UDC 629.4.067 DOI: 10.18372/2073-4751.71.16995

Aliyeva S.Y.,
orcid.org/0000-0003-1166-119X

PROTECTION OF INFORMATION DURING DATA TRANSFER IN OPEN NETWORKS

Azerbaijan Technical University

sevincalievaa@gmail.com

The rapid development of e-commerce and the growing number of users make us seriously think about the security problems of open data transmission channels. The absence of governing bodies has led to the fact that each member of the network must take care of its own security [1].

Connecting to open (global) networks, such as the Internet, significantly increases the efficiency of work and opens up many new opportunities. At the same time, care must be taken to create a system for protecting information resources from those who want to use, modify or simply destroy them. Information security involves maintaining the integrity, availability and, if necessary, confidentiality of information and resources used to enter, store, process and transmit data. To solve the complex problem of protection, a combination of legislative, organizational and software and technical measures is necessary [2].

This article is devoted to the influence of the specifics of open networks on solving the problems of ensuring the security of data transmission and creating libraries based on the best algorithms. It is shown how the features of the data affected the development of modern algorithms for their encryption [2,3].

Absolutely secure algorithms have existed for a long time, but have not yet found wide application in open networks.

The first absolutely secure algorithm was the Vernam algorithm, and its theoretical justification was developed by Shannon [4,5].

The maximum possible uncertainty of a fixed-size data block is reached when all possible values of this block are equally probable. In this case, it is equal to the block size in bits. Thus, the uncertainty of the key K does not exceed its length:

$$H(K) \leq K.$$

The absolute strength condition for ciphers that satisfy the Kirchhoff principle, $|K|\ H(K) = H(E)\ H(T) = |T|$, where E is the cipher and T is the original data.

In order for a cipher built according to the Kirchhoff principle to be absolutely secure, it is necessary that the size of the key used for encryption be no less than the size of the data being encrypted, i.e.

$$|K| \geq |T|.$$

Exact equality is possible only if all possible values of the key are equally probable. This is equivalent to the condition that the key bits are equally likely and statistically independent of each other [6,8].

An example of an absolutely secure cipher is Vernam's one-time gamma - imposing, with the help of some binary operation T °, on open data T a key K of the same size, composed of statistically independent bits that take on possible values with the same probability:

$$T' = T \circ K.$$

The operation used to overlay the gamma must satisfy the following conditions: the encryption equation must be uniquely solvable with respect to open data with known encrypted and key; key with known open and encrypted data.

For this reason, the most convenient operation to implement is usually chosen – bitwise modulo 2 summation (or bitwise exclusive OR), since it: 1) requires the simplest logic of all possible operations for its implementation; 2) inverse to itself, so the same procedure is used for encryption and decryption [6,7].

Another, absolutely stable algorithm is the family of notepad algorithms. At the preliminary stage, “notebooks” are exchanged,

representing an excess set of characters that can be used when transmitting a message. An example of such a notebook would be a book or a jumbled piece of information (preferred). At the stage of encryption, each character of the ciphered text is compared to different positions of this character in the notepad. Consequently, the absolute uncertainty of the transmitted information is achieved [8,9].

Thus, absolutely strong ciphers require the use of a key that is at least as large as the data being encrypted. Both the sender and the recipient must have this key, that is, it must first be delivered to them, and this requires a secure channel. Along with a potentially insecure channel for the transmission of encrypted data, the existence of a secure channel is necessary for the transmission of the same key size. This is not always acceptable for economic reasons, so such systems are used only in exceptional cases to protect information of particular value. The vast majority of real encrypted communication systems use algorithms that do not have absolute security and are therefore called imperfect ciphers [9,10].

The unavailability of the algorithm does not increase the security of the cipher; open algorithms are accepted as the standard.

