Journal of Education for Pure Science- University of Thi-Qar
Vol.12, No2 (December, 2022)
Website: jceps.utq.edu.iq *Email: jceps@eps.utq.edu.iq*

DOI: http://doi.org/10.32792/utq.jceps.12.02.35

Developing an algorithm and investigation its properties to generate protected data

Hyder Yahya Alshaeaa[1]

[1] Computer Science Department, College of Education for Pure Sciences, University of Thi-Qar.
Thi-Qar, Iraq.

Received 9/5/2022 Accepted 10/5/2022 Published 01/12/2022

Abstract:

Reliability plays an important role in increasing the protection of legal data exchanged between systems and devices, as well as between the sender and the receiver, where it is necessary the device receives legal data only from corresponding and authorized devices. Otherwise, the received data from illegal device can be processed by the device, therefore may lead to denial of service or errors in the device operation, or maybe possibly a loss of data or functioning. The potential attacker is the real threat to the trusted channel of data exchange, who deliberately generates and transmits data to the receiving device, which the device may mistakenly perceive as a trusted and authenticated source. In order to determine the difficulties associated with the transfer of legal data, this study focuses on development and investigation of the properties of algorithm for the formation of the protected data of small size. The study of algorithm for the presence of uneven frequencies of the meeting different words of the sources. It is shown that the absence of these uneven frequencies does not allow an attacker to use methods to increase either the probability of opening the parameters of the algorithm, or the probability of successful implementation of various types of attacks on the data transmission.

Keywords: Limited length of the data, Analysis of information blocks, Authentication control, Probability calculation of hash.

1. Introduction

The constantly growing level of computer technology, information and telecommunications technologies have led to the creation of a worldwide unified information space, where the resources of public computer networks used for storing, processing and transmitting information. One of the most important requirements for data transmission systems between two subscribers is to ensure the authenticity of the transmitted information and its confidentiality [1]. To ensure the authenticity of the transmitted information, it is not enough to use information security mechanisms based on cryptology methods. Due to restrictions on the use of cryptographic tools (Security tokens, Key-Based Authentication, Cryptography Architecture), it is necessary, along with encryption algorithms for transmitting information, to use methods to ensure and determine its authenticity [2]. There is a growing demand for those methods and mechanisms of information protection that used in the development of a generalized architecture for the transmission of protected data [3]. At the physical level, the most popular way of transmitting the open system interaction model is sequential data transfer between subscribers [4]. A similar method used in the block chain technology [5].

The existing methods of ensuring the authenticity of transmitted information over open communication channels based on the use of block encryption algorithms, which the result in the information is processed by frames (information blocks), which means that there is no continuous transmission [6]. The information block (IB) in this paper will understand as a set of data that was sent between the source and the recipient continuously over time, taking up the communication channel exclusively. Most current block-by-frame transmission protocols have feature mechanisms for structuring received frames, which is the conventional way of transferring information between the source and recipient devices [7], which divides the data stream into several packets [8].

If we assume that the information sent by the source was separated into a series of IB, and that the blocks were sented to the communication channel in order, starting with the first block, it does not follow that the receiver would receive them in the same order. There might have been a variety of causes for such an occurrence. First and foremost, they are the characteristics of information packet routing when sent over a switched channel [9]. Transmission mistakes might arise in this situation, necessitating the re-sent of a specific IB in the sequence. Map– oriented communication protocols [10], which emphasis on high-speed data transfer without loss, are a type of non-switched channel that occasionally involves re-sent of IB. In this case, most current block-by-block transmission methods include procedures for streamlining received frames.

Furthermore, if the data flow is separated into a series of blocks and the block reception time is unequal, as it is frequently when blocks arrive from various sources [11], worries about the authenticity of block frames would be arise. The solutions to these problems may be found in the transmission Protocol's procedures, which involve the application of unique algorithms in each case to manage the sequence and authenticity of the IB in the Protocol that allows the transfer. The topic of ensuring the authenticity and authenticity of the IB whose length differs significantly from that recommended by encryption standards will be a cornerstone in this research area, and the findings of such studies can have a wide range of applications, ranging from their use in a variety of heterogeneous systems to radio authentication systems and identification. The purpose of this work is to investigate the properties of an algorithm that ensures the privacy, orderliness, and authenticity of tiny IB (about tens of bits).

