Serdica J. Computing **2** (2008), 295–312

Serdica
Journal of Computing

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

STUDY OF THE INFORMATION SECURITY OF FILE OBJECTS UNDER INFORMATION ATTACKS WITH A RECORD OF EFFECT OF THE METHODS OF COMPRESSION

Dimitrina Liubomirova Polimirova-Nickolova

Abstract. This report examines important issues pertaining to the different ways of affecting the information security of file objects under information attacks through methods of compression. Accordingly, the report analyzes the three-way relationships which may exist among a selected set of attacks, methods and objects. Thus, a methodology is proposed for evaluation of information security, and a coefficient of information security is created. With respects to this coefficient, using different criteria and methods for evaluation and selection of alternatives, the lowest-risk methods of compression are selected.

ACM Computing Classification System (1998): D.4.6.

Key words: Information Security, Information Attacks, Methods of Compression, File Objects, Risk Assessment.

1. Review of the study. Modern society more than ever requires the use of information flows with substantial sizes. These information flows are usually presented in the form of file objects located in a network TCP/IP environment.

Information flows have been subject to various information attacks, and that has attracted the attention of scientists starting already in the 60s [14]. Chris Rodgers has explored computer and network attacks in a TCP/IP environment featuring viruses, worms, Trojan horses and DoS attacks [10]. Daniel Klein has researched one of the most frequently used attacks for accessing systems or file objects – an attack through a password [7]. Marco de Vivo and David Dittrich have examined the effect of various network attacks [8, 13]. In 2004, attacks on mobile phones became extremely popular, and those have been researched by Martin and Hsiao [9]. World organizations like CERT/CC, SANS and OIS research and analyze attacks and regularly publish related reports and bulletins.

Another group of scientists have been trying to find ways of reducing the size of the file objects by designing different methods of compression. For example, Cokus and Winkowski use methods of compression applicable to XML objects [2]. Butner, Iddan, Meron work on compressing images used in medicine and wireless telecommunications which also have a higher rate of compression [1, 5]. Gilbert and Haffner have conducted studies in the field of compression of complex images and video both with and without loss of information [3, 4].

2. Purpose and tasks of this study. One common strategy for file objects protection could include the use of methods of compression, thus simultaneously achieving a reduction in their size and an increase of their information security as information is presented in the form of exit codes.

According, the **main purpose** is to research and analyze the change in information security of file objects located in a TCP/IP environment and subject to information attacks by recording the effect of the methods of compression.

Therefore, the main tasks resulting from the purpose defined above are as follows:

1) to propose a method for reducing *maximum* three-way relationships among certain attacks, methods and objects to *real* three-way relationships that can exist among them;
2) to propose a methodology for evaluation of the information security of an object under attacks by recording the effect of the applied method of com-

pression. Furthermore, using this methodology, the task is to define the methods of compression achieving the highest values of the coefficient of information security for each object for a respective attack, and for all objects for respective attacks;

3) to propose a procedure for selecting methods of compression with the lowest risk with regard to the coefficient of information security of the respective objects for all attacks;

4) to conduct experiments proving the accuracy of the approach that entails the use of matrix transformations applied to an initially created base of two-way relationships among attacks, methods and objects which are to be transformed into three-way attack-method-object relationships.

3. Definition of sets of attacks, methods and objects.

3.1. Analysis and evaluation of the sets of *maximum* number of attacks, methods and objects. In order to explore the information security of file objects the available information of the known attacks, methods and objects until the moment of exploration is systemized. They form the sets of *maximum* number of attacks ($A_{\max} = \{a_1, a_2, \dots, a_i, \dots, a_n\}$),
methods ($M_{\max} = \{m_1, m_2, \dots, m_j, \dots, m_k\}$) and
objects ($O_{\max} = \{o_1, o_2, \dots, o_f, \dots, o_l\}$).

In the set $A_{\max}$ 89 different attacks are included, divided in 33 main groups. 20 of these are in the category “Malicious software (Malware)” and 13 in the category “Malicious attack (Malattack)”. 59 methods of compression take part in the set $M_{\max}$. They are divided in 9 groups: 5 of them belong to the category “Lossy methods of compression” and 4 to the category “Lossless methods of compression”. 42 file objects representing over 23000 file formats are organized in 10 main groups. 7 of these belong to the category “Directly executable” and 3 to the category “Indirectly executable”. They form the set $O_{\max}$.

3.2. Method of reducing the sets to *potential* number of attacks, methods and objects. After that the sets of *possible* relations attack-object (Ω), method-object (Ω) and attack-object (Ω) are determined. For each pair of relations an expert rating for correspondence is made. A logical processing with result logical 1 is labeled to those pair of relationships between the elements of which a connection exists. Thus the sets of *expert relations* are formed. For each pair of relations an experiment with simulation character is conducted, while for the successfully conducted experiments a result of logical entity is set. In this way the sets of *experimental relations* are formed. Combining the sets of expert relations with those of experimental relations the sets of *possible relations* are formed. Table 1 shows the results for the formed corresponding sets of *expert*, *experimental* and *possible relations* for each pair of relations.

