ABSTRACT

Kyryl Yu. Shekhanin. Development and analysis of steganographic methods of hiding data in the structure of file systems. – Qualification scientific work is as a manuscript.

Thesis submitted for obtaining the Doctor of Philosophy degree in Information Technologies, Speciality 125 – Cybersecurity. – V. N. Karazin Kharkiv National University, Ministry of Education and Science of Ukraine, Kharkiv, 2021.

The dissertation is devoted to the development and improvement of steganographic methods of hiding information in the structure of the file system by mixing clusters. Development of a model for evaluating these methods and possible uses.

The purpose of the dissertation is to increase the bandwidth of cluster steganosystems while ensuring the necessary resistance to unauthorized detection of hidden information.

The first chapter of the dissertation (*Research of media technologies and properties of file systems*) analyzes the current state of development of technologies of physical media. In particular, technologies such as HDD and SSD are analyzed. The forecast on development of technologies of data carriers is given. Common file systems are also analyzed, their comparative analysis is performed. Based on the results of the analysis, a file system is selected for further research. *The first partial research problem* is solved: conducted research on modern and promising methods of information storage, properties of physical media and types of file systems, analysis of existing methods of steganographic hiding information in the structure of file systems.

The second chapter (*Development of a method to increase the bandwidth of cluster steganosystems and study their properties)* analysis in detail the FAT32 file system as a root system from the family of cluster file systems. Possible properties of the file system structure that help to hide the message in the file system structure are also investigated and analyzed. A mathematical model of the method of hiding information by mixing clusters of cover files is given. *The second partial problem of the research* is solved: development of a method to

increase the bandwidth of cluster steganosystems. *For the first time is obtained* a method of increasing the bandwidth of cluster steganosystems on the basis of taking into account the additional dependence of cluster locations within single cover file of the system.

In the third chapter (*Analysis of methods of hiding information and improving the mathematical model for estimating the basic parameters of cluster steganosystems*) the researched methods on the possible size of the hidden message depending on various initial parameters are analyzed. The formulas by which it is possible to estimate the maximum possible size of the steganogram depending on key parameters of methods are received (number, order and size of cover files). Diagrams are given that clearly show the dependence of the size of the steganogram on the key parameters of the methods.

The level of security of the hidden message before detection is also assessed by analyzing the average level of file system fragmentation and fragmentation of each cover file. The evaluation results are obtained by statistical analysis of the level of fragmentation of computer systems from laboratories of V. N. Karazin Kharkiv National University. According to the results of statistical analysis, a cast of file systems in terms of the level of fragmentation, the most fragmented file types and how to use computer systems. Based on the results of this analysis, data on the possible size of the hidden message are provided, so that the level of fragmentation of the cover files is within the average level of the file system.

This chapter provides an estimate of the time parameters of the methods of hiding the message, performed using the developed program “SteganoFAT”. This estimate depends on the technical properties of the computer system and the physical storage medium, so the equation of the total computational complexity is derived depending on the number of cover files and the number of steganoblocks. The total hiding time depends on the number of read head movements, the number of reads and writes data in clusters. The result of this study are methods of hiding information: ПЗОП, ПчЗОП-I/II/III for the basic method and ПЗОПм, ПчЗОП-I/II/IIIм for a modified method of hiding information, respectively. For each method, an estimate of the computational complexity is theoretically obtained. The most optimal in terms of time and level of fragmentation is the method developed in this

dissertation ПчЗОП-II (alternately reading data from clusters to RAM with sequential ordering of only the clusters involved in the message).

The third partial task of the research is fulfilled: improvement of the mathematical model of estimation of the basic parameters of steganosystems. This is achieved by a comprehensive assessment of bandwidth, assessment of resistance to detection and assessment of computational complexity.

The fourth partial problem of the research is also solved: improvement of the method of hiding information in the structure of cluster steganosystems. The improvement is the rational use of RAM, which has reduced the amount of memory required, which has allowed the implementation of hiding methods on low-resource systems. Another improvement is the special calculation of permutation tables so as to reduce the number of moving clusters. This significantly reduced the time to hide the message.

The second scientifically substantiated result is obtained: the mathematical model of estimation of the basic parameters of cluster steganosystems due to additional consideration of elements of a configuration of a steganosystem that allows to estimate more completely capacity of steganosystem capacity is improved.

And *the third scientifically substantiated result* is obtained: improved method of hiding information in the structure of cluster steganosystems by generating an appropriate set of permutations of clusters, which reduces the time of hiding information.

In the fourth chapter (*Development of software implementation of methods of hiding information*), a software simulation is developed, which clearly demonstrates the principle of methods of hiding information in the structure of the file system by mixing clusters of cover files and empirically evaluate the effectiveness of hiding methods (ПЗОП, ПчЗОП-I/II/III). A full description of the developed software implementation is provided with reference to the public repository, algorithmic features of methods of hiding information are described and fragments of code of functionality of hiding of the message are given.

With the help of the developed program the empirical estimation of computational complexity concerning each method of hiding of the information, each way is received.

According to the results of comparison of theoretically obtained and empirically calculated estimation of computational complexity, it is found that the methods of the modified method in this software implementation have worse computational complexity than theoretically expected. However, an improved algorithm of the modified method is described, which allows to reduce the level of computational complexity.

The fifth partial task of the study is performed: development of software implementation of the proposed methods and conducting experimental research.

The fifth chapter (*Directions for using methods of hiding information*) describes possible systems that use methods to hide information in the file system structure.

Such systems are:

– system of hidden data storage;

– system of hidden data transmission;

– a system for verifying the physical medium of information or files stored on the medium.

For each system, a description is provided with an analysis of the effectiveness of such a system depending on the configuration parameters.

Also provided is a description of common software implementations that implement the above-described systems of methods of hiding information. A comparative analysis is made and an assessment of known software implementations is given as if they also used the methods of hiding information developed in this paper as a component in their algorithms. It is concluded that the described methods allow to significantly expand the functionality of the implemented programs and / or methods can provide an alternative in case it is unable to use internal program tools.

Key words: method of hiding information, steganographic methods, file system, storage device, FAT, statistical analysis, file system structure, bandwidth, computational complexity, simulation software, StarForce.