photographs. It is the most widespread because of its simplicity of implementation and the fact that a sufficiently large amount of secret message can be placed (every 8th byte is encoded). However, its proliferation affects its detectability and ease of decryption. Just like the previous method, the Echo Hiding divides the audio file into discrete chunks. Not difficult to implement, and undergoes many changes, including noise. Phase Coding also withstands noise, is undetectable and very difficult to decipher.

Table 2 – «Comparison of audio steganography techiques»

Method	Invisibility	Capacity	Detectability	Complexity	Advantages	Disadvantages
LSB	High	High	High	Low	Versatile, easy to implement.	Suffers from added noise and not secure
Echo Hiding	High	Low	Medium	High	Noise sensitivity is eliminated	Echo size is limited
Phase Coding	High	High	High	High	Eliminates the disadvantages of other noise reduction methods of audio steganography	Rarely used due to the complexity of implementation

1.3 Image steganography methods evaluation

The compression method is also useful in that it not only hides the message, but also performs compression. However, over-compression sometimes affects image quality. As described above, LSB is suitable for almost all file types, encodes messages quickly and requires little technical power [17, 18]. Parameters for image steganography methods we can observe in Table 3.

Table 3 – «Comparison of image steganography techiques»

Method	Invisibility	Capacity	Detectability	Complexity	Advantages	Disadvantages
JPEG Compression (Transform Domain)	High	Medium	Low	Medium	Compression resistant, Hard to break the algorithm, Need low processing power	Has blocks artifacts means loss of some information.
Least Significant Bit (Image Domain)	High	High	High	Low	Versatile, easy to implement.	Suffers from image compression

Conclusion

The relevance of Digital Steganography at present lies in hiding of transmitted data used both for peaceful purposes to transmit important information, protect property rights, and for terrorist purposes. However, there are also problems that steganographic methods of protecting information face, be it audio, which can be subject to noise, or pictures, which can be compressed or partially removed. For text files, it was found that methods such as Line-shift coding, Word-shift coding, Feature coding can be ineffective and visible to the reader. Language synonym system is very effective and invisible however it requires the sender's compilation. A method such as LBS is used for many types of files and is difficult to detect for an ordinary person, but it is not very reliable due to its easy decoding by an attacker. The most reliable for audio files is the Echo Hiding method. For photographs - JPEG compression since it is very difficult for an intruder to decode a secret message. Thus, for each file type, depending on the importance of the secret message and the required encoding capability, different types of steganography can be used.

REFERENCES

1. Urbanovich P.P. Protection of information by cryptography methods, steganography and obfuscations, Minsk, 2016.

2. Gribunin V.G., Okov I.N., Turintsev I.V. Digital steganography. - M.: Solon-Press, 2002.– 272 p.

3. Cox I., Miller M., Bloom J., Fridrich J., Kalker T., Digital Watermarking and Steganography, 2017.

4. Kumar A.S., Sahu M. Digital image steganography and steganalysis: Ajourney of the past three decades, Open Computer Science, October 2020.

5. Harpreet K., Jyoti R. A Survey on different techniques of steganography, Bathinda, Punjab, India, 2016.

6. Surana J., Sonsale A., Bhavesh J., Sharma D., Choudhary N. Steganography Techniques, India, 2017.

7. Djebbar F., Ayad B., Abed K.M., Hamam H. EURASIP Journal on Audio, Speech, and Music Processing volume 2012, Comparative study of digital audio steganography techniques.

8. Konakhovich G.F., Puzyrenko A.Y. Computer steganography. Theory and practice. - : MK-Press, 2006. - 288 p., Ill.

9. Gribunin V.G., Okov I.N., Turintsev I.V. Digital steganography. - M.: Solon-Press, 2002 .– 272 p., Ill.

10. Ryabko B.Ya., Fionov A.N. Fundamentals of modern cryptography and steganography. - 2nd ed. - M: Hotline - Telecom, 2013. - 232 p., Ill. - ISBN 978-5-9912-0350-0.

11. Zavyalov S.V., Vetrov Yu.V. "Steganographic methods of information protection": textbook. - SPb.: Publishing house of Polytechnic. University, 2012.– 190 p.

