UDC 004.72.056.52

DOI: 10.25140/2411-5363-2018-1(1)-115-130

Vitalii Lytvynov, Nikolai Stoianov, Ihor Skiter, Helen Trunova, Alla Hrebennyk

CORPORATE NETWORKS PROTECTION AGAINST ATTACKS USING CONTENT-ANALYSIS OF GLOBAL INFORMATION SPACE

Urgency of the research. *Further improvement of the security corporate networks in the conditions of massive influence of computer attacks requires an increase in the probability of detection of new computer attacks and a decrease in the recognition time for the signs of known attacks.*

Target setting. *Analysis of the texts of the global information space reduces the time of detection of possible threats.*

Actual scientific researches and issues analysis. *Recent publications about systems of defense from attacks and use of text analysis in detecting threats were considered.*

Uninvestigated parts of general matters defining. *It is necessary to improve the methods of working out the data sets of the body of network packets, content of Internet pages, information of mass media and social networks, which in turn raises the problem of semantic and syntactic processing of natural language texts.*

The research objective. *The aim of the paper is organization collective protection of corporate networks via the introduction of threat monitoring systems, active intelligence activities in the global information space in order to search collect and analyze data about attacks, abnormal behavior, and content of Internet resources.*

The statement of basic materials. *The requirements of security systems to reduce the time of a threat detection lead to the need for active intelligence assessment aimed at continuous monitoring of the surrounding cyberspace that consists of a variety of individual users and organizations' computer networks. The purpose of such monitoring is to determine the characteristics, interests, features of the security policy of a particular corporate network in the global information space. In this context, particular importance attaches to the analysis of text information from both fully and partially open digital sources. A rational solution to this task is the establishment of threat monitoring centers aimed at the organization of collective protection for corporate networks related to them.*

Conclusions. *The proposed method of protection allows both to detect cyber threats in the global information space and to customize their own corporate network security systems in accordance with their characteristic threat vectors.*

Keywords: *systems of defense from attacks; corporate network; intelligence assessment; text representation models; collective protection.*

Fig.: 5. Bibl.: 21.

Introduction. In the modern world, problems related to the use and spread of malicious software, information attacks and other types of cyber threats, which have received the general name "cybercrime" are becoming more and more relevant.

During its development, the information technology sector has accumulated various types of cybercrime, which causes a great damage to both companies and individuals. According to the ISTR report [1] provided by Symantec (one of the leading developers of information security software), the past 2017 was too active for the attackers and was marked by significant incidents in Europe, the United States and the Middle East. The harm that it caused significantly exceeded the figures for 2012, when the total loss inflicted by IT offenders amounted to $ 388 billion.

It is clear that IT specialists were first to realize that there were some problems with the fight against cybercrime. According to the survey, most incidents in the field of information security lead to a loss of payment data (13 %), intellectual property (13 %), customer bases (12 %) and staff information (12 %) [2]. Of course, the problem of improving the methods for analyzing network security and preventing violations in order to fight cybercrime remains relevant. Thus, in today's society, cybersecurity issues have become the defining task of protecting the global information space.

Analysis of recent studies and publications. Traditional approaches to detecting malware are either limited to the use of signatures – byte sequences that identify malicious software, or heuristic algorithms, but these methods are not capable of detecting new attacks in real time [3].

These days, content analysis of text information is used to prevent threats, along with the analysis of the network traffic characteristics, the behavior of corporate networks and their security policy. Existing systems of text analysis and modeling include different kinds of search engines and information-analytical systems. They are capable of solving such tasks as classification of documents by its subject matter, author identification, detection of plagiarism, modeling representations of the knowledge about the subject area and the content of text, classification and filtering of documents by specified queries, and much more [4; 5; 6].

Highlighting the previously unsolved parts of the problem. Enhancement in the effectiveness of security systems and reduced time of threats detection requires a further development in the methods of processing the data arrays of the network packets' body, content of Internet pages and information from mass media and social networks, which raises the problem of semantic and syntactic processing of text, written in natural language.

The purpose of this paper. Applying a wider range of information for assessment of cyber threat's level of danger and creation of collective protection for corporate networks through introduction of threat monitoring systems and active intelligence assessment in the global information space of the Internet.

Main text.

1. Global level of corporate networks security.

The IT community has a considerable amount of experience in solving the tasks of providing information security (cybersecurity) for computer systems. A number of freely distributed and commercial systems of defense from attacks (SDA) was developed and became widely accepted in the field of corporate computer networks building [7–11].

Typical components of SDAs are (Fig. 1):

- a control module designed to configure the system as a whole and issue control commands to its components;
- a sensor block for collecting the output data of network packages, settings, system states, events, messages in system logs, etc.;
- a subsystem of analysis, which identifies the facts of computer attacks and/or abnormal behavior in the information and telecommunication system of the corporate network;
- a storage, which holds the primary information from sensors and signatures, and templates of attacks that are generated by the subsystem of analysis;
- a response module, which is responsible for visualizing the results of the analysis, the generation of warnings, and, in the case of resistance, for the execution of the instructions for selected security methods.

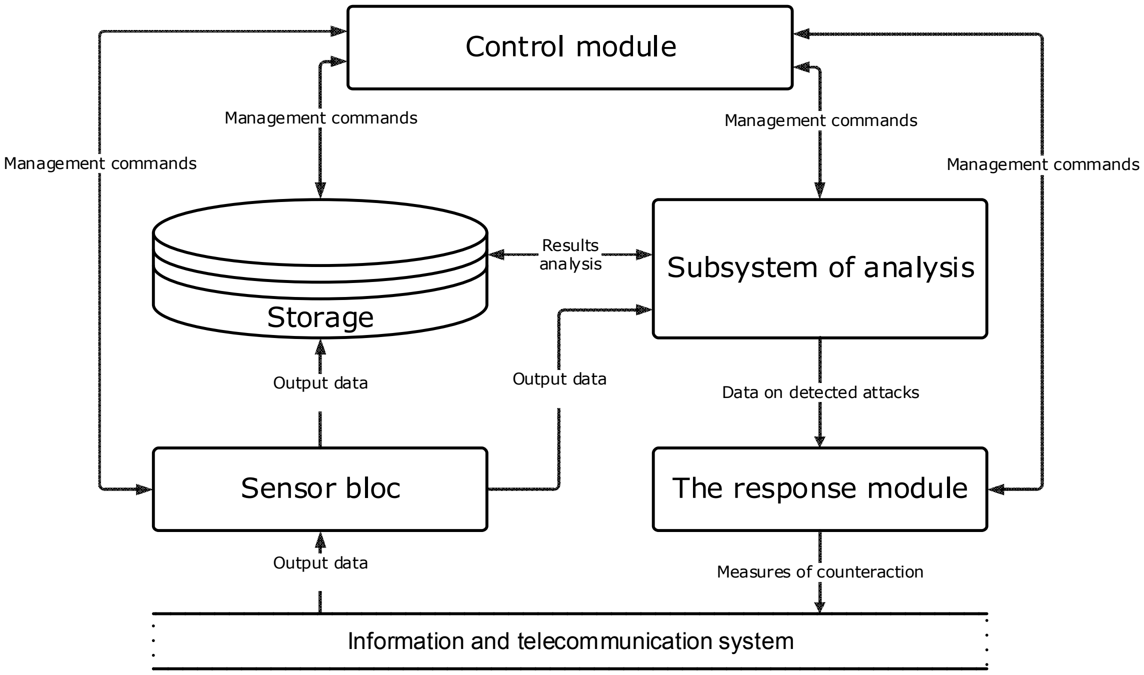


Fig. 1. General architecture of SDA

It is known, that there are two types of basic requirements for SDA:

1. Requirements for detecting non-standard behavior of the computer network and attacks, with aim of minimization of errors of the first and second kind (signaling of non-standard behavior or attack, when it is absent, detection missing of attack or unusual behavior of the network when it takes place);
2. Requirements for detecting attacks in real time.

