Managing the Security of the Critical Infrastructure Information Network

Serhii Toliupa[1], Anatolii Shevchenko[2], Serhii Buchyk[1], Ihor Pampukha[2], and Andrii Kulko[1]

[1] *Taras Shevchenko National University of Kyiv, 60 Volodymyrska str., Kyiv, 01033, Ukraine*
[2] *Military Institute of Taras Shevchenko National University of Kyiv, 60 Volodymyrska str., Kyiv, 01033, Ukraine*

Abstract

Critical infrastructure has a multi-level structure, encompassing technical components, social aspects, organizational elements, and government involvement, all interconnected through a distributed information system that demands protection. The development and implementation of cutting-edge information technologies have created unprecedented conditions for the gathering and utilization of data, resulting in a fundamental reliance on their uninterrupted operation across various sectors of society and the state, including the economy, politics, national security, and international affairs. This dependence also exposes vulnerabilities in the functionality of critical national infrastructure systems, enabling adversarial entities and groups to exploit it for illegal activities in cyberspace. They do so by compromising the integrity, availability, and confidentiality of information, thereby inflicting damage upon information resources and systems. This article proposes a method for managing information system security based on internal cyberattacks. The method relies on modifications to the support vector method, utilizing parameters typical for internal cyberattacks on information systems. Its primary objective is to identify the input parameters of internal cyberattacks and enhance the reliability of decision-making in assessing the state of IP security, all within the timeframe comparable to existing methods. The mathematical framework employed in this method reduces the volume of input data required for managing information system security while bolstering the reliability of decision-making in evaluating the security status of critical infrastructure's information resources.

Keywords

Cybersecurity, critical infrastructure, threat, cyberspace, information resource.

1. Introduction

Many countries are implementing the concept of critical infrastructure, allowing them to focus on systems, networks, and individual facilities that could have profound adverse effects on national security if destroyed or disrupted [1–3]. The rapid development of information and communication technologies over the past two decades has significantly impacted the operation of critical infrastructure facilities. These technologies are no longer just tools for exchanging and processing information; they have become instruments for causing harm. Critical infrastructure comprises multiple levels, including technical components (equipment and devices), the social level (personnel responsible for maintaining technical components), the organizational level (interaction among the services operating critical infrastructure), and the level of state administration (regulatory and supervisory bodies overseeing critical infrastructure). These levels are all interconnected through a distributed information system that requires protection. The complexity of critical infrastructures arises from the intricate

CPITS-2023-II: Cybersecurity Providing in Information and Telecommunication Systems, October 26, 2023, Kyiv, Ukraine
EMAIL: tolupa@i.ua (S. Toliupa); tolyamixailshevchenko75@gmail.com (A. Shevchenko); buchyk@knu.ua(S. Buchyk); pamp@ukr.net (I. Pampukha); kulko.andrii@gmail.com (A. Kulko)
ORCID: 0000-0002-1919-9174 (S. Toliupa); 0000-0003-2723-0378 (A. Shevchenko); 0000-0003-0892-3494 (S. Buchyk); 0000-0002-4807-3984 (I. Pampukha); 0009-0006-1185-0774 (A. Kulko)

structure intricate interdependencies and nonlinear relationships among components and levels of the system [4–6].

The development and implementation of cutting-edge information technologies have created unprecedented conditions for data accumulation and utilization, leading to fundamental dependence on their continuous operation across various aspects of society and the state, such as the economy, politics, and national and international security. This dependence exposes vulnerabilities in the functioning of critical national infrastructures, enabling adversarial entities and groups to exploit them for illegal activities in cyberspace by compromising the integrity, availability, and confidentiality of information, resulting in damage to information resources and systems. Furthermore, the potential use of information technologies in cyberspace for military-political power struggles, terrorism, and hacker attacks is a significant concern, especially in times of martial law [7].