There is no way to get the exact value of the complexity of cryptanalysis. All estimates are based on tests of cipher resistance to currently known types of cryptanalysis, and there is no guarantee that new analysis methods will not be developed in the near future that significantly reduce labor intensity. What has been said above means that, given the current state of affairs in cryptography, the security of absolutely all ciphers, with the exception of perfect ciphers, cannot be substantiated by evidence. Instead, it is empirically substantiated as resistance to the types of cryptanalysis known today, but no one can guarantee that tomorrow a type of cryptanalysis will not be invented that is successful for this particular cipher. That's why you should not trust the "latest ciphers" – they have not passed the test of time. For the same reason, it is not wise to trust cryptalgorithms that their authors keep secret – even in the absence of "hatchholes" maliciously left there, there is absolutely no guarantee that the algorithm has been investigated with all the necessary thoroughness. An example of a cipher proposed for implementation without disclosing its algorithm is Clipper from the US NSA. The requirement for the reconfigurability of the algorithm appears due to the fact that sooner or later the attacker may have at his disposal a description of the algorithm, its software or hardware implementation. In order not to have to completely replace the algorithm in this case, it must contain an easily replaceable part in all encryption nodes where it is used. This leads to the Kirchhoff principle: a cipher is defined as a parameterized algorithm consisting of a procedural part, that is, a description of exactly what operations and in what sequence are performed on the encrypted data, and parameters – various data elements used in the transformations. Disclosure of only the procedural part should not lead to an increase in the probability of successful decryption of the message by an attacker above the allowable limit. For this reason, and because declassification of this part is quite likely in itself, there is little point in keeping it secret. Some part of the parameters of the algorithm, which is called the cipher key, is kept secret:

$$T' = E (T) = EK (T),$$

where K is the cipher key.

Using the Kirchhoff principle allows us to draw some conclusions about the construction of ciphers. The disclosure of a specific cipher (algorithm and key) does not lead to the need to completely replace the implementation of the entire algorithm, it is enough to replace only the disclosed key. Keys can be alienated from other components of the encryption system (stored separately from the implementation of the algorithm, in a more secure place) and loaded only as needed and for the duration of the encryption (this significantly increases the reliability of the system as a whole). It becomes possible to accurately estimate the "degree of uncertainty" of the encryption algorithm – it is simply equal to the uncertainty of the key used:

$$H(EK) = H(K).$$

Accordingly, it becomes possible to estimate the probability and complexity of successful decryption, that is, the amount of computational work that needs to be done for this [8,9].

The vast majority of processors have a 32-bit architecture, which imposes its own limitations on the software implementation of the cipher. For comparison, we present the characteristics of Russian GOST 28147-89 (hereinafter referred to as GOST) and American (DES) standards developed at approximately the same time (tab. 1).

Table 1. DES Algorithm Data Encryption Standard. DES Features Block cipher, 64 bits per block 64-bit key, with only 56 bits effective ECB mode and CBC mode.

P A R A M E T R	GOST	DES
1. Encryption block size	64 bits	64 bits
2. Key length	256 bits	56 bits
3. Number of rounds	32	16
4. Substitution nodes (S-blocks)	are not fixed	fixed
5. Key length for one round	32 bits	48 bits
6. Scheme for generating a round key	Simple	Complex
7. Initial and final bit permutations	No	yes

Since DES was developed during the 8-bit architecture, today there are a number of inconveniences in the software implementation of this algorithm on modern PCs, which leads to the fact that GOST, despite the longer key length and more rounds, works faster [3,4].

Specially designed for 32-bit machines, and also significantly faster than DES, is the popular Blowfish.

The forthcoming transition to a 64-bit architecture creates a requirement for future algorithms to take into account the features of this architecture, although this is not necessary for hardware implementation [4,5].

Sometimes it is not possible to use a secure channel at the preliminary stage. This led to the division of algorithms into classical – symmetric and new, asymmetric encryption algorithms [4].