Table 1

Relation	***Maximum number of relations***	***Expert rated***	***Experimental rated***	***Possible relations***
Attack—object	3738	2231/3738	2861/3738	2188/3738
Method—object	2478	588/2478	845/2478	585/2478
Attack—method	5251	3540/5251	3835/5251	3540/5251

The set of *real* relations between attacks, methods and objects (X) represents (1):

$$\mathrm{X} = (\Omega \wedge \Xi) \wedge (\Omega \wedge \Theta) \wedge (\Xi \wedge \Theta) \tag{1}$$

In the set of potential number of attacks ($A_{pot} = \{a_1, a_2, \ldots, a_i, \ldots, a_p\}$) 60 attacks take part from a total of 89. In the set of potential methods ($M_{pot} = \{m_1, m_2, \ldots, m_j, \ldots, m_q\}$) 53 methods take part from a total of 59. In the set of potential number of objects ($O_{pot} = \{o_1, o_2, \ldots, o_f, \ldots, o_r\}$) 30 objects take part from a total of 42.

4. Analysis of the information security of objects, subject to attacks and treated with a method of compression.

4.1. Methodic for evaluation of the information security. When the analysis of the information security is conducted of objects are considered the following restriction conditions: 1) only the sets of potential number of attacks, methods and objects are analyzed; 2) the experiments are conducted under standard user and not corporate or governmental requirements; 3) in order to simplify the calculations the lossy methods of compression are excluded; 4) during the conducted experiments for determining of the coefficient of information security objects with equal (or insignificant differences in) initial sizes (1MB) are used.

4.2. Definition of the coefficient of information security. The information security of an object can be represented as a value depending on several main parameters. In the current analysis the parameters $TIME$ (T) and $SIZE$ (S) are chosen. The parameter $TIME$ represents the evaluation of the time needed for attack to fully complete the planned action toward the object in the cases when the object is not treated with a method of compression and when it is treated with a method of compression. The parameter $SIZE$ represents the evaluation of the size of an object before and after its treatment with a method of compression. Main characteristics are chosen, which influence the general rating of the parameter, toward which they correspond before and after the application of method of compression. For each characteristics a value ($V_{(charact.)}$) is defined. Set are weight coefficients (W) of the different characteristics, through application of the AHP (Analytic Hierarchy Process) method [12]. At the end the total value of the parameter ($V_{(parameter)}$) is defined as (2):

$$V_{(parameter)} = \sum_{i=1}^{n} \left(V_{(charact.i)}.W_i\right) \tag{2}$$

For the parameters $TIME$ and $SIZE$ a coefficient of the information security ($K^{IS(p)}$) is formed for evaluation of the parameter (p) as (3):

$$K^{IS(p)} = \frac{RV_{(p)}}{\max RV_{(p)}} \tag{3}$$

where: $RV_{(p)} = \dfrac{V''_{(p)} - V'_{(p)}}{V'_{(p)}}$ is the average value of the parameter, which show with how many times the value of the corresponding parameter is increased after the application of a method of compression on an object; $V'_{(p)}$ is the value of the information security of the object corresponding to the parameter before the application of method of compression; $V''_{(p)}$ is the value of the information security of the object corresponding to the parameter after its treatment with method of compression; $\max RV_{(p)}$ is the maximum average value corresponding to the parameter, achieved from the same object in one of the others *real* relations, in which it takes part.

The coefficient of the information security of an object (K^{IS}) can be represented as average value of the coefficients of evaluation of the parameters (4):

$$K_z^{IS} = \frac{1}{n}\sum_{p=1}^{n} K^{IS(p)} \tag{4}$$

where $K^{IS(p)}$ represents a coefficient of the information security of an object corresponding to a given parameter, n is the number of tested parameters corresponding to the information security of an object, and z varies in the boundaries of the multiplication of a_p, m_q and o_r.

Fig. 1 a), b), c), d), e), f) shows a graphical interpretation of the obtained values of K^{IS} for some of the most commonly used objects.

4.3. Methods of compression with highest values of the coefficient of the information security. From the obtained results the conclusion which method is most suitable for application on the object in order to obtain maximum information security from a certain attack can be made. Fig. 2 a), b), c), d), e), f) shows a graphical variation of the coefficient of information security for selected objects towards the corresponding attacks, derived from the application of the different methods of compression.

In practice though, one object can be subjected of several attacks simultaneously. For each objects a group of methods can be formed, achieving highest values of K^{IS} towards all attacks, to which it might be exposed. They are used to define the methods with lowest risk for compression in relation to the information security of the objects.

5. Methods of compression with lowest risk in relation to the coefficient of information security.

5.1. Description of a model for selection of alternatives. The model, which is proposed for selection of methods of compression with lowest risk in relation to the coefficient of information security is constructed in five stages.

Stage one from the construction of the model is connected with the definition of the objects which will be explored and the different alternatives, compiling of the different methods of compression, which can be applied to the corresponding object.