Earlier, the main efforts of developers were pointed to create effective detection algorithms, satisfying to the first type of requirements. These detection algorithms have used different mathematical basis: statistical methods, methods of automata theory, methods of interacting sequential processes calculus, methods of mathematical logics, neural networks, fuzzy logics, and other formalisms.

Some detection algorithms, in particular algorithms on basis of neural networks have cyberspace-adaptive properties. However, the rapid dynamics of the environment change (the variety of network structures, the variety of types of attacks, etc.) often reduced efforts of designers to zero.

As a rule, the main "bottleneck" of all previous approaches is the violation of time limits adopted for real-time systems. In the case of neural nets detection process adaptation is done by procedure of neural network learning. But it is very time consuming procedure. So, enforcement of adaptive capabilities of detection algorithms leads to slowing of overall detection process.

The way to avoid this dead end situation is following:

- for SDA of corporative information system to use a broader set of analyzed information about environment,that permits to predict behavior of IT system and it's environment;
- to do risk analysis and estimate current or predictable level of danger for corporate nets from known attackers;
- to have time and possibility for corporate system SDA be ready to reflect the most probable attacks.

The first two paragraphs from the above (list), subordinated to the field of intelligence or counterintelligence activity.

According to [12; 13], in the modern world the term political, economic, scientific-technical intelligence means active action, which are aimed at collecting, storage and processing of valuable information, that is closed to outsiders.

A similar definition can be given for counterintelligence activities. Concerning protection corporate computer network from unauthorized access to information, model of attack on computer network always contains step of intelligence activity, as well as protecting the computer network includes counterintelligence activities.

Consider possible approaches to the implementation of the above opportunities.

If the attacker has such information about atacker as: his address and qualification, his preferences regarding the use of certain types of harmful actions, the degree of activity, often gives the opportunity to build both passive and active protection. If the management of passive protection is comes down only to varying their own vulnerabilities, then in contrast to the latter, active protection allows you to carry out counterattacks to the source of the invasion.

In modern SDA, there are three levels of protection from attacks, having access to the processed information:

1. Network layer;
2. Layer of operating system;
3. Application layer (Fig. 2).

Application layer – is responsible for interacting with the end user, layer of OS – is responsible for maintenance of application software and DBMS, network layer – is responsible for the interaction of units of the information and telecommunication system. Each level has its own vulnerabilities.

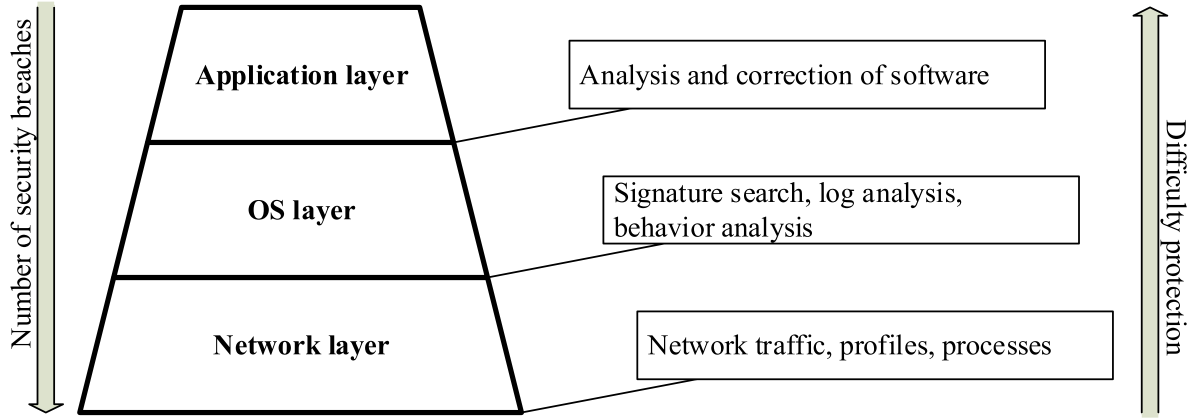
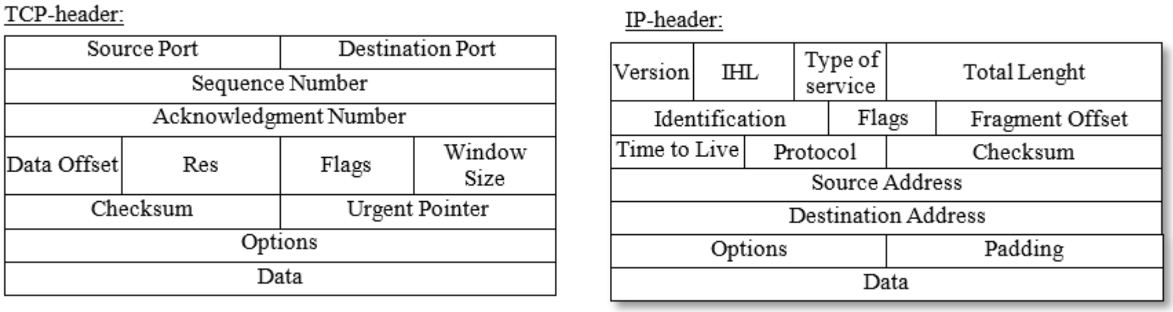

Number of security breaches
Application layer
OS layer
Network layer
Analysis and correction of software
Signature search, log analysis, behavior analysis
Network traffic, profiles, processes
Difficulty protection

Fig. 2. Levels of protection of computer networks by traditional SDA

At the network layer, the bottleneck is the used sharing protocols between the corporate network and outside environment, which tend to be oriented on the package delivery of the information. The packages have a fixed structure. TSP and IP packages can be an illustration of such structures (Fig. 3).


TCP-header:

Source Port		Destination Port	
Sequence Number			
Acknowledgment Number			
Data Offset	Res	Flags	Window Size
Checksum		Urgent Pointer	
Options			
Data			

IP-header:

Version	IHL	Type of service	Total Lenght
Identification		Flags	Fragment Offset
Time to Live	Protocol		Checksum
Source Address			
Destination Address			
Options			Padding
Data			

Fig. 3. Structure of TCP and IP headers

The analysis of the structure of circulating packages in the corporate network is the essence of the analysis at the network layer of protection in SDA. As a rule, the package flags, the port addresses for network nodes, the time intervals between specific events and so on are analyzed here.

The package contains the information about the sender, which is often represented as a DNS-address. This information is definitely of a great value as it can clearly point at the source of the attack. However, the truth of address information about the source of the attack is often questionable, since it can be easily corrected by the sender of the package. For some protocols, such as mail, the address of the attacker may also be obviously stated. However, as in the previous case, the address of the sender can easily be changed.

As a result, there is a need to allocate one more level of realization of the protective methods – the level of the global network.

At this level the information, which is contained in the text documents on web-sites, global network portals, social networks or other legitimate objects of the information space can be analyzed and both the sources of attacks and their information characteristics can be indirectly identified.

The concept of a text document here is multivalued: it is text information from websites and portals, and emails, and program codes that are entered into the computing environment of the victim's computer. In any case, this level is characterized by, on the one hand, methods used in intelligence activities, including business or competitive intelligence [12], and, on the other hand, methods of text processing [14].

In the latter case, studies in this area have significant scientific results and a stated number of tasks of text processing. These include:

- the task of determining the topic of texts in information-analytical and information retrieval systems. The essence of the task is the automatic classification of texts by thematic categories;
- the task of analyzing patents in information systems;
- the task of finding out the author of the text. This is the task of determining the authorship of an unknown text by selecting features of the author's style and comparing of these features with the peculiarities of other documents which authorship is known;
- the task of detecting plagiarism and incorrect borrowing in order to protect copyright. Its solution is to compare the proposed text with the texts of already known authors in order to determine the degree of coincidence;
- the task of automatic annotation and abstracting. It is a brief characteristic of the document, that shows the main content and is an important component of automatic text processing systems. Most existing annotation systems are based on detection of words and vocabulary units, calculation of their weights in the sentence and determining of sentences with the largest total weight. Compiling the abstract is based on these sentences.