2. Issues

The existence of close interconnections between critical infrastructure components is a fundamentally important feature that has a decisive impact on the nature of their functioning in regular and emergencies. On the one hand, the interconnection of the CI elements increases their efficiency and survivability, allowing the rational use and redistribution of available resources and capacities, and on the other hand, it makes them vulnerable to a large-scale disaster, the huge amount of damage that cannot be ignored regardless of the low probability of the risks. The analysis of modern hybrid warfare in the world and cyber operations shows the evolution of forms, methods, and techniques used to conduct cyberattacks on critical infrastructure management systems. Studies have shown that for the effective functioning of the information system, it is advisable to use an appropriate security subsystem as part of the IP management system with the ability to assess and manage the security status of the information system in real-time and in conditions of the constantly changing character of cyberattacks [8].

The critical infrastructure information system possesses distinct features, including varying network dimensions, geographically scattered information system components, and elements that extend beyond controlled areas. These features make it suitable for facilitating rapid interaction among geographically dispersed critical infrastructure units. However, they also render it vulnerable to exploitation by intruders who can conduct cyberattacks, causing destructive actions against the system and the management of critical infrastructure information resources in general. The operation of the information system in cyberspace is influenced by a range of factors, encompassing natural, human-made, and anthropogenic elements that disrupt the information transfer process. Nevertheless, the most significant harm to the information system can result from deliberate attacks that exploit current security weaknesses. These attacks can occur at various levels of the basic reference model for open system interaction [9–10].

Hence, one of the most critical challenges in operating the critical infrastructure information system is ensuring the security of information, software, and hardware. To effectively manage the security of information resources, it is imperative to utilize specialized equipment, algorithms, and methods that guarantee the secure operation of nodes, components, and the entire information system.

The existing methods of managing the state of the information system security do not take into account the peculiarities of cyberattacks by internal and external intruders and also have low reliability in making a management decision on assessing the state of the information system security and implementing security measures. At the same time, the main requirements for the methods of managing the state of the information resources security in an information system are real-time operation; consideration of threats characteristic of information and telecommunication systems; adaptive functioning of the information security system with self-organization; decentralization of management and hierarchical distribution

structure; increase in the reliability and completeness of management decision-making; reduction of mathematical complexity and resource burden of methods; flexibility of the mathematical apparatus; application of special samples on the state of the information system security; possibility of application in systems with high dynamics of topology change; decentralisation of management and availability of a hierarchical distribution structure; minimum network load with service information [11].

As a result, today there is a discrepancy between the above capabilities of the existing methods of managing the state of the information resources security and the requirements for methods of managing the state of the information resources security in the information system, to solve this scientific task, namely: development of methods of managing the state of the IP security against external and internal cyber-attacks on the information system using identification methods, dynamic programming and support vectors.

Due to the complex structure of the critical infrastructure facilities and the complex nature of the interactions between a significant number of elements, the possibilities of conducting scenario analysis using traditional tools (event trees, fault trees, Bayesian networks, neural networks) are limited [11]. To describe the development of disruptions in critical infrastructure, network models are applied, which extensively use the mathematical tool of graph theory. Networks are an extremely flexible notion that can be widely used in the study of infrastructure systems. In this case, a hierarchy of mathematical models of varying complexity can be built to describe various aspects of infrastructure system risks about potentially initiating impacts. With the help of these models, it is possible to describe many properties and characteristics of network systems: chaos, self-organization, statistical distributions, and criticality.

3. Quantum Authentication

The peculiarity of modern critical infrastructures is that they are becoming cross-border and, in some cases, global. The spatial dimension of critical infrastructures, along with the existence of close interconnections between them, makes their functioning dependent on a huge number of factors related to the state of the natural, technogenic, and social environment in different regions of the world. A significant amount of hazardous substances, energy, and information stored, transported, and processed by critical infrastructures, as well as their huge role in the economy and human life, determine the possibility of large-scale accidents at critical infrastructures and the severity of the consequences arising from such accidents for the population and economic facilities. The complexity of critical infrastructures significantly impedes the creation of effective protection systems, as it becomes almost impossible to conduct a detailed scenario analysis of the system, identify all major hazard scenarios, and determine a set of protection measures and barriers aimed at parrying all possible threats [12].