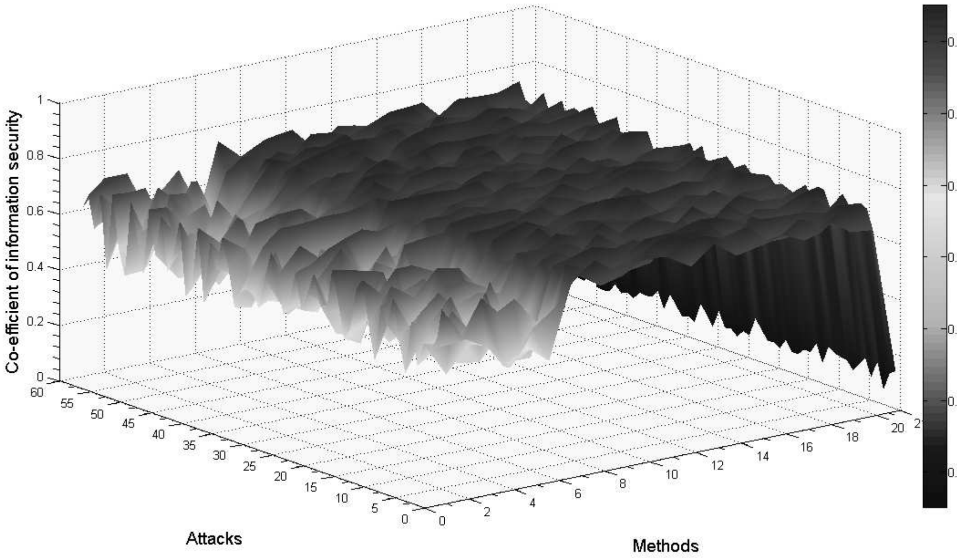

Fig. 1. a) Geographic Information System object

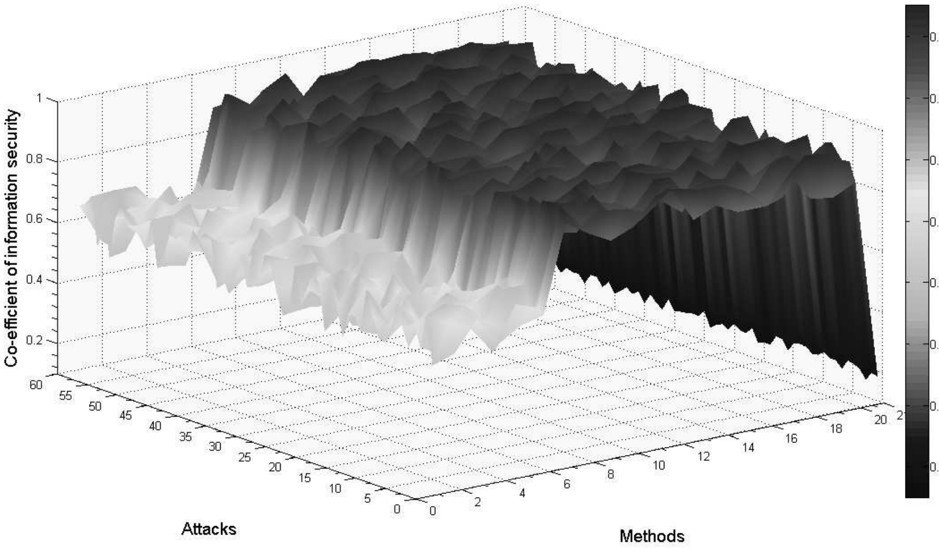

Fig. 1. b) Text/Document object

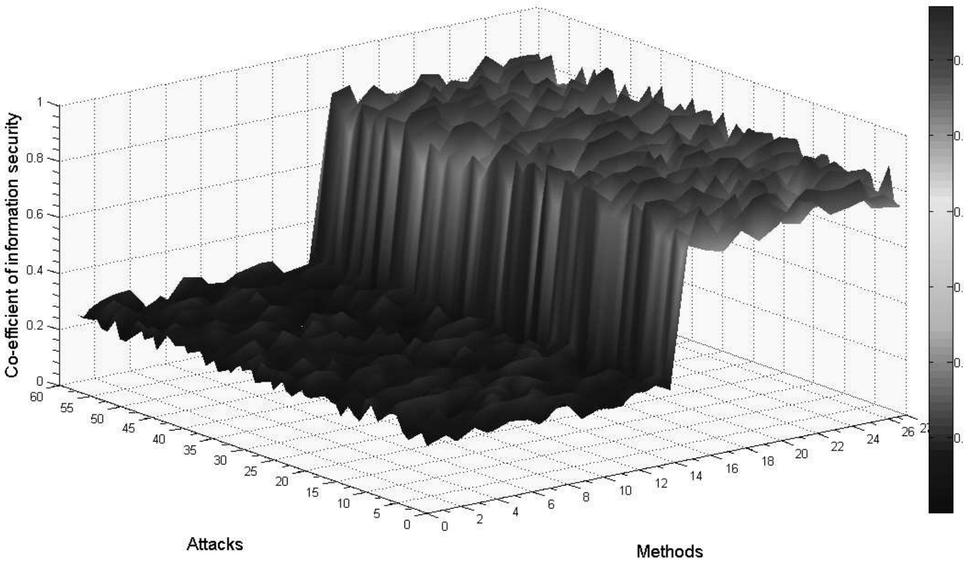

Fig. 1. c) Raster graphic

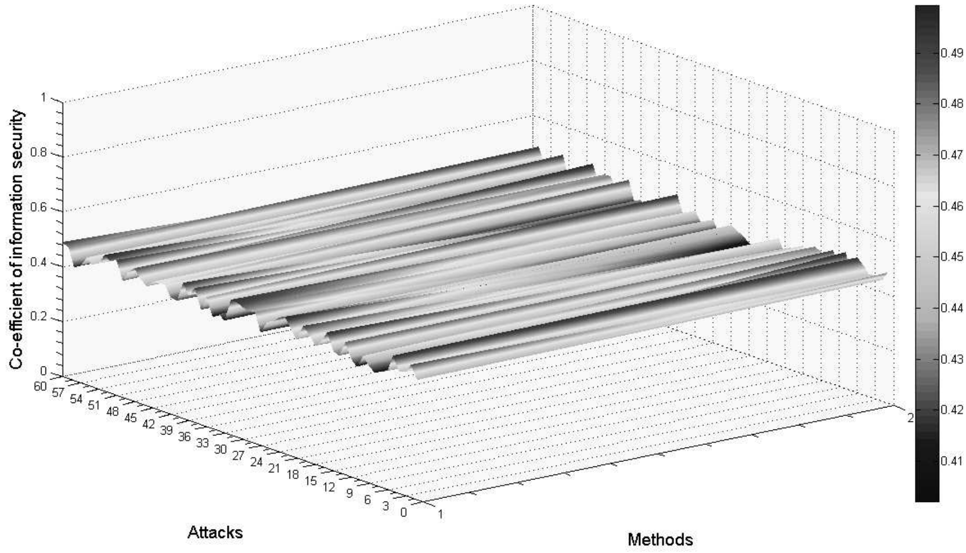