2. Proposed algorithm

To lessen the probability of mistakes in the transmission of tiny IB, it is suggested that algorithms based on the creation of linked IB chains be used, with a single block's integrity and authenticity verified by its membership in such a single chain [12]. Combining information into a single chain and building an interdependence of information block is the most optimum solution to transmit various IB over a communication channel, that allows you to uniquely determine the authenticity of each received block. This allows such chains to develop more flexible and effective methods to ensure and determining their authenticity.

Obviously, forging the entire chain, even if it only consists of a dozen packets, is more difficult than falsifying one or two sent packets. As a result, the simplest approach for an attacker to affect the proposed systems is to generate random blocks and send them to the recipient [13]. When building a chain of blocks, the receiver may accidentally add one or more matching illegal blocks [14]. As a result, a chain of legal IB will be sent with an error and not will be processed, which is a well-known denial-of-service (DoS) attack [15]. As a result, it's important to safeguard information sent between sender and recipient [16]. In the work [17] is detailed in full, the method for creating IB and their subsequent integration into chains. Data exchange is described in detail in [18,19], and the block diagram of the source algorithm is shown in Figure 1.

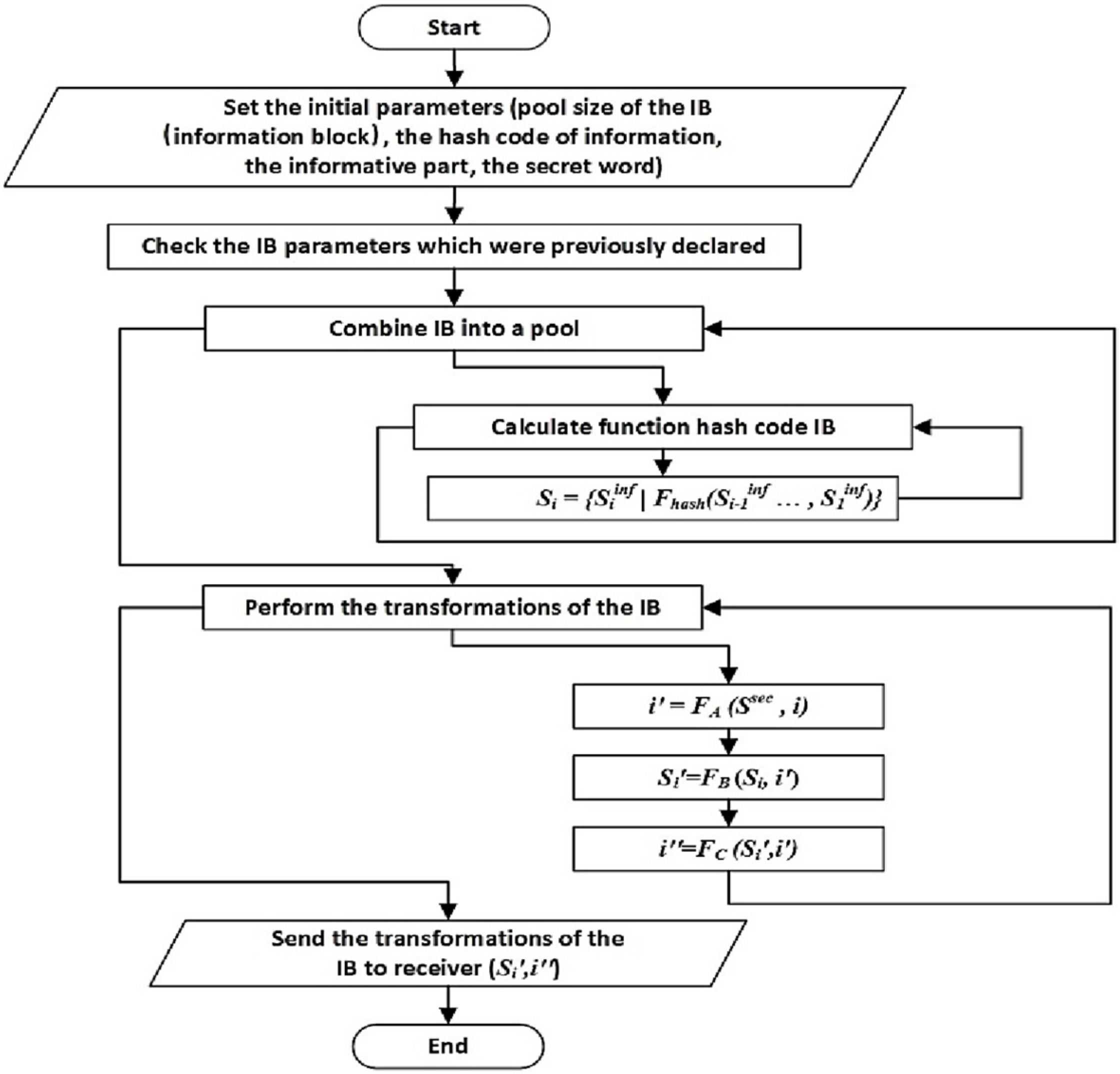
Start
Set the initial parameters (pool size of the IB (information block), the hash code of information, the informative part, the secret word)
Check the IB parameters which were previously declared
Combine IB into a pool
Calculate function hash code IB
$S_i = \{S_i^{inf} \mid F_{hash}(S_{i-1}^{inf} \dots, S_1^{inf})\}$
Perform the transformations of the IB
$i' = F_A(S^{sec}, i)$
$S_i' = F_B(S_i, i')$
$i'' = F_C(S_i', i')$
Send the transformations of the IB to receiver (S_i', i'')
End