In the IT area, tasks of text analysis acquire specific sense. In particular, some of the most popular are:

- the task of analyzing Internet texts and identifying users characteristics;
- Text Mining, including tasks of information impact on the emotional state of social media users;
- the task of analyzing source program code texts, etc.

IT professionals very often have problems with viruses and other malware. Actual threats include spreading spam, phishing, network attacks on enterprise infrastructure, including target and DDoS attacks, where use potentially dangerous software vulnerabilities.

These and other similar examples show a close relationship between cybersecurity systems and word processing systems: when detecting spam, data loss, detecting and tracking potentially dangerous messages, etc.

As it is pointed out in [14], the main source of the text data in the IT industry are posts of users in social networks, blogs, forums, etc.

Processing of the flows of text messages has different purposes:

- tracking of undesirable, potentially harmful messages, identifying the people behind them;
- determination of the emotional dimensions (tone) of the text messages is used during the advertisement campaigns, including the times when it is used during the creation of the contextual advertising;
- configuring of the information search systems interfaces for each specific user.

The relevant task is the authorship identification of small texts, which appears a way more frequent than the task of the authorship identification of the significant size texts [14]. It is mainly due to the widespread of the *instant messenger* programs for message exchange over the Internet, increasing the role of email during the business communication process, vast popularity of the Internet forums and blogs. Users have an opportunity to send messages without completing the registration forms and without inputting any kind of information about themselves; in this case, the registration is more a formality and the address of the sender can be changed easily.

The tasks of the creator identification of the software, including the identification of the malware creator are closely knitted with the tasks of the information security.

This field of the research is actively evolving lately. From one side, it is connected with intellectual property protection, from another, it is connected with the necessity of cyber threats prevention, which arises because of the malware usage. In the latter case, it is hard to overestimate the possible damage, which can be caused to control systems by the key infrastructure, including to the military targets. Because there are new kinds of malware being created all over the globe, there is a necessity of the identification of the malicious code creators and bringing them to justice.

As it was mentioned before, we can get the benefit from the methods and approaches used by competitive intelligence for securing the computer networks, as well as its automatization approaches [12; 15], such as:

1. Objectives classification (like questions, topics, avenues for enquiry).
2. Groups of search bots (in the Ukrainian segment of the Internet using the Ukrainian language, in the international web using the main European languages).
3. Programs for automatic information ranking by classifiers.
4. Employees and units classifiers.
5. Programs for automatic information distribution by consumers.
6. Interactive reference books on information-based topics, collected at the present time.

These tools, as well as the presence in the arsenal of cyber security software for word processing tasks, combined with powerful tools for searching information on the Internet, allow the automated support of a number of competitive intelligence scenarios for the purpose of protecting computer networks.

For example, on fig. 4 presents one of the possible scenarios for determining the address of the attacker on the corporate computer network and possible automated support for it.

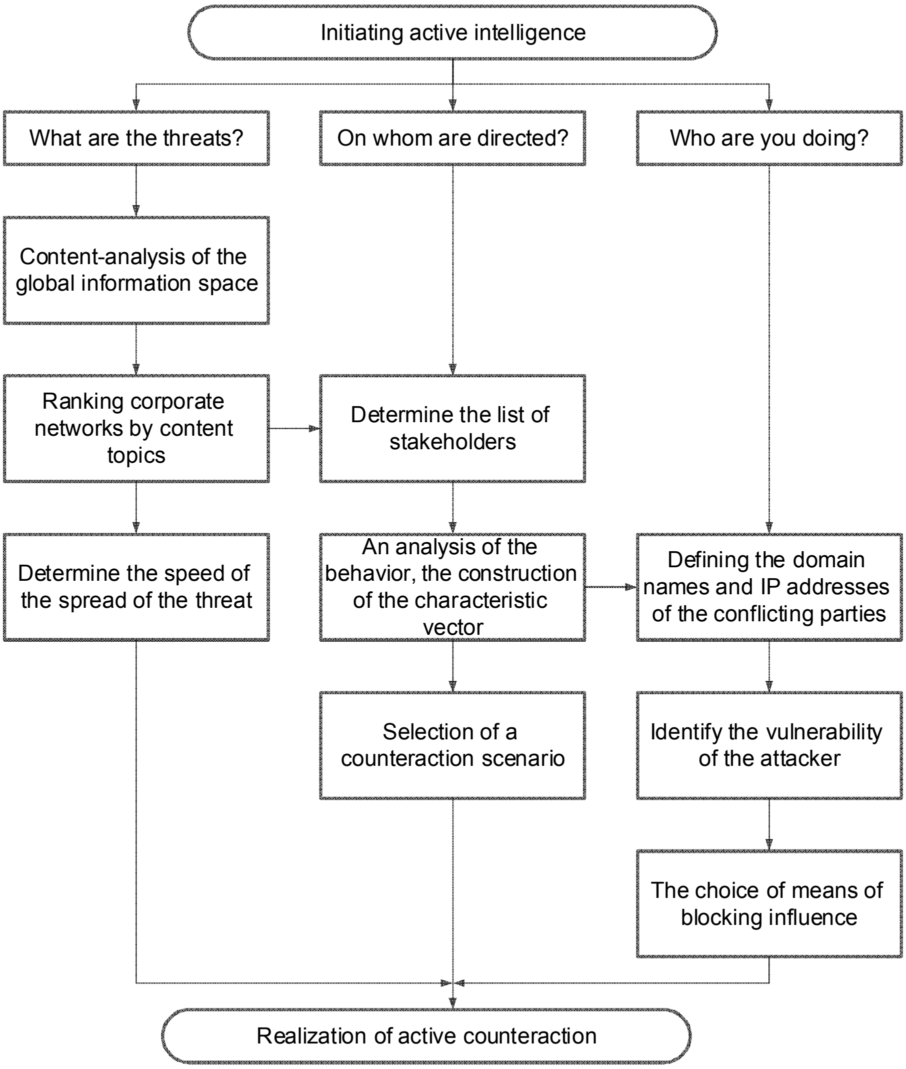

Initiating active intelligence
What are the threats?
On whom are directed?
Who are you doing?
Content-analysis of the global information space
Ranking corporate networks by content topics
Determine the list of stakeholders
Determine the speed of the spread of the threat
An analysis of the behavior, the construction of the characteristic vector
Defining the domain names and IP addresses of the conflicting parties
Selection of a counteraction scenario
Identify the vulnerability of the attacker
The choice of means of blocking influence
Realization of active counteraction

Fig. 4. Implementation of active intelligence on global level of network with using text processing tools

2. The formal models of texts representation.

The basis of all above-mentioned tasks of text processing is the formal models of text representation.

Let us consider that a text is a sequence of characters of an alphabet A, its structure is set by a formal grammar G, which defines its syntactic construction. Furthermore, words and its forms such as objects, subjects, verb constructions, simple sentences, complex sentences, etc. are highlighted. All the sequences of characters, which are described by grammar, form language. Even grammar involvement for text description gives an opportunity to carry out its characterization, since entry of every next element depends on the previous elements. Statistical dependency between elements of the text can be described with a help of informational portrait of the text, which is made on the basis of mutual information between elements of the texts. On that is pointed out in the works of A. Kolmogorov [16] and R. Piotrovsky [17], where the definition of amount of information in one last object relatively to another is being introduced.

The statistical models of the text.