Fig. 1. d) Uncompressed sound

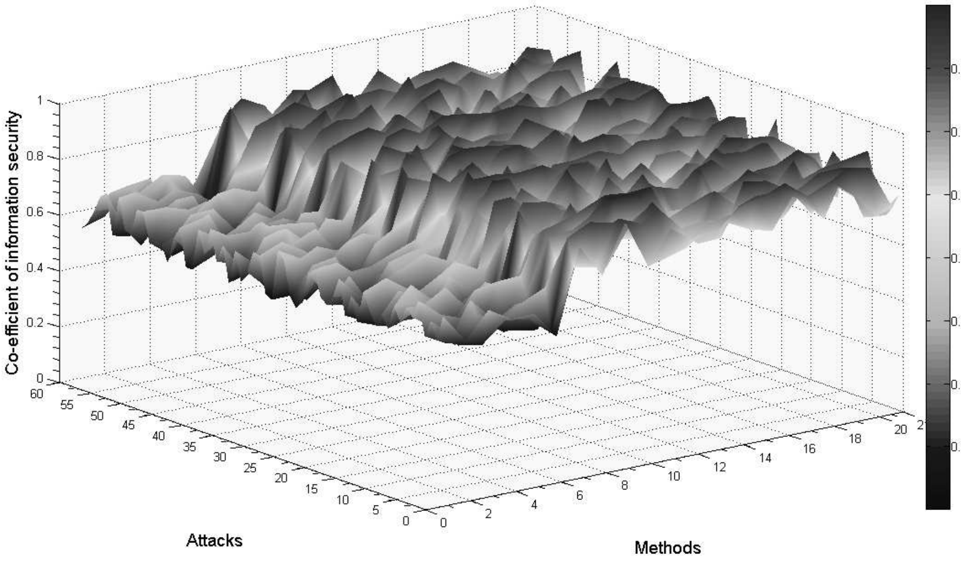

Fig. 1. e) Dynamic web page

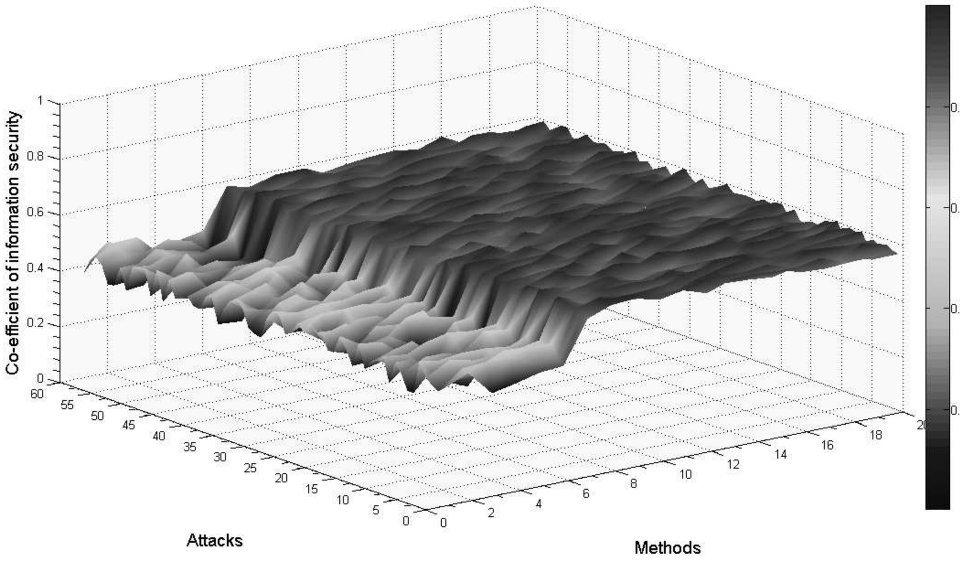

Fig. 1. f) Source code

Fig. 1. Graphic interpretation for determined values of the coefficient of information security for different file objects