In "symmetric" crypto algorithms (DES, GOST, Blowfish, RC5, IDEA), the same key is used both for encryption and for restoring an open message. Therefore, this key is secret. The advantage of these algorithms is their good theoretical knowledge, including the rationale for cryptographic strength. Compared to "asymmetric" algorithms, one should note the relative simplicity of both software and hardware implementation, a higher speed of operation in the forward and reverse directions, as well as the provision of the necessary level of protection when using significantly shorter keys. The main disadvantages include the need to provide additional secrecy measures when distributing keys, the problems associated with this and, possibly, the costs, as well as the fact that secret key algorithms work only in conditions of full trust of correspondents to each other, because do not allow real "digital signature". In "asymmetric" methods (RSA, ECC), direct and reverse crypto transformations are performed using different semi-keys that do not have easily traceable links between themselves, allowing one semi-key to calculate another. Therefore, one of the semi-keys is publicly published so that anyone can encrypt a message or verify a digital signature. Only someone who knows the second – secret – semi-key can decrypt such a message or put a signature. Such algorithms, in comparison with "symmetric" ones, are more demanding on computing resources and, therefore, their implementation and use is more expensive. To date, the cryptographic strength of "asymmetric" algorithms is less justified than the cryptographic strength of "symmetric" algorithms. But they work where "classical" crypto schemes are inapplicable: they allow you to implement various ingenious protocols such as digital signature, open key distribution and reliable authentication in a network that is resistant even to complete interception of traffic. However, the use of asymmetric methods has led to a new set of problems. Chief among them is the problem of

obtaining a reliable public key of the addressee [12].

The encryption process consists of a set of rounds-steps, at each step the following actions are performed. The input block is divided in half into the older (L) and younger (R) parts. The value of the encryption function is calculated from the lower part (R) and the round key (k) X=f(R,k). The function used at this step is called the round encryption function. It can be one for all rounds or individual for each round. In the latter case, the encryption functions of different rounds of the same cipher differ, as a rule, only in details. The output block is formed, its high part is equal to the low part of the input block L'=R, and the low part is the result of the bitwise XOR operation (let's denote it (+)) for the high part of the input block and the result of calculating the encryption function R'=L(+)f(R,k) [3,4].

Pretty Good Privacy (PGP) is a software package developed by Philip Zimmerman that provides mail and file encryption. Zimmerman took existing cryptosystems and cryptographic protocols and developed a freeware program for various platforms. It provides message encryption, digital signatures, and email compatibility [5].

The algorithms used for message encryption are RSA for key transfer and IDEA for message encryption itself. Digital signatures are achieved by using RSA for the signature and MD5 for calculating the message digest. PGP uses ZIP compression, and also masks the coordinates and data of the sender, which slightly complicates the process of traffic analysis. Mail compatibility is achieved by using Radix-64 conversion [5,6].

PGP has been verified by a huge number of people, its source code is published on the Internet, but since it uses RSA, the degree of protection of the message depends on the key length (the more the better).

The MIT PGP package version 2.6 and later is completely legal free software for non-commercial use. Viacrypt package version 2.7 and later is a commercial product [4].

To date, the PGP package provides the best degree of data protection for home users when working on open networks. To protect data at the level of organizations, it is advisable to use commercial versions of encryption packages based on asymmetric algorithms (packages based on RSA, ECC algorithms). Maximum security can be obtained by choosing the optimal algorithm for a specific situation [4,12].

Methods for protecting information during transmission over communication channels

Organizational measures are not able to fully prevent unauthorized access attempts, since they apply exclusively to the scale of the organization, do not cover communication channels, and do not involve the use of technical means to combat the threats of interception of information messages. In this regard, along with the use of different priority modes and access control systems, developers of information systems pay attention to various cryptographic methods of information processing [4,5].

A proven method of protecting information from unauthorized access is encryption (cryptography). Encryption is the process of converting open data (plaintext) into encrypted (ciphertext) or encrypted data into open data according to certain rules using certain rules contained in the keys (cipher).

Known cryptographic methods for protecting information can be divided into two classes:

1. information processing by replacing and moving letters, in which the amount of data does not change (encryption);

2. compression of information by replacing individual combinations of letters, words or phrases (coding).

There are certain requirements for encryption algorithms:

- high level of data protection against decryption and possible modification;
- security of information should be based only on the knowledge of the key and not depend on whether the algorithm is known or not (Kirckhoff's rule);
- a small change in the source text or the key should lead to a significant change in the ciphertext (the "collapse" effect);

- the range of key values should exclude the possibility of data decryption by enumeration of key values;
- cost-effective implementation of the algorithm with sufficient speed;
- the cost of decrypting data without knowing the key must exceed the cost of the data [6,7].