Talking about the models of the texts that were founded on using statistical and informational approach, the view of C. Shannon about the source of information [18] can be used. If we consider the text as a sequence of symbols or other elements, so their occurrence is not random. Any meaningful words or phrases, which form text completely, have statistical structure. In the tasks of analyzing the text its must be accounted.

This approach, which is relied on the views of C. Shannon and fundamental concepts of information theory, was developed in the works of A. Kolmogorov [16] in the probabilistic plan.

It can be used if consider the text as holistic complex system. Any text has a certain meaning that is invariant to the methods of texts presentation. As a complex system text has a semiotic (full of linguistic) nature of informational relations between its subsystems [14].

Let x_i , $i = \overline{1, L}$– is elements of the text, L – is a number of different meanings that element x_i can obtain. Then: $p(\, x_i)$ – is a probability of occurrence of element x_i in the text, $p\big(\, x_i, x_j\big)$ – is a probability of occurrence of a pair of elements x_i and x_j. For well-known texts $T_1, T_2, \ldots, T_m$ the authors $A_1, A_2, \ldots, A_k$ find the value of the selected parameter: the number of inputs of the selected elements in different ways and in their combination then calculate the probability of their appearance in the text, which can be written in the matrix of the probabilities of the collisions of the pair of elements:

$$B_{T_k} = \begin{bmatrix} p(\, x_1, x_1) & \ldots & p(\, x_1, x_L) \\ \ldots & \ldots & \ldots \\ p(\, x_L, x_1) & \ldots & p(\, x_L, x_L) \end{bmatrix}, k = \overline{1, m}.$$

Then, for each pair of elements, a quantitative measure of mutual information between them can be brought into conformity, the results of this are presented in the form of the matrix MI_{T_k} (information portrait of the text T_k) of the mutual information between the elements,

$$MI_{T_k} = \begin{bmatrix} a_{11} & \ldots & a_{1L} \\ \ldots & \ldots & \ldots \\ a_{L1} & \ldots & a_{LL} \end{bmatrix}, k = \overline{1, m}.$$

where $a_{ij} = I\big(x_i, x_j\big)$ denotes mutual information between the elements x_i and x_j, which is calculated by the formula:

$$I\big(x_i, x_j\big) = log_2 \frac{p(\, x_i, x_j)}{p(\, x_i) p(x_j)}, i, j = \overline{1, L}.$$

Informational portraits can be constructed for each text T_k on a plurality of different text elements for each level of the structural-hierarchical model of the text.

In the work [4] the notion of informational portrait is defined as a set of words and phrases selected automatically, which are important for the chosen sample within a framework of general array of documents.

Informational portrait in this case is based on the identification of the relationship of terms and calculation of the weight coefficients of these terms.

There are two algorithms evaluating the relationship between concepts [14]:

1) the algorithm of joint occurrence, which is based on the calculation of the common occurrence of concepts in the same documents (I type);

2) the context proximity algorithm, which is based on the calculation of the correlations of the sets of keywords included in the documents in which the concepts where mentioned (II type).

Different methods of cluster and factor analysis can be used to regularize the concepts and identify their relationships. As a result of their functioning, the relationship tables will take the form of block-diagonal matrices. Thus, the informational portrait of a text can be regarded as its formalized model.

Markov models of texts.

A text is not a random sequence of independent usage of its elements. There are syntactic, semantic, and other dependencies between the elements of the coherent text. An extension of the approach in which symbols are used independently of each other (a probabilistic model of the text) is the Markov model of the generation of text elements [5]. The probability of appearance of an arbitrary element in a text presented in the form of the Markov`s chain depends on the previous element.

Consider some arbitrary text T as a system. Its elementary units (letters, letter combinations, words): s_i, $(i = 1, \dots, N)$. S_q denotes a state of the system at time q. The simplest Markov`s chain is determined by the set of transition probabilities:

$$P[S_q = s_i] = P[S_q = s_i | S_{q-1} = s_{i-1}].$$

With the complication of this model, the probability of occurrence of this element is considered to be dependent on the group of previous elements. Assume that the appearance of some element s_i depends on k previous elements, then:

$$P[S_q = s_i] = P[S_q = s_i | S_{q-1} = s_{i-1}, \dots, S_{q-k} = s_{i-k}].$$

A similar model allows a more complete characterization of the structure of the text.

Relational Model of Text.

Much of the text processing literature a formalized model of text was seen as $\langle E, R\rangle$ pair, where E – set of essence that establish a construction of the text, R – finitary relations which are usually verb form in the text. Based on the model ontologies are built [19] which comprise the description of subject areas. The latter sometimes given as a way of presenting knowledge that enshrined in the text.

In the IT sector practice of using relational model of text is quite extensive: from designing applications to the use of information search mechanisms.

Logical and linguistic model of text.

The logic-linguistic model of the text is widely used in a mathematical linguistics [6, 20]. It allows to present arbitrary sentences as the conjunction of atomic predicates, each of which describes the indivisible content of the sentence:

$$L^S = \wedge_{p \in P^S} \wedge_{h \in H_p^S} L_p^S(h), \tag{1}$$

$$L_p^S(h) = \wedge_{x \in X_p^S(h)} \wedge_{g \in G_p^S(x,h)} L_p^S(x, g, h), \tag{2}$$

$$L_p^S(x, g, h) = \wedge_{y \in Y_p^S(x,g,h)} \wedge_{q \in Q_p^S(x,g,y,h)} L_p^S(x, g, y, q, h), \tag{3}$$

$$L_p^S(x, g, y, q, h) = \wedge_{z \in Z_p^S(x,g,y,q,h)} \wedge_{r \in R_p^S(x,g,y,q,z,h)} L_p^S(x, g, y, q, z, r, h), \tag{4}$$

where S – sentence of natural language;

p – relation that connects actors, objects and subjects of relations in the sentence that connects actors, objects and items of relations in the sentence S, $p \in P^S$ – set of relations included in the sentence S;

h – characteristic of the p-th sentence S relation, $h \in H_p^S$ – the set of characteristics of the p-th relation in sentence S;

$L_p^S(h)$ – predicate that describes p-th relation to the characteristic h and connects actors, objects and items of relation p in sentence S;

x – sentence subject S, $x \in X_p^S(h)$ – set of entities associated with the objects of sentence S by p-th relation that has a characteristic h;

g – characterization of the subject x of the sentence S, $g \in G_p^S(x,h)$ – set of characteristics of the subject $x \in X_p^S(h)$;

$L_p^S(x,g,h)$ – predicate that describes the p-th relation with the characteristic h between the subject $x \in X_p^S(h)$ with the characteristic $g \in G_p^S(x,h)$, the objects and items of the p-th relation in sentence S;

y – sentence object S, $y \in Y_p^S(x,g,h)$ – set of entities associated with the objects of sentence S by p-th relation that has a characteristic h;

q – characteristic of the object y of the sentence S, $q \in Q_p^S(x,g,y,h)$ – set of characteristics of the object $y \in Y_p^S(x,g,h)$;

$L_p^S(x,g,y,q,h)$ – predicate that describes the p-th relation with the characteristic h between the subject $x \in X_p^S(h)$ with the characteristic $g \in G_p^S(x,h)$, the objects $y \in Y_p^S(x,g,h)$ with the characteristic $q \in Q_p^S(x,g,y,h)$ and objects of the p-th relation in sentence S;

z – subject of the p-th relation of the sentence S, $z \in Z_p^S(x,g,y,q,h)$ is the set of objects of the p-th relation, which has the characteristic h, between the subject $x \in X_p^S(h)$ with the characteristic $g \in G_p^S(x,h)$ and the object $y \in Y_p^S(x,g,h)$ with the characteristic $q \in Q_p^S(x,g,y,h)$;

r – characteristic of the subject of the p-th sentence relation S, $r \in R_p^S(x,g,y,q,z,h)$ – set of characteristics of an object $z \in Z_p^S(x,g,y,q,h)$;

$L_p^S(x,g,y,q,z,r,h)$ – simple, atomic predicate that describe a sentence part that has a finished content and describes in the sentence S the p-th relation with the h-th characteristic between the subject $x \in X_p^S(h)$ with the characteristic $\in G_p^S(x,h)$ and the object $y \in Y_p^S(x,g,h)$ with the characteristic $q \in Q_p^S(x,g,y,h)$, whose subject $z \in Z_p^S(x,g,y,q,h)$has the characteristic $r \in R_p^S(x,g,y,q,z,h)$.