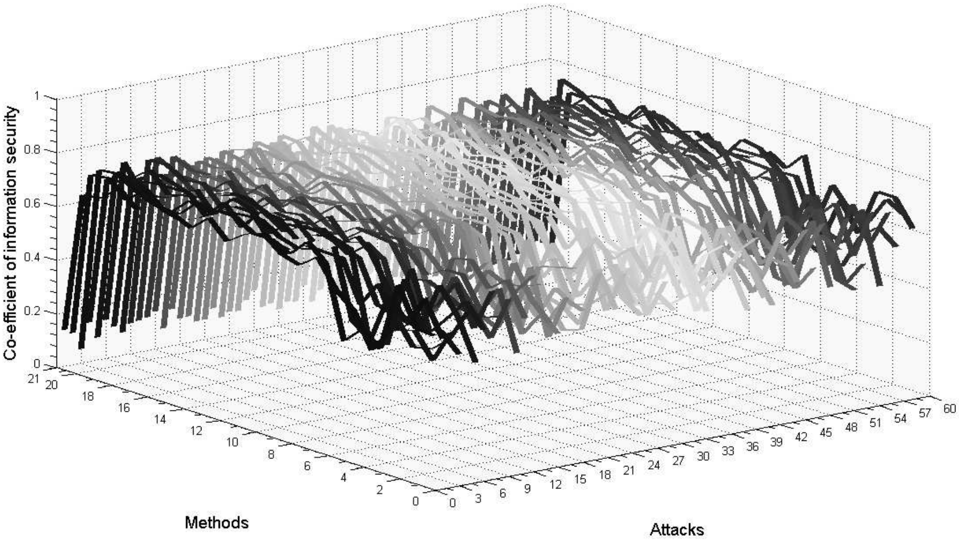

Fig. 2. a) Geographic Information System object

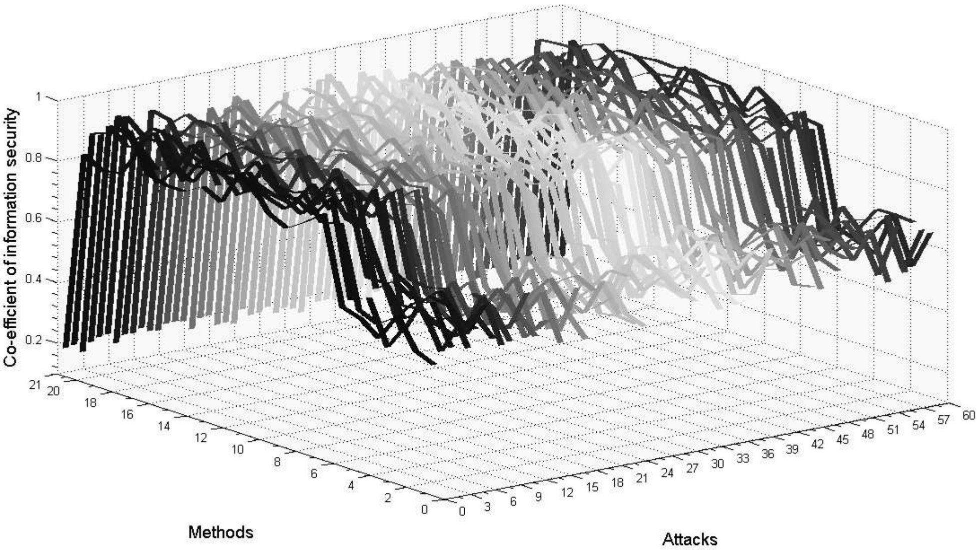

Fig. 2. b) Text/Document object

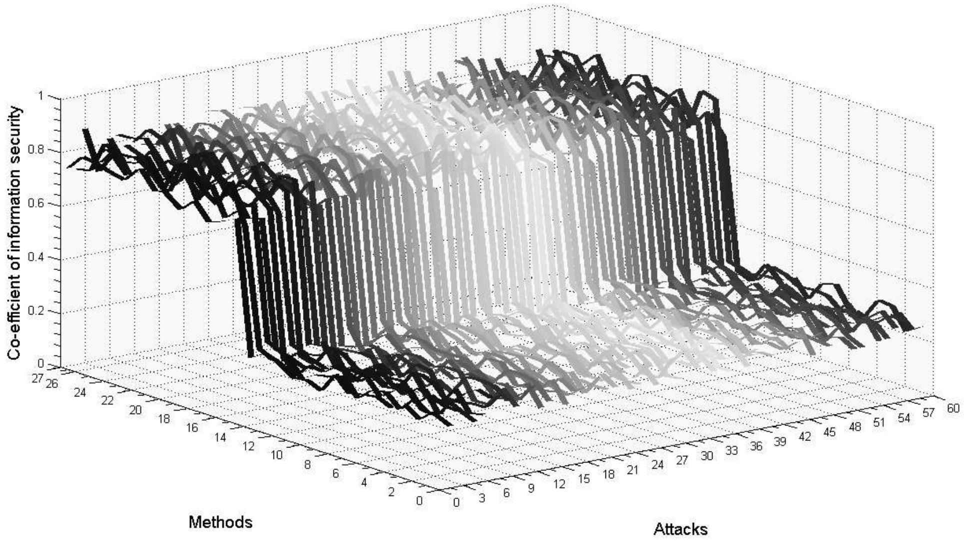

Fig. 2. c) Raster graphic

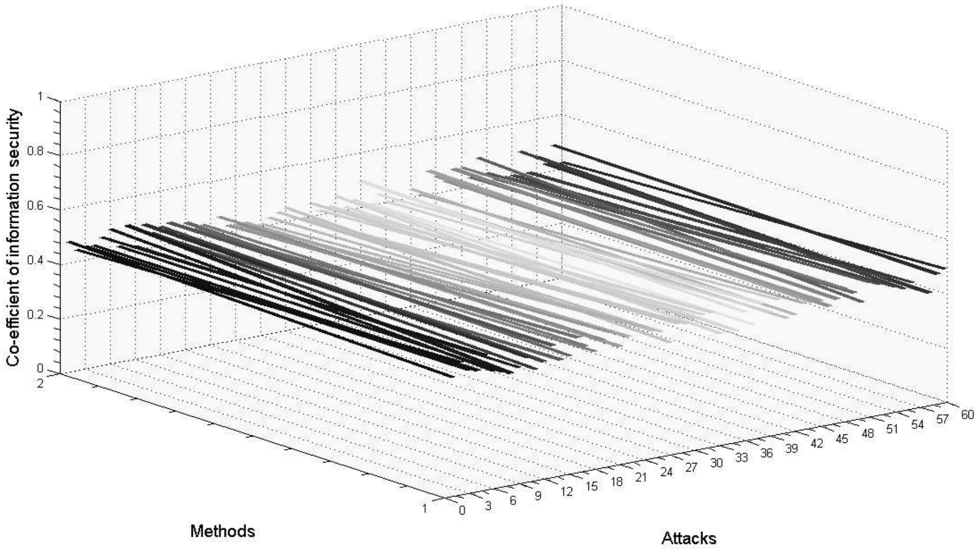

Fig. 2. d) Uncompressed sound

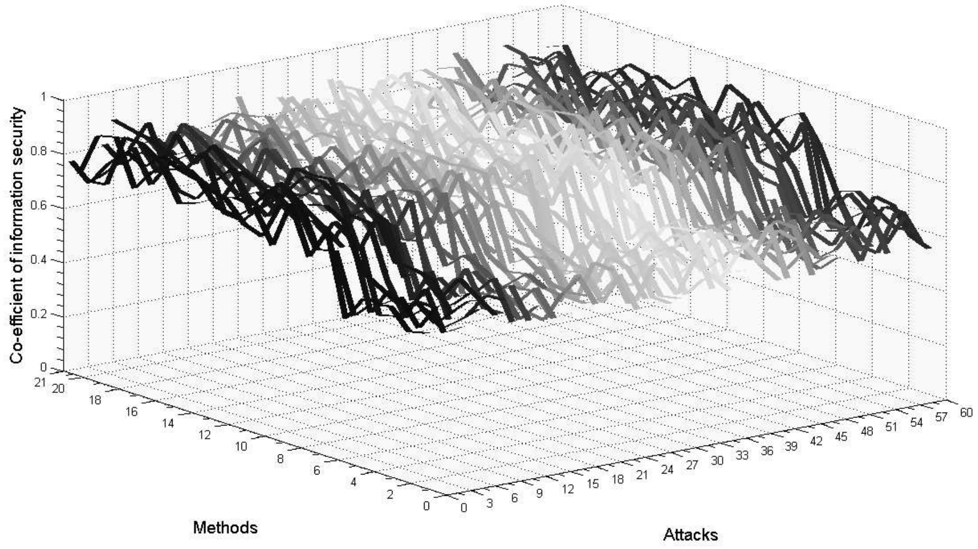

Fig. 2. e) Dynamic web page

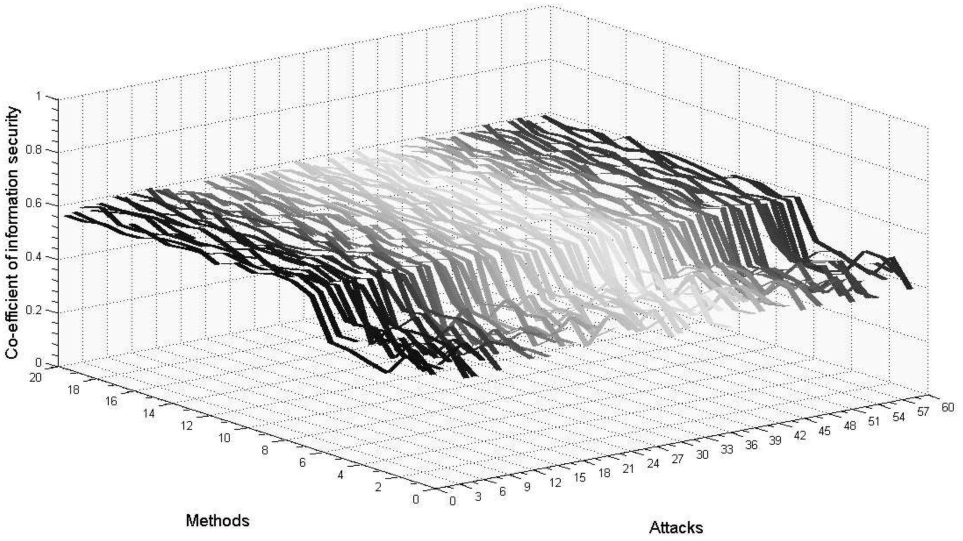

Fig. 2. f) Source code

Fig. 2 Distribution of the coefficient of information security for given object and attack, when a given method of compression is applied

Stage two is connected with the selection of the different characteristics/situations compiling of the different attacks, which can attack the corresponding compressed object.

Stage three is connected with definition of the weight of the different characteristics/situations, e.g. to define for each attack the possibility to attack the chosen compressed object.

On stage four with the help of chosen criteria and methods for selection of alternatives a selection of a alternative with lowest risk (method of compression) is made and the rest alternatives are sorted in descending order in relation to the information security of the object.

The last stage five from the construction of the model is connected with the selection of the best alternative (method of compression) for each object, which is with lowest risk in relation to the information security of the object in question towards all attacks, to which it can be exposed.