Figure 1. The block diagram of the sender algorithm.

In the sender algorithm, performs the following operations:

1) Find the variable i' from $i' = F_A\left(S^{sec}, i\right)$;

2) Find the variable S_i' from $S_i' = F_B\left(S_i, i'\right)$;

3) Find the variable i'' from $i'' = F_C\left(S_i', i\right)$;

4) Sends the variables $\{ S_i' \mid i'' \}$ to receiver.

In the receiver algorithm, performs the following operations:

1) return the variable *i* from $i^{rec} = F_C^{-1}(S_i', i'')$;

2) return the variable i' from $i^{rec'} = F_A\left(S^{sec}, i^{rec}\right)$;

3) return the contents IB $S_i = F_B^{-1}\left(S_i', i^{rec'}\right)$;

where: i' - irreversible transformation, S_i' and i'' – reversible transformation. One block received as a result of sequencing S_i' and i'' of the pool's sequence number that was sent to the recipient [20]. The transmitted IB can be written as: $S=\{S^{inf} | S^{index} | S^{hash}\}$. Figure 2 depicts the format of the IB viewed by the receiver.

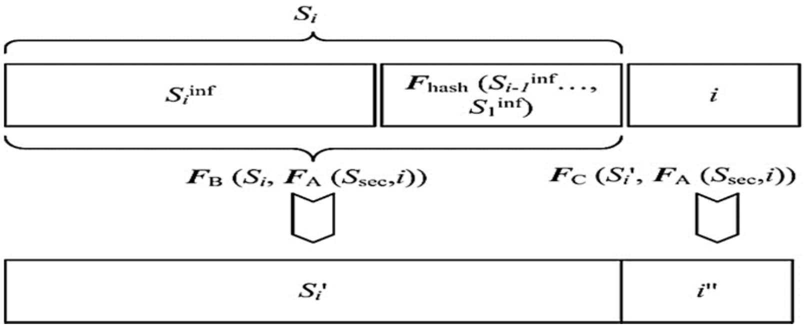
S_i
S_i^{inf}
$F_{hash}(S_{i-1}^{inf}..., S_1^{inf})$
i
$F_B(S_i, F_A(S_{sec}, i))$
$F_C(S_i', F_A(S_{sec}, i))$
S_i'
i''

Figure 2. Format of information blocks (IB).

When the performed experiment on the proposed algorithm, the initial parameters for the data transmission algorithm are the secret key S^{sec} (length 8-32 bits), the information sequence S_i^{inf} (16-32 bits) and the number of the transmitted data in the pool *i* (3-6 bits long enough to describe the sequence number in the pool) [21]. After the transformations completed, a data sent to the receiver, the length of which we have defined 16 – 32 bits. This is the standard bits width of the machine word, the algorithm under study focused primarily on the use of mutual recognition systems in software and hardware components of complex heterogeneous information systems. Because of the small size of the information processed and output words, it's difficult to apply well-known irreversible (cryptography hashing) and reversible algorithms (encryption). As a result, original algorithms were necessary to provide appropriate results on the entropy of output words on tiny amounts of original information [22], as well as to decrease the time and material costs of developing methods for service data analysis.