The logic-linguistic model L^S of sentence S is represented by the set of formulas (1-4) presented above and is formally described by the sequence of the eight conjunctions included in these formulas. The transition from the general formula L^S to the predicate $L_p^S(x,g,y,q,z,r,h)$ is a decomposition of the problem of the formal description of the arbitrary sentence of the natural language and reflects a systematic approach to its solution. Therefore, the complex expression L^S is true if and only if all elementary predicates of the type $L_p^S(x,g,y,q,z,r,h)$ are included.

Multidimensional text model.

Every text object can be set with a set of some values. Sign selection depends on the processed texts, aims and tasks of the data analysis and other factors. The character of the signs also can be different, qualitative and quantitative, binary (dichotomous), ordinal, etc. However, in any case their complex can be treated as appropriate -dimensional space of signs, and given objects as points of this space. In some tasks, including text information analysis tasks, data is often presented by not the separate signs values, but with probability values of some variable $\rho\,(x_i, x_j\,)$, which characterizes objects pairwise mutual accordancex_i i x_j. Depending on the aims of tasks the degree of similarity or difference is examined, in last case such description denotes distance between objects. Anyway when solving data analysis problems geometrical closeness of two or more points in this -dimensional space means the closeness of corresponding objects, i.e. their homogeneity. The separate classes (clusters) of objects will be represented by coherent areas in this space.

As an example, it is possible to point the next possible signs of every level.

For the level of letters as signs can come forward: frequencies of separate letters appearance, frequencies of separate syllables and signs appearance, frequencies of *n*-gram subsequences of characters from text appearance. For the level of words: frequencies of appearance of separate words, word-parts, bases of words or a few words.

For the level of sentences: frequencies of appearance of sentences with the fixed amount of words, with a certain grammatical construction, using special turns, etc.

In the semantic representation of the text, the value of different attributes as well can be defined at all levels of the semantic hierarchy. Then a collection of documents can be presented in the form of a matrix “Object- sign” $KT = [x_{ij}]$, in which lines correspond to texts ($i = \overline{1,m}$), columns - to signs ($j = \overline{1,G}$), and matrix elements – to the value of sign for each text. . Matrix “Term- document” is formalized by an expression, which is a separate case of transposed matrix “Object- sign”.

To reduce the dimension of the matrix “Text-sign” and detection the most informative features can be used singular decomposition of the matrix (SVD – singular value decomposition). An arbitrary matrix can be represented as:

$M = UWV^T$,

where U i V^T – are orthogonal matrices,

W – diagonal matrix, in addition, its elements are sorted in descending order. Elements of the matrix W – are singular numbers.

Columns and rows of matrices U i V^T, which correspond to a small singular numbers, make the smallest contribution to the final text, so their exclusion will allow to reduce the dimension of the matrix M without significant losses for further calculations [21]. Large singular numbers are main information characteristics, others contain random noise.

When using methods as analysis of main components, factor and discriminatory analysis and others in classical multidimensional data analysis, the “Object-sign” matrix is converted into covariance (correlation) matrix. In this case, the covariance matrix is a square matrix of the “sign-sign” type and it characterizes the degree of proximity (similarity) of signs. However, in practice, to describe text objects is often used representation form of an objects proximity matrix (matrix of “object-object” type).

The correlation matrix "object-object" defines the degree of similarity of the objects, and its elements are determined by the formula:

$$r_{ik} = \frac{\sum_{j=1}^{M}(x_{ij}-\bar{x_\imath})(x_{kj}-\overline{x_k})}{\sqrt{\sum_{j=1}^{M}(x_{ij}-\bar{x_\imath})^2 \sum_{j=1}^{M}(x_{kj}-\overline{x_k})^2}},$$

where $\bar{x} = \frac{1}{M}\sum_{j=1}^{M} x_j$ – the average value.

Formulas are used to calculate the coefficients of the rank correlation when not quantitative values of signs are considered. At the same time, using the developed methods of data multidimensional analysis, it is necessary to take into account the features of the text as a real object and it is essential to consider the process of text structures formation, when compiling models and presenting texts in the form of a multidimensional object.

3. Evaluation of cyberspace from the perspective of threats to corporate computer networks.

Sure, active intelligence of cyberspace in the interests of cyber security of corporate computer networks needs to calculate some threat indicators. For corporate computer networks these indicators can be considered as a vector of threats from different attacks:

$$R(t) = \langle r_1(t), r_2(t), \dots, r_n(t) \rangle,$$

where $r_i(t) = P_i(t) * C_i$ - risk of i-type attack during t-time,

$P_i(t)$ – corporate network's probability of being attacked by i-type attack during t-time,

C_i – cost of lost cause of i-type attacks.

Calculations of risks from various attacks require the identification of sources of attacks on indirect grounds, determining their inclinations to attacks or undesirable influences of one kind or another, determining the characteristics of attack activity, calculating predictive activity indicators based on time series analysis, and the like.

The ordering of the elements of this vector in descending risk values is reduced to the construction of the vector $R^*(t)$, the first elements of which indicate the attacks, which should strengthen the protection of the computer network.

This protection becomes possible or by configuring the corporate network SDA to prepare the activation of attack detection algorithms in accordance with the vector $R^*(t)$, or by eliminating the vulnerabilities that use this type of attack. Given the temporary limitations of the attack detection process, such actions should be performed based on predictions of the activity of potential attack sources, the detection of which is the task of the global network security level of the corporate network.

4. Collective protection of corporate networks against computer attacks.

As can be seen from the previous arguments, the task of text processing and the task of assessing cyber threats indicators for corporate networks, inherent to the global network level, are complex resource-intensive tasks.

Given the temporary requirements for the SDA, it can be assumed that including them in the latter will entail a slowdown in the performance of basic functions and an unjustified increase in resource consumption. At the same time, in our opinion, assigning functions of the globally-lingual level of protection of the corporate network to the functions of a separate computer complex that manages this level of protection of several corporate networks and determines the threat indicators for each of them is a promising solution. We may call this complex as System Monitoring Unit (SMU).

In addition to the parallelism in performing certain functions of SMU and SDA, this solution allows for the collective protection of subordinate corporate computer networks against computer attacks. The essence of this protection is to conduct self-diagnostics of corporate computer networks with the help of SDA, exchange of information about attacks and non-standard behavior with partners, about interference in work. Here you can solve the problem of determining the speed of the spread of external interventions, the coordination of the parameters of the SDA, including the coordination of efforts to analyze unknown invasions.

The structure of SMU complex is shown in Fig. 5.