5.2. Main database of the model. The main database of the model consists of matrixes, created for each object. They include in the vertical the separate alternatives (compiling of the different methods of compression) and on the horizontal – the different situations in which can fall the object (compiling of the different attacks).

During description of the model, the following terms will be used:

- ✓ risk – when we speak of risk we will have in mind the risk of achieving a lower value of K^{IS} of the object when applying a method of compression as means of protection from different attacks;
- ✓ profit – the profit of application of the method of compression on an object is connected with the achievement of higher value of K^{IS}.

5.3. Calculating the evaluation criteria. For the finding of alternative with lowest risk the following criteria from game theory are used:

- ▷ Maximum of the mathematical expectation for the profit (E_a);
- ▷ Minimum of the mathematical expectation for the risk (E_r);
- ▷ Criteria of Laplace for the profit (L_a);
- ▷ Criteria of Laplace for the risk (L_r);
- ▷ Criteria of Wald (WA);

- ▷ Criteria of Savage (SA);
- ▷ Criteria of pessimism-optimism (H);
- ▷ Criteria of pessimism-optimism of risk (F).

The methods of compression can be sorted according level of preference with the help of the following two methods for multi-criteria evaluation (multi-criteria evaluation methods):

- ▷ Method of the linear combination of formal criteria (S_j);
- ▷ Method of maximum guaranteed result (t_j).