At the same time, an analysis of the current practice in the field of critical infrastructure functioning shows that their design, construction, and operation are carried out by the traditional paradigm of ensuring the security of technical systems (technical services) and information security systems. This paradigm includes: analyzing possible failure scenarios in the system; identifying the most critical scenarios; and creating protective barriers to prevent these scenarios.

It should be noted that efforts to protect critical infrastructure facilities are traditionally focused on technical aspects. In virtue of that significant progress has been made in ensuring the reliability of technical components of critical infrastructure. However, the capabilities of this approach are close to being exhausted. This is because critical infrastructures can no longer be considered as predominantly technical systems, but are becoming more and more techno-social systems.

Due to the rapid development of information technology in recent decades, critical infrastructure facilities are becoming more complex. Thus, there are a lot of factors

to consider when assessing the security of critical infrastructure, and some operating modes of critical infrastructure are fully applied. The reason for this is the complex nonlinear interactions between critical infrastructure components, the strong degree of interconnectedness between different subsystems, and the fact that critical infrastructure and the environment are beginning to change faster than they can be described and studied. This creates a situation where there is a lack of information about critical infrastructures and, as a result, limited opportunities to predict their performance and manage them. At the same time, in certain modes, it is impossible to describe in detail the principles of critical infrastructure functioning and develop management rules [13].

The distinction between fully determined and underdetermined systems becomes extremely important in developing a set of security measures. The peculiarity of underdetermined systems is that it is impossible to fully describe their performance and predict their state under different conditions and in different operating modes. As a result, for complex systems such as critical infrastructures, it is almost impossible to create a closed list of projected impacts to which the system may be exposed during its operation. In this regard, the traditional strategy of ensuring critical infrastructure security, based on the development of a set of protective barriers designed to shut back projected impacts, cannot be successful [14].

In our case, we consider the situation of an equiprobability of the system being in a state of security breaches of the Information System (IS) of critical infrastructure. At the same time, there are both security breaches from internal cyberattacks on the information resource of critical infrastructure, the protection of which is required by law, in the information system, and the search for countermeasures to detect changes in the state of security. To simulate this situation, a training sample is created that contains 20% of normal messages and 80% of anomalous messages with the types of attacks. A base with response options for a variety of detected breaches is also being built.

The input data are: $X = XH \cup XM \cup XL$ — parameters of inbound traffic;

$XM = \{x_m(t), m = \overline{1,18}\}$ is a set of traffic parameters that are specific to external cyber-attacks;

$XH = \{x_h(t), h = \overline{1,15}\}$ is a set of traffic parameters that are specific to internal cyber-attacks;

XL is a set of traffic parameters that are not used in the implementation of a method;

$S(t), s = \overline{1,10}$ are parameters of cybersecurity sensors;

$XV = \{XH \cup S(t)\}$ is input data specific to internal cyberattacks.

Limitations and assumptions: Attack types are identified: DoS, U2R, R2L, Probe, and Side. To identify the behavior, the attack signatures that are threats to the information system are considered. Abnormal behavior is identified as a newly detected security status. The process of managing the security status is quasi-stationary on a time frame $(t_0 ... T)$.

It is required: to increase the reliability of management decision-making regarding the assessment of the security status of the critical infrastructure information resource, the requirement to protect which is established by law from external cyberattacks, if the time for management decision-making will not exceed that of similar methods.

The principle of the method is to distribute control of the security status of the information system based on a set of input parameters specific to internal cyberattacks and a set of parameters of cybersecurity sensors using the description of the information system and the support vector machines.

Support Vector Machine (SVM) is a set of similar algorithms of the "supervised learning" models used in analyzing data for classification and regression analysis. This method belongs to a family of linear classifiers. A peculiarity of the support vector machine is a continuous minimization of the empirical classification error and maximizing the width of the gap between the classes. Hence, this method is often called the maximum margin classifier [15].

The method searches for elements located on the boundaries between two classes, which are called support vectors.