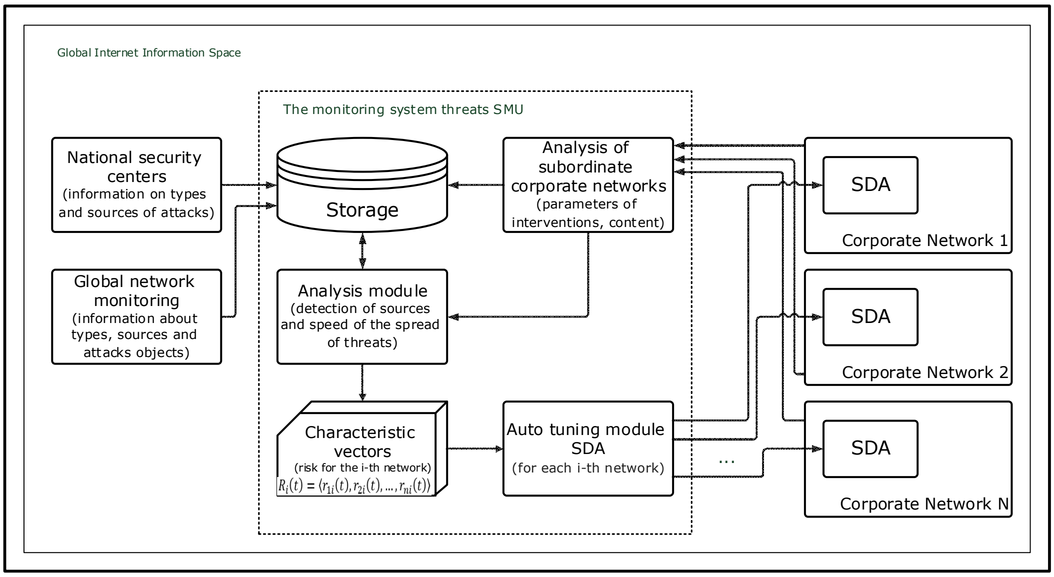

Global Internet Information Space
The monitoring system threats SMU
National security centers (information on types and sources of attacks)
Global network monitoring (information about types, sources and attacks objects)
Storage
Analysis of subordinate corporate networks (parameters of interventions, content)
Analysis module (detection of sources and speed of the spread of threats)
Characteristic vectors (risk for the i-th network)
$R_i(t) = \langle r_{1i}(t), r_{2i}(t), \ldots, r_{ni}(t) \rangle$
Auto tuning module SDA (for each i-th network)
...
SDA
Corporate Network 1
SDA
Corporate Network 2
SDA
Corporate Network N

Fig. 5. Architecture of SMU

In our opinion, the rational use of the proposed complex is to support the activities of the regional cybersecurity center, which is designed not only to perform the functions of operative protection of wards of corporate networks, but also to support their audit.

Conclusions and suggestions. Further improvement of the security and stability in functioning of the information and telecommunication systems of corporate networks in the conditions of massive influence of computer attacks requires an increase in the probability of detection of new computer attacks and a decrease in the recognition time for the signs of known attacks.

To solve this problem, it is not enough to use only traditional methods that utilize identification characteristics of network traffic and information about the work of corporate networks and security devices. The processing of data sets of the body of network packages, content of Internet pages, information from mass media and social networks is very valuable in this area.

Processing, careful analysis and synthesis of information collected from Internet resources is made using content and/or rapid analysis methods, bibliometric and/or cluster analysis, as well as expert and/or situational methods.

However, a tight time limit for the search, collection, extraction and processing of information circulating in the global information space of the Internet, its accumulation, classification by certain attributes, further analysis, synthesis, compilation and making it accessible to the concerned users, as well as transformation into synthesized conclusions and recommendations necessitates some arrangements. First, the automation of all measures in the complex of risks monitoring system associated with these processes. Second, the configuration of SDAs subordinate to the SMUs of corporate networks according to their risk vectors.

The development of a corporate networks protection model with a collective SMU defense module, methods for detecting and identifying computer attacks with help of content analysis of the global information space and the architecture of SDA, related to it, will provide a basis for the synthesis of a reliable and high-performance adaptive cyber threats detection systems and will shorten the detection time of the computer attacks of the new generation.

References

1. Internet Security Threat Report. (2017). *www.symantec.com*. Retrieved from https://www.symantec.com/security-center/threat-report.

2. Chekunov, I. G. (2012). Sovremennye kiberugrozy. Ugolovno-pravovaya i kriminologicheskaya kvalifikatsiya kiberprestuplenii [Modern cyber threats. Criminally-legal and criminological expertise cybercrime]. *Pravo i kiberbezopasnost – Law and Cybersecurity*, 1, 9-22 [in Russian].

3. Uwe, A. & Dasgupta, D. (2005). *Artificial immune systems. In: Introductory Tutorials in Optimisation, Decision Support and Search Methodology (E. Burke and G. Kendall Eds.).* Retrieved from http://eprints.nottingham.ac.uk/336/1/05intros_ais_tutorial.pdf [in English].

4. Surkova, A. S. (2014). Identifikatsiia avtorstva tekstov na osnove informatsionnykh portretov [Identification of authorship of texts based on information portraits]. *Vestnik Nizhegorodskogo universiteta im. N. I. Lobachevskogo – Bulletin of the Nizhny Novgorod University named after. N. I. Lobachevsky,* 3 (1), 145–149 [in Russian].

5. Hmelev, D. V. (2000). Raspoznavanie avtora teksta s ispolzovaniem tsepey A.A. Markova [Recognition of the author of the text using chains A. A. Markov]. *Vestnik MGU. Seriia: Filolohiia – Bulletin of the Moscow State University. Series: Philology*, 2, 115–126 [in Russian].

6. Vavilenkova, A. I. (2015). Porivnialnyi analiz rechen pryrodnoi movy za zmistom [Comparative analysis of sentences of natural language in content]. *Matematychni mashyny i systemy – Mathematical Machines and Systems*, *2*, 97-103 [in Ukrainian].

7. Gamayunov, D. Ju. (2007). Obnaruzhenie kompyuternykh atak na osnove analiza povedeniya setevykh obektov [Detection of computer attacks on the basis of the analysis of the behavior of network objects]. *Candidate's thesis* [in Russian].

8. Chi, S.-D., Park, J. S., Jung, K.-C. & Lee, J.-S. (2001). Network security modeling and cyber-attack simulation methodology. Lecture Notes in Computer Science. *Springer-Verlag, 2119.*

9. Kotenko, I. V., Stepashkin, M. V. & Bogdanov, V. S. (2006). Arkhitektury i modeli komponentov aktivnogo analiza zashchishchennosti na osnove imitatsii deistvii zloumyshlennikov [Architectures and models of active security analysis components based on simulated actions of intruders]. *Problemy informatsionnoi bezopasnosti. Kompjuternye sistemy – Problems of information security. Computer systems,* 2, 7–24 [in Russian].

10. Kotenko, I. V., Stepashkin, M. V. (2005). Analyzing Vulnerabilities and Measuring Security Level at Design and Exploitation Stages of Computer Network Life Cycle. Lecture Notes in Computer Science. *Springer-Verlag, 3685.*

11. Lukatskiy, A. V. (2001). *Obnaruzhenie atak [Detection of attacks].* St. Petersburg: BHV-Peterburg [in Russian].

12. Prilukov, M. V. (2006). Rol delovoi (konkurentnoy) razvedki v obespechenii natsionalnoi bezopasnosti i politicheskoi stabilnosti v Rossiiskoy Federatsii [The role of business (competitive) intelligence in ensuring national security and political stability in the Russian Federation]. *Candidate's thesis* [in Russian].

13. Buriachok, V. L. (2013). Metodolohiia formuvannia derzhavnoi systemy kibernetychnoi bezpeky [Methodology of forming a state system of cybernetic security]. *Doctor's thesis* [in Ukrainian].

14. Surkova, A. S. (2016). Kontseptualnyi analiz, printsipy modelirovaniia i optimizatsiia algoritmov sinteza tekstovykh struktur [Conceptual analysis, principles of modeling and optimization of algorithms for the synthesis of text structures]. *Doctor's thesis*. Nizhniy Novhorod [in Russian].

15. Dodonov, V. O. (2017). Informatsiini tekhnolohii analizu ta vyiavlennia informatsiinoho vplyvu v sotsialnykh merezhakh na osnovi multyahentnykh modelei rozpovsiudzhennia informatsii [Information technologies for analyzing and detecting information influence in social networks on the basis of multiagent information dissemination models]. *Candidate's thesis*. Kyiv [in Ukrainian].