3. Analysis and discussion the results of the functions transformations

The task of applying the combination of reversible and irreversible functions transformations that described above is that the attacker could impose false data only by randomly forming their false information blocks, and not use the characteristics of the obtained codes to increase the likelihood of a various attacks on the considered information transmission system. The transformations described above, like any other transformations on the data, are the mapping of the set of source data sets $\{S^{sou}\}$ to the set of result values $\{S^{res}\}$. Since the total length of the words S^{sec}, *Si* and *i* greater than total word length *Si'*, *i"*, then several elements of the set $\{S^{sou}\}$ will be displayed in one element of the set $\{S^{res}\}$. The purpose of the study of the above algorithms for uneven display, when different elements of the resulting set corresponds to a different number of elements of the original. Such unevenness can allow an attacker to analyze the transmitted information, increase the likelihood of attacks on the system under study [21].

Introducing into consideration an attacker who, as we have agreed, can intercept information from the source to the receiver, we must consider its possible methods of investigation, that is, to investigate the properties of transformations in the conditions in which the attacker will conduct their observations. The analysis of the Protocol of interaction between the source and the receiver of data allows us to conclude that the attacker can, firstly, carry out permanent interception of all transferred data generated in different communication sessions based on the different session secret words S^{sec} and, secondly, it can intercept information within a single session, during which information is generated based on a single value session secret word S^{sec}.

As part of the study, we obtained graphs of the frequency of the meeting of certain bits sequences of forming output words $\{Si',i''\}$ full or partial (depending on the situation under study) iterate through the values of one or several fields of input words S^{sec}, Si and i. At the first stage, we studied the influence of the irreversible transformation function on the characteristics of the output values.

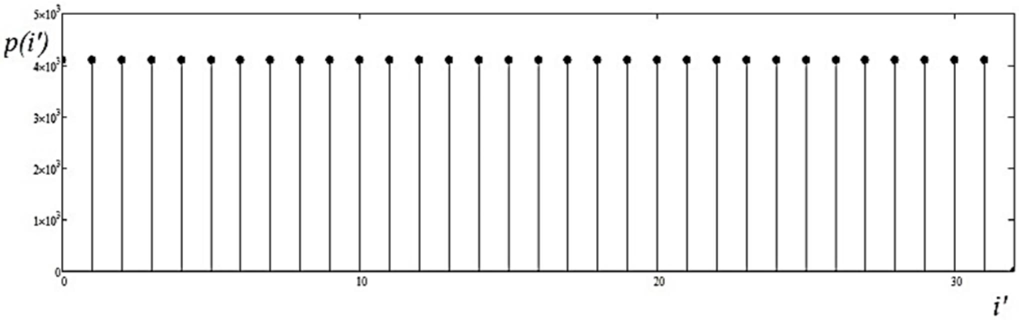

Figure 3. The frequency p of the meeting of the results of the execution function F_A over the set of words S^{sec} and i.

Figure 3 shows the frequency $\boldsymbol{p}$ of the meeting of words i' – result of an irreversible function F_A over S^{sec} and i. On the x-axis - the value of the function output F_A that presented in decimal format, the y-axis - the number of the input sets of words that give the specified value during transformation F_A. The graph was plotted with the length of the secret word S^{sec} 12 bits and length of the field i 5 bits. It can be seen that each of the 32 output words i' received at occurs 4096 times. A similar result is obtained for the other input word lengths of the function F_A. Hence, we conclude that the algorithm of this function satisfies the condition of uniformity of the frequency of occurrence the output words.

The frequency of occurrence of various output words was investigated $\{S_i',i''\}$ as the final result of execution F_A, F_B and F_C over the complete set of all possible values of the input words S_i and i, with a fixed value of the S^{sec}. In other words, data transmission was simulated within a single communication session, during which a single secret word S^{sec} is active. Figure 4 shows a graph of this relationship for the length of the secret word S^{sec} 12 bits and length of the field i 5 bits.