In accordance with the reference model of computer networks OSI (Open Systems Interconnection) seven levels of system interaction are distinguished. The OSI model is an abstract model of how computers, applications, and other devices interact in a telecommunications network (fig. 1).

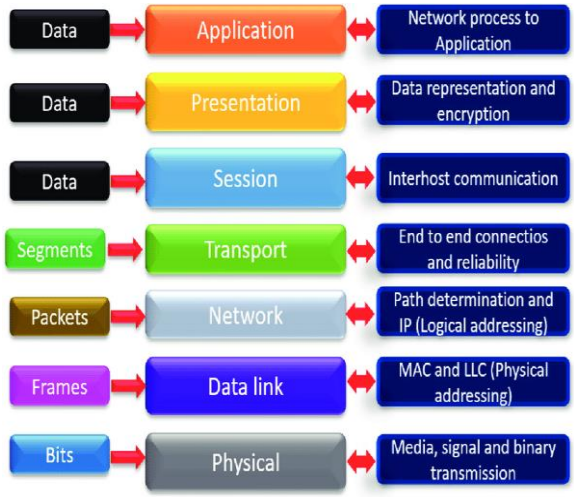

Fig. 1. OSI Model

The top three levels are designed to communicate with the end user, and the bottom four are focused on real-time communication functions [9,10].

With channel encryption, i.e., data encryption at the lower level of the network, the data sent over each communication channel is processed, and each incoming information stream must be decrypted, followed by re-encryption of the outgoing one.

Sending information in the clear over any channel will jeopardize the security of the entire network as a whole. In this regard, the cost of implementing channel encryption in large networks can be quite high. In addition, when using this type of encryption, it will be necessary to protect each node of the computer network through which the transmitted information passes. This is due to the need to protect confidential information, which can only be accessed by certain employees, and for the rest it is necessary to restrict access to this data [10,11].

With end-to-end encryption performed at the upper levels of communication networks, the processing of the transmitted information is carried out at one of the upper levels only in relation to the content of the transmitted message with the addition of service information necessary for routing the information package. After this processing, the generated packet is sent to lower levels for transmission to the destination. With such encryption, there is no need for additional "pair" processing (decryption and encryption) at each intermediate network node, since the information message remains encrypted throughout the entire data transmission route [11].

The disadvantage of this method of protection is the transmission of supplemented service information in unencrypted form, therefore, unauthorized receipt of a number of useful data (for example, about the schedule of communication sessions) is possible. In addition, if end-to-end encryption is used, there may be difficulties in encoding methods due to differences in the communication protocols and interfaces used, depending on the types of computer networks and the elements of these networks [7].

With combined encryption, using the capabilities of both channel and end-to-end encryption, the transmitted information can be best protected, but at the same time, the cost of data protection increases proportionally.

Most cryptographic data protection tools are implemented using specialized hardware devices installed on the transmitting and receiving sides - an encoder and a decryptor, which encrypt and decrypt the transmitted information, respectively. The use of specialized equipment for encryption causes a relatively high cost of implementation, however, there is a certain predominance of hardware compared to software methods. The advantages of hardware solutions are mainly considered, related to the speed of

information processing and to ensuring the physical protection of components. It is believed that hardware is able to more quickly carry out the necessary data processing operations than software is able to implement complex cryptographic algorithms.

Software encryption is the result of implementing a cryptographic algorithm by software. The advantages in using software tools are the possibility of replication by ordinary copying, the relative ease of their modification and use [8,9].

The transformation of information, which results in a change in the amount of memory occupied by the data, is called encoding. Encoding of textual information can actually be carried out using code tables by replacing some characters with others. In this case, a certain compression of the transmitted information packet can also be carried out . If the information is encrypted using a simple substitution, then it could be decrypted by determining the frequencies of occurrence of each letter in the ciphertext and comparing them with the frequencies of the letters of the Russian alphabet. Thus, it is possible to define a substitution alphabet, as a result of which the text is deciphered [9].

An analysis of the encryption methods currently used shows that, despite their fairly widespread use, they are not completely free from shortcomings and leave a certain field for improving and developing new methods for protecting information transmitted over communication channels. Given the dynamics of the development of document management methods, including their formation, it is advisable to integrate the processes of creating and protecting generated documents [10,11].