The support vector machine searches for a linear function that enables to assignment of the elements of a dataset to one of two classes. The task of binary classification can be defined as the search for a linear function f(x) that takes a value less than zero for elements of one class and greater than zero for elements of another [16].

The separating hyperplane has the following form:

$$f(x) = \varpi \times x - b = 0, \quad (1)$$

where w is a vector perpendicular to the separating hyperplane, the parameter b determines the distance of the hyperplane from the origin.

The hyperplane parallel to the optimal hyperplane and closest to the support vectors of the two classes can be defined by the following equations:

$$\begin{aligned} \varpi x - b &= 1 \\ \varpi x - b &= -1 \end{aligned} \quad (2)$$

If the training data is linearly inseparable, then we can select hyperplanes to prevent data points from falling into the margin between them and then maximize the distance between the hyperplanes. In this case, the distance between the planes is $\frac{2}{\|\varpi\|}$, so we should minimize it $\|\varpi\|$. To exclude all points from the line, the following conditions must be satisfied [11]:

$$c_i(\varpi x_i - b) \geq 1, \ 1 \leq I \leq n, \quad (3)$$

where c_i is a class label that takes a value –1 and +1, and x_i is a sample vector with class label c_i.

This quadratic optimization problem is equivalent to the problem of finding the saddle point of the Lagrange function [12]:

$$\begin{aligned} &-L(\lambda) = \sum_{i=1}^{n} \lambda_i + \frac{1}{2} \sum_{i=1}^{n} \lambda_i \lambda_j c_i c_j (x_i x_j) \to \min_{\lambda} \\ &\lambda_i \geq 0, 1 \leq i \leq n \\ &\sum_{i=1}^{n} \lambda_i c_i = 0 \end{aligned} \quad (4)$$

where L is the Lagrange function, λ_i are Lagrange multipliers.

To generalize the SVM to the case of linear inseparability, the constant C is introduced—an internal parameter of the method that allows you to adjust the ratio between maximizing the width of the separation band and minimizing the total error.

The main problem of using the support vector method in a binary classification task is the difficulty of finding a linear boundary between two classes. If it is not able to construct such a boundary, one solution is to increase the dimensionality (transferring data to another space of higher dimensionality), where it is possible to construct a plane that divides the set of elements into two classes [17].

Thus, the problem of timely detection of changes in the state of information security of the critical infrastructure and information system's object is solved by managing the state of information system security based on a set of parameters of internal attacks in the conditions of limited samples of current observation data.

Management of the state of protection of the information system of critical infrastructure facilities against internal cyberattacks occurs in the case of identification of the parameters of violations that are realized by a set of multidirectional and different attacks.

Identify the input data (data parameters) of the traffic.

I. By identification, we mean finding a model that is optimal in some sense, based on the results of observations of the input and output variables of the object, namely, a set of traffic parameters. Identification is the reverse task of system synthesis.

Parametric identification will be used to identify input data based on parameters that are typical of internal cyberattacks.

In parametric identification, data about a critical infrastructure facility is processed to obtain posterior information about it. The parameters of the selected model are estimated. In the simplest cases, such an assessment can be performed using a transient response graph.

The task of parametric identification can be formulated as follows: to select such values $\tilde{X}$ on the set of $\{X\}$ possible parameter values so that the differences of the indicators reach their minimums, i.e. the purpose of this analysis is to search:

$$\tilde{X}(t) = \arg\min_{y \in \hat{X}} \sum_{i=1}^{18} (y_i(t) - x_i(t))^2 \quad (5)$$

where $\hat{X}$ is traffic parameters described in the database; $x_i(t)$ is parameters that describe the flow of incoming traffic data and are obtained from the data distribution block.

II. The next step is to retrieve from the database a set of data on the state of the information system security $X(t)=\{x_1,...,x_{25}\}$, a set of possible security breaches $\Lambda=\{\lambda_1,...,\lambda_n\}$, and a set of possible means of counteracting breaches (management decisions) $U=\{u_1,...,u_n\}$, where n is the number of options in the sets contained in the database; assessing security based on the dependence of the average value of the set of optimal values of means of counteracting breaches $\overline{U}$ on the value Λ and establishing the state of security.