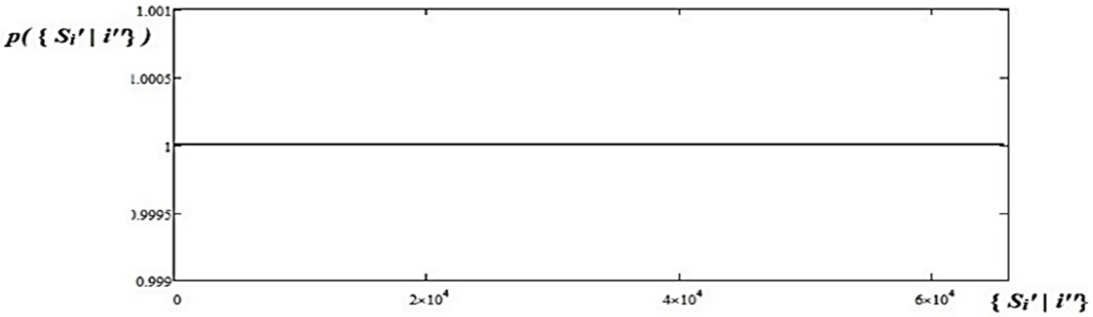

Figure 4. The frequency p of the results of execution F_A, F_B, and F_C over the complete set of all possible values of the input words *Si* and *i*.

It can be seen that within a specific data transfer session, each word {*Si, i*} is converted to one and only one word {*Si', i''*}, which also meets the requirements for codes that ensure confidentiality and authenticity. Similar results were obtained for other values of the length of the input fields *Si* and *i*, and for any values of the session word of the word S^{sec}.

It should be noted that the described algorithms are sensitive to the number of rounds and the order in which operands are fed to the input. For example, Figure 5 shows a graph similar to the graph in Figure 4, but constructed from a system in which the F_A algorithm is applied twice (the graph does not show the entire range of all values of the words {*Si, i*}, but only part of it).

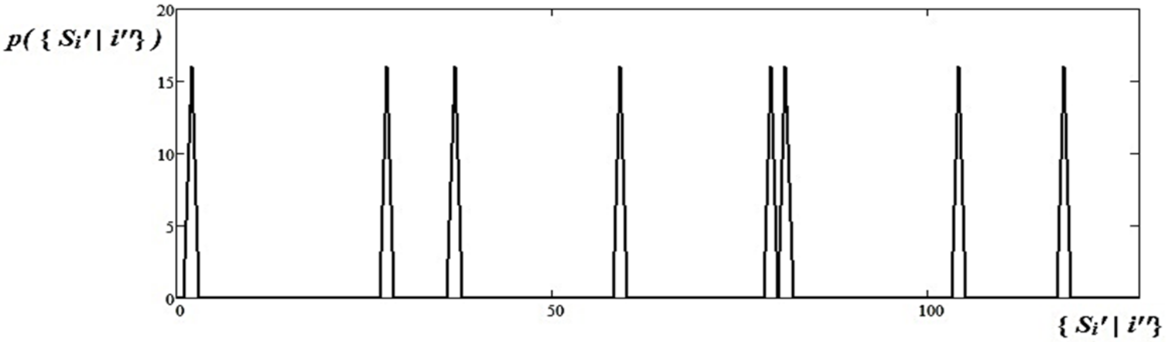

Figure 5. An example of the uneven frequency of meeting the results of execution functions over the set of values of the input words *Si* and *i*.

It can be seen that there is a clear unevenness in the output words, some values of the words {*Si', i''*} do not occur at all, and others every 16 times. Such unevenness, which the described algorithms lack, allows the attacker to carry out some types of attacks on the transmission process described in [22], due to the formation of extraneous information blocks not randomly, but based on the unevenness of the frequency of occurrence of words formed by a legal source in the communication channel.

4. Conclusions

The results that obtained allow us to conclude that the described transformation algorithms are suitable for use in transmission systems with a tiny size of information packet. The results of such study can have a wide range of applications, they provide uniform frequency characteristics of the output words, as well as easy to implement in hardware, which makes them suitable for use in communication systems with high requirements for response speed and power consumption.