References

1. *Nichols R.K.* ICSA guide to cryptography / R.K. Nichols. – New York: McGraw Hill, 1999. – Part 2, 4. – 837 p.

2. *Вербіцький О.В.* Вступ до криптографії / О.В. Вербіцький. – Львів: Видавництво науково-технічної літератури, 1998. – 247 с.

3. *Bo Zhu.* Analysis and Design of Authentication and Encryption Algorithms for Secure Cloud Systems: A thesis presented to the University of Waterloo in fulfillment of the thesis requirement for the degree of Doctor of Philosophy in Electrical and Computer Engineering. – Waterloo, Ontario, Canada, 2015. – 163 p.

4. European Payments Council (EPC). Guidelines on cryptographic algorithms usage and key management. EPC342-08 / Version 11.0 / Produced by PSSG / Date issued: 8 March 2022. – 73 p.

5. *Volna E.* Cryptography Based On Neural Network / E. Volna, M. Kotyrba, V. Kocian, M. Janosek // Shaping reality through simulation: 26th European Conference on Modelling and Simulation (Koblenz, Germany, May 29 – June 1 2012) / ECMS 2012 proceedings edited by: K.G. Troitzsch, M. Moehring, U. Lotzmann. European Council for Modeling and Simulation. – P. 386-391.

6. *Howard B.* Information security: Threats and protection mechanisms / B. Howard, O. Paridaens, B. Gamm. – 2003. P. 117-121.

7. *Bellare M.* Introduction to Modern Cryptography / M. Bellare, P. Rogaway. – 2005. – 283 p.

8. *Ashchenko V.V.* Cryptography: an introduction / V.V. Ashchenko. – American Mathematical Soc., 2002. – 229 p.

9. *Schneier B.* Applied Cryptography: Protocols, Algorithms, and Source Code in C / B. Schneier. – Wiley, 2016. – 769 p.

10. *Zimmermann P.* A Proposed Standard Format for RSA Cryptosystems / P. Zimmermann // Advances in Computer Security. – V. III. – Edited by Rein Turn, Artech House, 1998. – 376 p.

11. *Garfinkel S.* Pretty Good Privacy / S. Garfinkel. – O'Reilly & Associates, 1995. – 430 p.

12. *Gasimov V.A.* Information search methods and systems [Textbook] / V.A. Gasimov. – Baku, 2015. – 310 p.

Aliyeva S.Y.

PROTECTION OF INFORMATION DURING DATA TRANSFER IN OPEN NETWORKS

Connecting to open (global) networks such as the Internet significantly increases work efficiency and opens up many new opportunities. At the same time, care should be taken to create a protection system against those who want to use, change or simply destroy information resources. Information security involves protecting the integrity, availability and, if necessary, confidentiality of information and resources used for data entry, storage, processing and transmission. A combination of legislative, organizational and software and technical measures is needed to solve the complex security problem.

This article is devoted to the influence of the characteristics of open networks on the solution of the problems of ensuring the security of data transmission and creating libraries based on the best algorithms. It is shown how the characteristics of data affect the development of modern algorithms for their encryption.

Absolutely secure algorithms have been around for a long time, but have not yet found widespread use in open networks.

Exact equality is only possible if all possible values of the key are equally likely. This is equivalent to the condition that the key bits are equally likely and statistically independent of each other.

Thus, absolutely strong ciphers require the use of a key at least as large as the data being encrypted. Both the sender and the receiver must have this key, meaning it must be delivered to them first, and this requires a secure channel. In addition to a potentially insecure channel for transmitting encrypted data, it is necessary to have a secure channel for transmitting the same key size. This is not always acceptable for economic reasons, so such systems are used only in exceptional cases to protect information of special value. The vast majority of truly encrypted communication systems use algorithms that do not have absolute security and are therefore called imperfect ciphers.

The unavailability of the algorithm does not increase the security of the password; open algorithms are considered standard.