Pattern recognition is taught as follows. There is a set of observations (security states) that belong to p different classes. The components of the vector are individual security threats to the information system's elements. Using information about observations and their classification, it is necessary to find a rule that would allow classifying changes in the security state (new observations) with a minimum number of errors.

Classes of observations can be situations of changes in the state of security of information system elements. For example, for two classes: the state of security of information system elements is deteriorating; the state of security of information system elements is improving. An example for three classes: the state of protection of information system elements is deteriorating; the state of protection of information system elements remains unchanged; the state of protection of information system elements is improving. The number of classes can be arbitrary and determined by the condition of unambiguous classification of the current situation.

We assume that the observation is given by the vector x, and its classification is given by the number ω (ω can take p values: 0, 1, ..., $p-1$). In practice, the vector of observations will be a vector whose components will be numerical estimates of the information system's security. The dimension of the vector will correspond to the number of threats submitted for consideration.

Thus, given a sequence of l observations and classifications $x_1,\omega_1;...;x_l,\omega_l$, it is necessary to construct a decisive rule $\omega=F(x)$ that would classify new observations with the least possible number of errors.

To formalize the word "error", we assume that there exists (although it is unknown) some rule Φ that defines for each vector x a classification $\omega=\Phi(x)$, which is called "true". An error in the classification of the vector x using the rule $F(x)$ is a classification in which $F(x)$ and $\Phi(x)$ do not coincide.

To be able to use mathematical analysis, we will assume that the rule $F(x)$ is one of the functions of some given set of functions $\{F(x)\}$, and the classification rule $\Phi(x)$ is determined by the conditional probability $P(\omega|x)$.

It is commonly assumed that there exists an unknown probability measure on the space of x vectors (we denote it by the density $P(x)$). By $P(x)$, situations x appear randomly and independently, which are classified using the $P(\omega|x)$ rule. Thus, the training sequence is determined

$$x_1,\omega_1;...;x_l,\omega_l. \qquad (6)$$

For any deciding $F(x)$ rule, let's define quality as the probability of different classifications using rule $F(x)$ and rule $P(\omega|x)$. The lower this probability, the higher the quality. Formally, the quality of the deciding rule can be written in the form:

$$I(F)=\sum_{i=0}^{p-1}\int\Theta(F(x)-\omega_i)P(\omega_i|x)P(x)dx \qquad (7)$$

where $\Theta(z)=\begin{cases}0, & z=0\\ 1, & z\neq 0\end{cases}$.

It is not possible to calculate directly the probability of an error-free classification for any deciding rule $F(x)$, since the densities $P(x)$ and $P(\omega|x)$ are not known.

Using the sample (6), find a rule in the class $\{F(x)\}$ that minimizes the functionality (7).

For convenience, we will assume that:

1. The variable ω takes only two values: 0 and 1 (i.e., that the situation x belongs to one of the two classes); this restriction is not fundamental, since a sequential division into two classes can be obtained by dividing into any finite number of classes.

2. The class of indicator functions $\{F(x)\}$, i.e. functions that take two values: 0 and 1, is parametric $\{F(x,\alpha)\}$ (here α is a parameter that belongs to the set Λ, the specific value of which $\alpha=\alpha^*$ determines a specific function $F(x,\alpha^*)$ of the class $F(x,\alpha)$; to find the required function in the class means to set the required value of the parameter in the class; studying only the parametric class of functions does not reduce the generality in the definition of the class of functions, since the set Λ is arbitrary: it can be a set of scalar values, a set of vectors, or a set of abstract elements).

3. Write the functionality (2) in the form

$$I(\alpha)=\int(\omega-F(x,\alpha))^2\,P(x,\omega)\,dx\,d\omega, \quad (8)$$

where the function $P(x,\omega)=P(\omega|x)P(x)$ is called the joint density of pairs x,ω given on the space X,Ω.