With the use of the matrix and with the help of different criteria from game theory [6] and methods of multi-criteria evaluation [11] the best variant for the person taking the decision can be determined. This variant includes the method of compression, which will have the lowest risk in relation to the coefficient of information security of the objet towards the studied attacks.

Under lowest risk alternative we assume the method of compression, which best satisfies the execution of the target of the model, namely to achieve the best information security of the object towards all attacks through application of a method of compression.

Table 2 shows the obtained results of the calculated criteria for evaluation

Table 2

	GIS object	***Text/ Document***	***Raster graphic***	***Uncompres-sed sound***	***Dynamic web page***	***Source code***
$\boldsymbol{E}_a$	LZX	LZRW1	FELICS	MLP audio	LZ78	LZMW
$\boldsymbol{E}_r$	LZX	LZRW1	FELICS	MLP audio	LZ78	LZMW
$\boldsymbol{L}_a$	LZRW1	LZMW	FELICS	MLP audio	XMill	LZRW4
$\boldsymbol{L}_r$	LZRW1	LZMW	FELICS	MLP audio	XMill	LZRW4
WA	LZSS	LZMW	MLP	MLP audio	XMill	LZRW4
SA	DEFLATE	LZSS	DPCM	MLP audio	LZP	LZMW
H	LZSS	LZX, LZMW	DPCM, MLP	MLP audio, SHN	Xmill, LZY	LZRW4, DEFLATE
F	DEFLATE, LZSS	LZSS, LZMW	DPCM	MLP	LZP	LZMW

of the risk for some of the most commonly used objects. In most cases these are methods of *dictionary* compression.

Table 3 shows a part of the results, obtained for some objects after the calculation of the methods of multi-criteria evaluation. In most cases these are methods of *dictionary* compression.