16. Kolmohorov, A. N. (1991). Tri podhoda k opredeleniju ponjatija «Kolichestvo informacii» [Three approaches to defining the concept of «Quantity of Information»]. *Novoe v zhizni, nauke, tehnike. Seriia «Matematika, kibernetika» – New in life, science, technology. Series «Mathematics, cybernetics»*, 1, 24–29 [in Russian].

17. Piotrovskiy, R. H. (1975). *Tekst, mashina, chelovek [Text, machine, man*]. Leninhrad: Nauka [in Russian].

18. Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical,* (Vols. 27), 379–423 [in English].

19. Web Ontology Language (n.d.). *www.w3.org*. Retrieved from https://www.w3.org/2001/sw/wiki/OWL [in English].

20. Vavilenkova, A. I. (2015). Informatsiina tekhnolohiia obrobky tekstovoi informatsii na osnovi pobudovy lohiko-linhvistychnykh modelei [Information technology for processing textual information on the basis of constructing logical-linguistic models]. *International Scientific Journal Acta Universitatis Pontica Euxinus,* (Vols. II), 377–380 Varna, Bulharia [in Ukrainian].

21. Vorontsov, K. V. *Mashinnoe obuchenie [Machine learning]*. Retrieved from http://www.machinelearning.ru/wiki/index.php?title=Mo [in Russian].

References (in language original)

1. *Internet* Security Threat Report [Online]. – Available : https://www.symantec.com/security-center/threat-report.

2. *Чекунов И. Г.* Современные киберугрозы. Уголовно-правовая и криминологическая квалификация киберпреступлений / И. Г. Чекунов // Право и кибербезопасность. – 2012. – № 1. – С. 9–22.

3. *Aickelin*, Uwe and Dasgupta, D. Artificial immune systems. In: Introductory Tutorials in Optimisation, Decision Support and Search Methodology (eds. E. Burke and G. Kendall). Kluwer. Report [Online]. – Available : http://eprints.nottingham.ac.uk/336/1/05intros_ais_tutorial.pdf.

4. *Суркова А. С.* Идентификация авторства текстов на основе информационных портретов / А. С. Суркова // Вестник Нижегородского университета им. Н. И. Лобачевского. – 2014. – № 3 (1). – С. 145–149.

5. *Хмелев Д. В.* Распознавание автора текста с использованием цепей А. А. Маркова / Д. В. Хмелев // Вестник МГУ. Сер. 9: Филология. – 2000. – № 2. – С. 115–126.

6. *Вавіленкова А. І.* Порівняльний аналіз речень природної мови за змістом / А. І. Вавіленкова // Математичні машини і системи. – 2015. – № 2. – С. 97–103.

7. *Гамаюнов Д. Ю.* Обнаружение компьютерных атак на основе анализа поведения сетевых объектов : дис. ... канд. физ.-мат. наук : спец. 05.13.11 / Гамаюнов Денис Юрьевич. – М., 2007. – 89 с.

8. *Network* security modeling and cyber-attack simulation methodology / Chi S.-D., Park J.S., Jung K.-C., Lee J.-S. // Lecture Notes in Computer Science. Springer-Verlag. – 2001. – Vol. 2119.

9. *Котенко И. В.* Архитектуры и модели компонентов активного анализа защищённости на основе имитации действий злоумышленников / И. В. Котенко, М. В. Степашкин, В. С. Богданов // Проблемы информационной безопасности. Компьютерные системы. – 2006. – № 2. – С. 7–24.

10. *Kotenko I. V.* Analyzing Vulnerabilities and Measuring Security Level at Design and Exploitation Stages of Computer Network Life Cycle / I. V. Kotenko, M. V. Stepashkin // Lecture Notes in Computer Science. Springer-Verlag. – 2005. – Vol. 3685.

11. Лукацкий А. В. Обнаружение атак / А. В. Лукацкий. – СПб. : БХВ-Петербург, 2001. – 624 с.

12. *Прилуков М. В.* Роль деловой (конкурентной) разведки в обеспечении национальной безопасности и политической стабильности в Российской Федерации: дис. ... канд. полит. наук : спец. 23.00.02 / Прилуков Михаил Витальевич. – М., 2006. – 351 с.

13. *Бурячок В. Л.* Методологія формування державної системи кібернетичної безпеки : дис. … д-ра техн. наук : спец. 21.05.01 / Бурячок Володимир Леонідович. – К., 2013. – 397 с.

14. *Суркова А. С.* Концептуальный анализ, принципы моделирования и оптимизация алгоритмов синтеза текстовых структур : дис. … д-ра техн. наук : спец. 05.13.01 / Суркова Анна Сергеевна. – Нижний Новгород, 2016. – 343 с.

15. *Додонов В. О.* Інформаційні технології аналізу та виявлення інформаційного впливу в соціальних мережах на основі мультиагентних моделей розповсюдження інформації : дис. ... канд. техн. наук : спец. 05.13.06 / Додонов Вадим Олександрович. – К., 2017. – 143 с.

16. *Колмогоров А. Н.* Три подхода к определению понятия «Количество информации» / А. Н. Колмогоров // Новое в жизни, науке, технике. Сер. «Математика, кибернетика». – 1991. – № 1. – С. 24–29.

17. *Пиотровский Р. Г.* Текст, машина, человек / Р. Г. Пиотровский. – Л. : Наука, 1975. – 327 с.

18. *Shannon C. E.* A mathematical theory of communication / C. E. Shannon // Bell System Technical. – 1948. – Vol. 27. – P. 379–423.

19. *Web Ontology* Language (OWL) [Online]. – Available : https://www.w3.org/2001/sw/wiki/OWL.

20. *Вавіленкова А. І.* Інформаційна технологія обробки текстової інформації на основі побудови логіко-лінгвістичних моделей / А. І. Вавіленкова // International Scientific Journal Acta Universitatis Pontica Euxinus. Special number for XI international conference «Strategy of quality in industry and education» (Varna, Bulgaria, 1 – 5 June 2015). – Varna, Bulgaria, 2015. – Vol. II. – P. 377–380.

21. *Воронцов К. В.* Машинное обучение (курс лекций) [Электронный ресурс] / К. В. Воронцов. – Режим доступу : http://www.machinelearning.ru/wiki/index.php?title=Mo.

УДК 004.72.056.52

Віталій Литвинов, Микола Стоянов, Ігор Скітер,
Олена Трунова, Алла Гребенник

ЗАХИСТ КОРПОРАТИВНИХ МЕРЕЖ ВІД АТАК З ВИКОРИСТАННЯМ КОНТЕНТ-АНАЛІЗУ ГЛОБАЛЬНОГО ІНФОРМАЦІЙНОГО ПРОСТОРУ

Актуальність теми дослідження. *Подальше вдосконалення захищеності корпоративних мереж в умовах масованого впливу комп'ютерних атак вимагає підвищення ймовірності виявлення нових комп'ютерних атак і зниженням часу розпізнавання ознак відомих атак.*

Постановка проблеми. *Аналіз текстів глобального інформаційного простору дозволяє скоротити час виявлення можливих загроз.*

Аналіз останніх досліджень і публікацій. *Були розглянуті останні публікації щодо систем захисту від атак та використання аналізу текстів у разі виявлення загроз.*

Виділення недосліджених частин загальної проблеми. *Вдосконалення методів опрацювання масивів даних тіла мережевих пакетів, вмісту інтернет-сторінок, інформації СМІ та соціальних мереж, що у свою чергу порушує проблему семантико-синтаксичної обробки текстів природної мови.*

Постановка завдання. *Організація колективного захисту корпоративних мереж шляхом впровадження систем моніторингу загроз, активної розвідувальної діяльності в глобальному інформаційному просторі з метою пошуку, збору та аналізу даних про атаки, аномальну поведінку, вміщуваний контент ресурсів мережі Інтернет .*