Thus, the task of pattern recognition training is to find one in the class of indicator functions $F(x,\alpha)$ that would minimize the functional (8) under conditions when the joint density $P(x,\omega)$ is unknown, but a probable and independent sample of pairs obtained according to this density is given.

Pattern recognition learning algorithms are based on a special method of finding a decisive rule based on the construction of a separating hyperplane.

To build the guide vector ϕ_0 we will use the results of our work. Consider the finite set of vectors Z, which consists of all possible differences formed by the vectors of the set X and the vectors of the set $\bar{X}$:

$Z=\{z_{ij}=x_i-\bar{x}_j\}$, $i=\overline{1,a}$, $j=\overline{1,b}$

($a\times b$ elements in total).

Let's find the minimal module vector ψ_0 that satisfies the inequality:

$$z_{ij}\psi\geq 1,\ z_{ij}\in Z. \quad (9)$$

The vector ψ_0 coincides in direction with the optimal vector ϕ_0, and the value $\frac{1}{\|\psi_0\|}$ is the distance between the projections of the sets X and $\bar{X}$ the direction of the vector ψ_0.

Thus, to find the vector ψ_0 and use it to construct the optimal separating hyperplane, it is necessary to minimize the functional

$$I=\psi^T\psi, \quad (10)$$

when the constraints (9) are met.

Finding the minimum of (10) under the constraints (9) is a quadratic programming problem, the solution of which is based on the Kuhn-Tucker theorem, which specifies the necessary and sufficient conditions for the minimum. The following theorem follows from the above.

Theorem 1. The minimal module vector ψ_0 satisfying (9) can be given as

$$\psi_0=\sum_{i=1}^{a}\sum_{j=1}^{b}z_{ij}\alpha_{ij}^0,\ \alpha_{ij}^0\geq 0, \quad (11)$$

and

$$\alpha_{ij}^0\left[z_{ij}^T-1\right]=0,\ i=\overline{1,a},\ j=\overline{1,b}. \quad (12)$$

Among all the vectors ψ satisfying (7), the vector ψ given in the form (11) and (12) is minimal in the module.

Let us call vectors z_{ij}^*, for which conditions are being made $z_{ij}^*\psi_0=1$ extreme vectors. According to Theorem 1, a minimal modulo directing vector can be provided in the form of a linear combination of extreme vectors. Vectors x_i^*, $\bar{x}_j^*$, forming extreme vectors z_{ij}^*, will be called informative.

Let us consider the problem, a solution that is equivalent to the composition of the optimal directing vector. Assume that α is parameters vector α_{ij}. Consider the function

$$W(\alpha)=\sum_{i=1}^{a}\sum_{j=1}^{b}\alpha_{ij}-\frac{1}{2}\psi^T\psi, \qquad (13)$$

where $\psi=\sum_{i=1}^{a}\sum_{j=1}^{b}z_{ij}\alpha_{ij}$.

In this case, α_0 is the maximum point of a function $W(\alpha)$ in a positive quadrant ($\alpha_{ij}^0\geq 0$) and determines the optimal separating vector.

Indeed, necessary and sufficient condition for function $W(\alpha)$ maximum in the point α_0 are the following

$$\frac{\partial W(\alpha_0)}{\partial\alpha_{ij}}=\begin{cases}0, & if \quad \alpha_{ij}^0>0, \quad i=\overline{1,a}\\ \leq 0, & if \quad \alpha_{ij}^0=0, \quad j=\overline{1,b}\end{cases}.$$

Let us write these conditions down and mark $\psi_0=\sum_{i=1}^{a}\sum_{j=1}^{b}z_{ij}\alpha_{ij}^0$.

We will obtain:

$$1-z_{ij}^T\psi_0=\begin{cases}0, & if \quad \alpha_{ij}^0>0, \quad i=\overline{1,a}\\ \leq 0, & if \quad \alpha_{ij}^0=0, \quad j=\overline{1,b}\end{cases}.$$

These conditions may be rewritten in the form of inequation

$$z_{ij}^T\psi_0\geq 1,\ \alpha_{ij}^0\geq 0,\ i=\overline{1,a},\ j=\overline{1,b}$$

and the equations $\alpha_{ij}^{0T}\left(1-z_{ij}\psi_0\right)=0$, $i=\overline{1,a}$, $j=\overline{1,b}$.