There is no way to get an exact value of cryptanalysis complexity. All estimates are based on tests of cipher resistance to currently known types of cryptanalysis, and there is no guarantee that new analysis methods will not be developed in the near future that significantly reduce labor intensity. The above means that, given the current state of cryptography, the security of absolutely all ciphers, except perfect ones, cannot be substantiated by evidence. Instead, it is empirically justified as a resistance to the types of cryptanalysis known today, but no one can guarantee that a successful type of cryptanalysis will not be invented tomorrow for this particular cipher.

The analysis of currently used encryption methods shows that, despite being used quite widely, they are not completely free of shortcomings and leaves a certain area for improvement and development of new methods of data protection transmitted through communication channels. Taking into account the development dynamics of document management methods, including their formation, it is appropriate to integrate the processes of creation and protection of created documents.

Алієва С.Я.

ЗАХИСТ ІНФОРМАЦІЇ ПІД ЧАС ПЕРЕДАЧІ ДАНИХ У ВІДКРИТИХ МЕРЕЖАХ

Дана робота присвячена проблемі оцінки ризиків у комп'ютерних мережах, які притаманні критичним інфраструктурам. У роботі показано місце оцінки ризиків у

глобальному процесі управління ризиками, а також його Підключення до відкритих (глобальних) мереж, таких як Інтернет, значно підвищує ефективність роботи та відкриває багато нових можливостей. Водночас слід подбати про створення системи захисту від тих, хто хоче використати, змінити чи просто знищити інформаційні ресурси. Інформаційна безпека передбачає захист цілісності, доступності та, якщо необхідно, конфіденційності інформації та ресурсів, які використовуються для введення, зберігання, обробки та передачі даних. Для вирішення складної проблеми безпеки необхідна сукупність законодавчих, організаційних і програмно-технічних заходів.

Дана стаття присвячена впливу характеристик відкритих мереж на вирішення завдань забезпечення безпеки передачі даних і створення бібліотек на основі найкращих алгоритмів. Показано, як характеристики даних впливають на розробку сучасних алгоритмів їх шифрування.

Абсолютно безпечні алгоритми існують вже давно, але ще не знайшли широкого застосування у відкритих мережах.

Точна рівність можлива тільки в тому випадку, якщо всі можливі значення ключа однаково ймовірні. Це еквівалентно умові, що ключові біти однаково вірогідні та статистично незалежні один від одного.

Таким чином, абсолютно надійні шифри вимагають використання ключа, щонайменше такого ж розміру, як дані, що шифруються. І відправник, і одержувач повинні мати цей ключ, тобто він повинен бути доставлений їм спочатку, і для цього потрібен безпечний канал. Крім потенційно незахищеного каналу для передачі зашифрованих даних, необхідно мати захищений канал для передачі такого ж розміру ключа. Це не завжди прийнятно з економічних причин, тому такі системи використовуються лише у виняткових випадках для захисту інформації особливої цінності. Переважна більшість справді зашифрованих систем зв'язку використовують алгоритми, які не мають абсолютної безпеки, і тому їх називають недосконалими шифрами.

Відсутність алгоритму не підвищує безпеку пароля; відкриті алгоритми вважаються стандартними.

Немає способу отримати точне значення складності криптоаналізу. Усі оцінки базуються на тестах на стійкість шифру до відомих на даний момент типів криптоаналізу, і немає гарантії, що в найближчому майбутньому не будуть розроблені нові методи аналізу, які значно зменшать трудомісткість. Вищевикладене означає, що, враховуючи сучасний стан криптографії, безпека абсолютно всіх шифрів, крім ідеальних, не може бути підтверджена доказами. Натомість це емпірично виправдано як опір відомим сьогодні типам криптоаналізу, але ніхто не може гарантувати, що завтра для цього конкретного шифру не буде винайдено успішний тип криптоаналізу.

Аналіз використовуваних в даний час методів шифрування показує, що, незважаючи на досить широке використання, вони не повністю позбавлені недоліків і залишають певну область для вдосконалення і розробки нових методів захисту даних, що передаються по каналах зв'язку. Враховуючи динаміку розвитку методів документообігу, в тому числі їх формування, доцільно інтегрувати процеси створення та захисту створених документів.