12. Morkel T., Eloff J.H.P., Olivier M.S. An overview of image steganography, Information and Computer Security Architecture (ICSA), Pretoria, South Africa

13. Goel S., Rana A., Kaur M. A Review of Comparison Techniques of Image Steganography, 2017

14. Bykov S.F. JPEG compression algorithm from the standpoint of computer steganography // Information Security. Confident. - SPb.: 2000, No. 3.

15. Gribunin V.G., Zherdin O.A., Martynov A.P., Nikolaev D. B., Silaev A. G., Fomchenko V. M. Fundamentals of steganography // Ed. Dr. tech. Sci. V.G. Gribunin, Trekhgorny, 2012.

16. Gabidullin E.M., Pilipchuk N.I. Lectures on information theory. Moscow, MIPT, 2007 .– 213 p. : ill. - ISBN 978-5-7417-0197-3

17. Hoffman R. (2012). Data Compression in Digital Systems. Springer Science & Business Media. p. 255. ISBN 9781461560319.

18. Ojaas H. Image Compression — DCT Method, DCT based Image Compression, 2021.

Ахметова Д.
Ғылыми жетекші: Аманжолова С.Т.

Түрлі стеганографиялық әдістерді шифрлеу тиімділігі

Аңдатпа. Қазіргі уақытта ғаламторда ақпарат пен файлдарды қауіпсіз жеткізуге үлкен көңіл бөлінуде. Осылайша, әртүрлі арналар арқылы құпия ақпарат алмасу сөзсіз болып қалады және шифрлауды да, беру фактісін жасыруды да қажет етеді. «Стеганография» сөзі гректің «стеганос» сөзінен шыққан, бұл жабық дегенді білдіреді. Стеганография - бұл құпия хабарламаны жіберуді жасыру әдісі. Құпия хабарды басқа файлға, мейлі ол сурет, бейне, аудио немесе мәтін болсын, ендіру үшін осы мақсатқа жетудің бірнеше әдістері бар. Бұл мақала файл түрлеріне байланысты әртүрлі стеганография әдістерін талдауға және ең тиімділерін анықтауға тырысады.

Түйін сөздер: Стеганография, шифрлау, стеганографиялық әдістер, стегожүйе, хабарламаны жасыру.

Ахметова Д.
Научный руководитель: Аманжолова С.Т.

Эффективность шифрования различных стеганографических методов

Аннотация. В настоящее время большое внимание уделяется безопасной доставке информации и файлов в сети Интернет. Таким образом, обмен конфиденциальной информацией по различным каналам остается неизбежным и возникает необходимость как в шифровании, так и в сокрытии факта передачи. Слово «стеганография» происходит от греческого слова steganos, что означает «скрытый». Стеганография — это метод сокрытия передачи секретного сообщения. Есть несколько методов для достижения этой цели по выстраиванию секретного сообщения в другой файл, будь то изображение, видео, аудио или текст. Эта статья анализирует различные методы стеганографии в зависимости от типов файлов и выделяет наиболее эффективные из них.

Ключевые слова: стеганография, шифрование, стеганографические методы, стегосистема, скрытие сообщений.

Авторлар туралы ақпарат:

Аманжолова Сауле Токсановна - п.ғ.к., Халықаралық ақпараттық технологиялар университетінің киберқауіпсіздік бөлімінің меңгерушісі.

Ахметова Дарья - Халықаралық ақпараттық технологиялар университетінің «Ақпараттық коммуникациялық технологиялар» дайындық бағытының 1-курс магистранты.

Сведения об авторах:

Аманжолова Сауле Токсановна - к.т.н., заведующая кафедрой Кибербезопасность Международного университета информационных технологий.

АхметоваДарья–магистрант1-гокурсанаправленияподготовки«Информационныекоммуникационные технологии» Международного университета информационных технологий.

About the authors:

Amanzholova Saule Toksanovna - Ph.D., Head of Cybersecurity Department, International Information Technology University.

Akhmetova Darya - 1st year mMaster's student of the of "Information Communication Technologies", International Information Technology University.