According to the theorem 1 assertion, these conditions determine the optimal directing vector.

Thus, the problem of building a hyperarea that divides two vectors' multiplicities has been reduced to the search for the function maximum $W(\alpha)$ in the positive quadrant.

An important issue for the search for a maximum for quadratic form (13) is the following.

Theorem 2. If dividing hyperarea exists (that is vector ψ_0, for which inequation is fulfilled (13)), then the function maximum $W(\alpha)$ in a positive quadrant equals half of the squared absolute value of the optimal directing vector $W(\alpha_0)=\frac{\|\psi_0\|^2}{2}$.

From the theorem arises consequence which is important for algorithm development.

Consequence. Valid estimate

$$\rho(\psi)\geq\frac{1}{\sqrt{2W(\alpha_0)}}, \qquad (14)$$

where $\rho(\psi)$—distance between set projections X and $\bar{X}$ to the direction ψ.

In this case, estimate equality (14) is reached if $\psi=\psi_0$ or if $\alpha=\alpha_0$.

This consequence is used for the criteria construction of vector inseparability. Two finite sets of vectors are virtually not separated by hyperarea, or simply inseparable if the distance between these set projections in any direction is less than preassigned ρ_0. This means there is no inseparability, if $\alpha_{ij}^*>0$, that $W(\alpha^*)>\frac{1}{2\rho_0^2}=W_0$.

Therefore, when generating an optimal directing vector it is necessary to calculate maximum of positively obtained quadratic form $W(\alpha)$ of a positive quadrant $\alpha_{ij}\geq 0$ or find out that the function maximum $W(\alpha)$ exceeds a prescribed value W_0. The latter means that the generation of separating hyperarea is impossible.

One of the most effective maximization algorithms of nonpositively defined quadratic form is a method of conjugate gradients. Using it one can attain maximum in n steps (n is form dimensionality). Let us consider the conjugate gradients method to maximize negative square form $F(y)=b^Ty-y^TAy$, where A is positively determined matrix, b, y are vectors.

Using conjugate gradient, function maximum search begins with an arbitrary point $y_0=y(0)$. The first step is taken in the direction of the gradient of function $F(y)$ at the point of $y(0)$.

Let us mark the gradient of a function at the point $y(0)$ through $g(1)$, and motion direction from the point $y(0)$ through $z(1)$.

Therefore, $z(1)=g(1)$.

A step is being taken in the direction of $z(1)$ attaining a maximum in this direction. Direction $z(1)$ maximum is set by equation

$$y(1)=y(0)+\frac{z^T(1)g(1)}{z^T(1)Az(1)}z(1)\cdot \qquad (15)$$

Beginning with the second step, motion direction is obtained by vector

$$z(t+1)=g(t+1)+\frac{\|g(t+1)\|^2}{\|g(t)\|^2}z(t), \qquad (16)$$

where $g(t+1)$ and $g(t)$ is gradient of function $F(y)$ at the points $y(t+1)$ та $y(t)$ respectively; $z(t)$ is motion direction at the point $y(t-1)$.

Motion in the direction of $z(t)$ is made before attaining conditional maximum. This maximum is reached at the point

$$y(t)=y(t-1)+h(t)z(t), \qquad (17)$$

where the motion step is provided by the value:

$$h(t)=\frac{z^T(t)g(t)}{z^T(t)Az(t)}. \qquad (18)$$

Thus, formulas (15)–(18) assign a search algorithm for the maximum quadratic form $F(y)$.

Method modification is guided to limit the region of search by a positive quadrant.