Table 3

GIS object		***Text/Document***		***Raster graphic***	
S_j	T_j	S_j	T_j	S_j	T_j
LZX	LZRW1	LZRW1	LZY	FELICS	PPPM
LZMW	LZP	LZP	LZP	Progressive FELICS	JPEG-LS
LZRW4	LZMW	LZMW	LZRW1	Block code	CTW
Uncompressed sound		***Dynamic web page***		***Source code***	
S_j	T_j	S_j	T_j	S_j	T_j
MLP audio	MLP audio	LZ78	LZSS	LZMW	LZW
SHN	SHN	LZY	LZY	LZRW4	LZMW
		LZMW	XMill	LZRW1	DEFLATE

6. Software realization and experimental studies.

6.1. Methods and equipment of experimental studies. The methods used for conducting of the experiments include the use of corresponding hardware and software instruments, which is connected with matrixes, matrix transformations, methods for evaluation of the risk, multi-criteria valuation, multi-criteria choice and other.

6.2. Description of the experimental study. For the software realization the program system for scientific studies of the company "*The MathWorks*" is used. The software realization and the experimental studies are conducted in four stages.

Stage I is connected with defining the set of potential attacks, methods and objects.

Stage II is connected with defining the values of the coefficient of information security of the objects.

Stage III is connected with defining the methods of compression with highest values of the coefficient of information security.

Stage IV is connected with defining the methods of compression with the lowest risk in relation to the coefficient of information security.

The results obtained at the end of each stage are used as starting data for the following stage.

7. Conclusion. Contemporary issues are reviewed, connected with the methods to influence the information security of file objects, subjected to information attacks through methods of compression. In this connection, the new moment in the studies is analysis of the three-sided relations, which can exist between a given set, from attacks, methods and objects know up to the time of the study.

A methodology is proposed for evaluation of the information security of objects, exposed to attacks, considering the influence of different methods of compression, which can be applied to such objects. Solving of the problems, connected with the security of file objects, is not a trivial task, hence in the general case this task concerns the problems, which cannot be solved algorithmically. One solution to this problem in user and not corporate or governmental requirements is the application of methods of compression on the objects, through which at the same time is achieved reduction of their size and increase in their security, because the content of the object is presented in the form of codes. In this connection, for every relation attack-method-object is defined a coefficient of information security, depending on two main parameters *TIME* and *SIZE*, which describe accordingly the attack and the object. The parameters are presented as two separate values correlation before and after application of the methods of compression.

Since on one object more than one methods of compression can be applied, different criteria and methods are used for selection of the best method of compression in relation to the information security of the studied objects. Analysis of the received results is made based on which are given recommendations for selection of methods of compression, during the application of which is achieved lowest risk in relation to the information security of file objects, exposed to information attacks.

REFERENCES

[1] Butner S., M. Ghodoussi. Transforming a Surgical Robot for Human Telesurgery. *IEEE Trans. on Robotics and Automation*, **19**, iss. 5 (Oct. 2003), 818–824.

[2] Cokus M., D. Winkowski. XML Sizing and Compression Study For Military Wireless Data, Proceedings of XML Conference & Exposition 2002, Baltimore Convention Center, Baltimore, MD, USA, December 8-13, 2002.

[3] Gilbert J., R. Brodersen. A lossless 2-D image compression technique for synthetic discrete-tone images. In: Proceedings of Data Compression Conference, DCC'98, 30 March – 1 April 1998, 359–368.

[4] Haffner P., Y. LeCun, L. Bottou, P. Howard, P. Vincent, B. Riemers. Color documents on the Web with DjVu. Proc. IEEE Int. Conf. Image Processing, Kobe, Japan, Oct. 1999.

[5] Iddan G., G. Meron, A. Glukhovsky, P. Swain. Wireless Capsule Endoscopy. *Nature*, **405** (25 May 2000), 417.

[6] Keeping E. Introduction to Statistical Inference. Courier Dover Publications, 1995, ISBN 0486685020, 151–173.

[7] Klein D. Foiling the Cracker: A Survey of, and Improvements to, Password Security. In: UNIX Security Workshop II, August 1990.

[8] de Vivo M., O. G. de Vivo, I. Germinal. Internet Security Attacks at the Basic Levels. *ACM SIGOPS Operating Systems Review* **32**, No 2 (1998), 4–15.

[9] Martin T., M. Hsiao, D. Ha, J. Krishnaswami. Denial-of-Service Attacks on Battery-powered Mobile Computers. Proceedings of the Second IEEE International Conference on Pervasive Computing and Communications (PerCom'04), 2004, ISBN 0769520901, 309–318.

[10] Rodgers C. Threats to TCP/IP Network Security, 2001.

[11] Sandblom C.-L., H. Eiselt. Decision Analysis, Location Models, and Scheduling Problems. Springer, 2004, ISBN 3540403388, 19–150.

[12] Vinze A., H. Chen, T. Raghu, D. Zeng, R. Ramesh. National Security. Elsevier, 2007, ISBN 0444519963, 69.

[13] Dittrich D. The DoS Project's "trinoo" Distributed Denial of Service Attack Tool, 1999
http://staff.washington.edu/dittrich/misc/trinoo.analysis, 2008.

[14] http://www.sptimes.com/Hackers/history.hacking.html, 2008.

Dimitrina Polimirova-Nickolova
National Laboratory of Computer Virology
Bulgarian Academy of Sciences
Acad. G. Bonchev Str., Block 8
1113 Sofia, Bulgaria
e-mail: polimira@nlcv.bas.bg

Received October 13, 2008
Final Accepted October 23, 2008