5. References:

1. W. Zheng, S. Liu, Z. Liu и Q. Fu, «Security transmission of FTP data based on IPSec,» в *1st IEEE Symposium on Web Society*, Lanzhou, China, 2009.
2. T. Masaru, "An HTTP Extension for Secure Transfer of Confidential Data," in *IEEE International Conference on Networking, Architecture, and Storage*, Guilin, China, 2009.
3. J. Amit, "Enhancement of secure data transmission," in *Euro American conference on Telematics and information systems*, Faro, Portugal, 2007.
4. B. Wang, H. Qian, X. Sun, J. Shen and X. Xie, "A Secure Data Transmission Scheme Based on Information Hiding in Wireless Sensor Networks," *International Journal of Security and Its Applications,* vol. 9, no. 1, pp. 125-138, 2015.
5. M. Bellare, J. Kiliany and P. Rogawayz, "The Security of the Cipher Block Chaining Message Authentication Code," *Journal of Computer and System Sciences,* vol. 61, no. 3, p. 362–399, 2000.
6. W. Stallings, "NIST Block Cipher Modes of Operation for Authentication and Combined Confidentiality and Authentication," *Cryptologia,* vol. 34, p. 225–235, 2010.
7. L. Maruhnenko, "Analysis of potential vulnerabilities and modern methods of protecting multi-user resources," in *infocommunication and space technologies: status, problems and solutions*, Russia, Kursk, 2018.
8. M. A. Efremov, I. V. kaluutsky, M. O. Tanygin and I. I. Rudak, "Security of personal data, social networks and advertising in the global INTERNET," *News of Southwestern state University. Series: Management, computer engineering, computer science. Medical instrumentation.,* vol. 1, no. 22, pp. 27-33, 2017.
9. R. Sumathi, "A Secure Data Transfer Mechanism Using Single-Handed Re-Encryption Technique," in *International Conference on Emerging Trends in Science, Engineering and Technology*, Tiruchirapalli, India, 2012.
10. P. Hao and Q. Shi, "Matrix factorizations for reversible integer mapping," *IEEE Transactions on Signal Processing,* vol. 49, no. 10, pp. 2314 - 2324, 2001.
11. M. Bellare, J. Kilian and P. Rogaway, "The security of the cipher block chaining message authentication code," *JCSS,* vol. 3, no. 3, p. 341–358, 1994.
12. A. Gervais, O. Ghassan, K. Wüst, V. Glykantzis, H. Ritzdorf and S. Capkun, "On the Security and Performance of Proof of Work Blockchains," 2006.
13. M. O. Tanygin and A. P. Tipikin, "System architecture for hardware restriction of access to information on a computer hard disk," *Telecommunications,* vol. 3, pp. 44 - 46, 2006.
14. M. O. Tanygin, H. Y. Alshaeaa and V. Altukhova, "Establishing Trusted Channel for Data Exchange between Source and Receiver by Modified One-time Password Method," in *International Russian Automation Conference, RusAutoCon*, Sochi, Russian Federation, 2019.
15. J. Black, Authenticated Encryption. In Encyclopedia of Cryptography and Security, Tilborg: Springer, 2005.
16. V. P. Dobritsa, A. Plugatarev, A. A. Zakharyuta and L. A. Evlanova, "An Approach to Creating an Adaptive Algorithm for Transmitting Information in Open Channel with Mutual Authentication of the Source and Receiver," in *Intelligent Information Systems: Trends, Problems, Prospects; Materials of Reports of the Vi all-Russian Scientific and Practical Conference "IIS-2018".*, Kursk, 2018.

17. M. Tanygin, "Method of Control of Data Transmitted Between Software and Hardware," in *Computer Science and Engineering: Materials of the IV International Conference of Young Scientists CSE-2010*, Lviv, 2010.
18. Tanygin, M. O., Alshaeaa, H. Y., & Kuleshova, E. A. (2020). A METHOD OF THE TRANSMITTED BLOCKS INFORMATION INTEGRITY CONTROL. Radio Electronics, Computer Science, Control, (1), 181–189. https://doi.org/10.15588/1607-3274-2020-1-18
19. Tanygin M.O., Efremov M.A., Hyder Y.A. (2020) Analysis of the Secure Data Transmission System Parameters. In: Radionov A., Karandaev A. (eds) Advances in Automation. RusAutoCon 2019. Lecture Notes in Electrical Engineering, vol 641. Springer, Cham. https://doi.org/10.1007/978-3-030-39225-3_74.
20. A. G. Sivakov, M. O. Tanygin and V. S. Panishev, Information security, Kurck: Southwest state University, 2017, p. 196.
21. M. O. Tanygin, "Calculation of the probability of collisions when using the algorithm for controlling the authenticity of messages," *News of the southwest state University. series: management, computer engineering, Informatics,* vol. 2, pp. 179-182, 2012.
22. M. A. Ivanov, Cryptographic methods of security information in computer systems and networks, Moscow: MEPhI, 2012, p. 400.