Виклад основного матеріалу. *Вимоги систем захисту щодо скорочення часу виявлення загроз призводять до необхідності ведення активної розвідувальної діяльності, спрямованої на проведення постійного моніторингу оточуючого кіберпростору, що складається з множини комп'ютерних мереж окремих користувачів та організацій. Метою такого моніторингу є визначення характеристик, інтересів, особливостей політики безпеки конкретної корпоративної мережі в глобальному інформаційному просторі. Особливе значення при цьому набуває аналіз текстової інформації з відкритих та умовно-відкритих електронних джерел. Раціональним вирішенням такого завдання є формування центрів моніторингу загроз, спрямованих на організацію колективного захисту підлеглих корпоративних мереж.*

Висновки відповідно до статті*. Запропонований метод захисту дозволяє як виявляти кіберзагрози в глобальному інформаційному просторі, так і налаштовувати власні системи захисту корпоративних мереж згідно з їх характеристичними векторами загроз.*

Ключові слова: *системи захисту від атак; корпоративна мережа; розвідувальна діяльність; моделі представлення текстів; колективний захист.*

Рис.: 5. Бібл.: 21.

УДК 004.72.056.52

Виталий Литвинов, Николай Стоянов, Игорь Скитер,
Елена Трунова, Алла Гребенник

ЗАЩИТА КОРПОРАТИВНЫХ СЕТЕЙ ОТ АТАК С ИСПОЛЬЗОВАНИЕМ КОНТЕНТ-АНАЛИЗА ГЛОБАЛЬНОГО ИНФОРМАЦИОННОГО ПРОСТРАНСТВА

Актуальность темы исследования. *Дальнейшее усовершенствование защищенности корпоративных сетей в условиях массированного влияния компьютерных атак требует роста вероятности выявления новых компьютерных атак и уменьшения времени распознавания признаков известных атак.*

Постановка проблемы. *Анализ текстов глобального информационного пространства позволяет сократить время обнаружения возможных угроз.*

Анализ последних исследований и публикаций. *Были рассмотрены последние публикации по системам защиты от атак и использованию анализа текстов для обнаружения угроз.*

Выделение неисследованных частей общей проблемы. *Совершенствование методов обработки массивов данных тела сетевых пакетов, содержания интернет-страниц, информации СМИ и социальных сетей, что в свою очередь поднимает проблему семантико-синтаксической обработки текстов естественного языка.*

Постановка задачи. *Организация коллективной защиты корпоративных сетей путем внедрения систем мониторинга угроз, активной разведывательной деятельности в глобальном информационном пространстве с целью поиска, сбора и анализа данных об атаках, аномальном поведении, размещаемом контенте ресурсов сети Интернет.*

Изложение основного материала. *Требования систем защиты относительно минимизации времени выявления угроз приводят к необходимости ведения активной разведывательной деятельности, направленной на ведение постоянного мониторинга окружающего киберпространства, состоящего из множества компьютерных сетей отдельных пользователей и организаций. Цель такого мониторинга - определение характеристик, интересов, особенностей политики безопасности конкретной корпоративной сети в глобальном информационном пространстве. Особое значение при этом приобретает анализ текстовой информации из открытых и условно-открытых электронных источников. Рациональным решением такой задачи является формирование центров мониторинга угроз, направленных на организацию коллективной защиты подчиненных корпоративных сетей.*

Выводы в соответствии со статьей. *Предложенный метод защиты позволяет как выявлять киберугрозы в глобальном информационном пространстве, так и настраивать собственные системы защиты корпоративных сетей в соответствии с их характеристическими векторами угроз.*

Ключевые слова: *системы защиты от атак; корпоративная сеть; разведывательная деятельность; модели представления текстов; коллективная защита.*

Рис.: 5. Библ.: 21.

Lytvynov Vitalii – Doctor of Technical Sciences, Professor, Head of Department of Information Technology and Software Engineering, Chernihiv National University of Technology (95 Shevchenka Str., 14035 Chernihiv, Ukraine).

Литвинов Віталій Васильович – доктор технічних наук, професор, завідувач кафедри інформаційних технологій та програмної інженерії, Чернігівський національний технологічний університет (вул. Шевченка, 95, м. Чернігів, 14035, Україна).

Литвинов Виталий Васильевич – доктор технических наук, профессор, заведующий кафедрой информационных технологий и программной инженерии, Черниговский национальный технологический университет (ул. Шевченко, 95, г. Чернигов, 14035, Украина).

E-mail: vlitvin@ukrsoft.ua

ORCID: http://orcid.org/0000-0003-2334-2275

Stoianov Nikolai – Doctor of Technical Sciences, Associate professor, Deputy Director, Bulgarian Bulgarian Defence Institute Prof.. Tsvetan Lazarov (2 Professor Tsvetan Lazarov Blvd., 1592, Sofia, Bulgaria).

Стоянов Микола – доктор технічних наук, доцент, заступник директора, Болгарський інститут оборони ім. Цвєтана Лазарова (бульвар Професора Цвєтана Лазарова, 2, м. Софія, Болгарія, 1592).

Стоянов Николай – доктор технических наук, доцент, заместитель директора, Болгарский институт обороны им. Цветана Лазарова (бульвар Профессора Цветана Лазарова, 2, г. София, Болгария, 1592).

E-mail: n.stoianov@di.mod.bg

Skiter Igor – PhD in Physical and Mathematical Sciences, Assistant Professor, Doctoral Candidate, the Institute of Mathematical Machines and Systems Problems National Academy of Science of Ukraine (42 Academician Glushkova Av., 03187, Kyiv, Ukraine).

Скітер Ігор Семенович – кандидат фізико-математичних наук, доцент, докторант, інститут проблем математичних машин і систем НАН України (просп. Глушкова, 42, м. Київ, 03187, Україна).

Скитер Игорь Семенович – кандидат физико-математических наук, доцент, докторант, институт проблем математических машин и систем НАН Украины (просп. Глушкова, 42, г. Киев, 03187, Украина).

E-mail: skiteris@ukr.net

ORCID: http://orcid.org/0000-0003-2334-2276

ResearcherID: F-5950-2014

Trunova Helen – PhD in Pedagogical Sciences, Assistant Professor, Assistant Professor of Department of Information Technology and Software Engineering, Chernihiv National University of Technology (95 Shevchenka Str., 14035 Chernihiv, Ukraine).

Трунова Олена Василівна – кандидат педагогічних наук, доцент, доцент кафедри інформаційних технологій та програмної інженерії, Чернігівський національний технологічний університет (вул. Шевченка, 95, м. Чернігів, 14035, Україна).

Трунова Елена Васильевна – кандидат педагогических наук, доцент, доцент кафедры информационных технологий и программной инженерии, Черниговский национальный технологический университет (ул. Шевченко, 95, г. Чернигов, 14035, Украина).

E-mail: e.trunova@gmail.com

ORCID: http://orcid.org/0000-0003-0689-8846

Hrebennyk Alla – PhD student, the Institute of Mathematical Machines and Systems Problems National Academy of Science of Ukraine (42 Academician Glushkova Av., 03187, Kiev, Ukraine).

Гребенник Алла Григорівна – аспірант, Інститут проблем математичних машин і систем НАН України (просп. Глушкова, 42, м. Київ, 03187, Україна).

Гребенник Алла Григорьевна – аспирант, Институт проблем математических машин и систем НАН Украины (просп. Глушкова, 42, г. Киев, 03187, Украина).

E-mail: grebennik.alla@gmail.com

Lytvynov, V., Stoianov, N., Skiter, I., Trunova, H., Hrebennyk, A. (2018). Corporate networks protection against attacks using content-analysis of global information space. *Technical sciences and technologies*, 1 (11), 115-130.