Define the function

$$\hat{g}_i(t)=\begin{cases}\dfrac{\partial F(y)}{\partial y_i}, & if \quad y_i>0 \quad чи \quad \dfrac{\partial F(y)}{\partial y_i}>0;\\ 0 \qquad if \quad y_i=0 \quad ma \quad \dfrac{\partial F(y)}{\partial y_i}\le 0.\end{cases}$$

Vector $\hat{g}(y)=\left(\hat{g}_1(y),...,\hat{g}_n(y)\right)^T$ is a conditional gradient of the function $F(y)$ on the multitude $y_i\ge 0$.

We conduct ascendancy to the maximum using formulas (15)–(18), where $g(y)$ is replaced with $\hat{g}(y)$. Motion begins with an arbitrary point of the positive quadrant and continues until it reaches the restriction in the point y_0. Then ascendancy begins again by the method of conjugate gradients, but from the point y_0. The search for maximum ends if inequity is solved $\left|\hat{g}_i(y)\right|\le\varepsilon$.

To keep the trajectory within the borders of the positive quadrant, step size $\hat{h}(t)$ is selected if two dimensions are minimal:

$$\hat{h}(t)=\min\left(h(t),h^*(t)\right),$$

where $h^*(t)=\min\limits_i\dfrac{y_i(t)}{\left|z_i(t+1)\right|}$.

When calculating $h^*(t)$ minimum is only determined by the coordinates i, for which $z_i<0$. If $z_i\ge 0$, step equals $h(t)$.

To form an optimal directing vector it is necessary either to determine that the function maximum $W(\alpha)=\sum\limits_{i=1}^{a}\sum\limits_{j=1}^{b}\alpha_{ij}-\dfrac{1}{2}\psi^T\psi$, where $\psi=\sum\limits_{i=1}^{a}\sum\limits_{j=1}^{b}\alpha_{ij}\left(x_i-\overline{x}_j\right)$, in the positive quadrant is bigger than a prescribed value W_0 (this means that faultless vectors' separation is possible), or, if it is not this way, to find a point α_0 of maximum $W(\alpha)$ in the positive quadrant. In this case equation of separating hyperarea is $x^T\psi=c_0$, where

$$c_0=\frac{\min\limits_{x_i\in X}x_i^T\psi_0+\max\limits_{\overline{x}_j\in\overline{X}}\overline{x}_j^T\psi_0}{2}.$$

We will maximize quadratic form $W(\alpha)$ in the positive quadrant using a modified method of conjugate gradients, where we take into account that

$$\hat{g}_i(t)=\begin{cases}1-\psi^T\left(x_i-\overline{x}_j\right), & if \quad \alpha_{ij}\ge 0 \quad or \quad 1-\psi^T\left(x_i-\overline{x}_j\right)>0;\\ 0 \qquad if \quad \alpha_{ij}=0 \quad and \quad 1-\psi^T\left(x_i-\overline{x}_j\right)\le 0.\end{cases}$$

and that $z^TAz=\psi^T\psi$.

Let us determine a maximum point $W(\alpha)$ through iterations. For the first iteration, we

point out the group of Z_1 vectors $z_{ii} = x_i - \overline{x}_i$, made up of l_1 vectors $x_1,...,x_{l_1}$ of learning consequence which belong to the first class, and l_1 vectors $\overline{x}_1,...,\overline{x}_{l_1}$ of learning consequence which belong to the second class. By vectors z_{ii} we build a quadratic formula $W(\alpha)$, find its maximum point α_{ii}^{01} in the positive quadrant, and determine a correspondent vector $\psi_{01} = \sum_{i=1}^{l_1} \alpha_{ii}^{01}(x_i - \overline{x}_i)$.

To receive the second iteration we form a group of differences Z_2. For this purpose, we pick out among the vectors of learning consequence such vectors as x_* and $\overline{x}_*$ where extreme values are achieved:

$$x_*^T\psi_{01} = \min_{x_i \in X} x_i^T\psi_{01},\ \overline{x}_*^T\psi_{01} = \max_{\overline{x}_j \in \overline{X}} \overline{x}_j^T\psi_{01}.$$

If it turns out, that inequities done

$$x_*^T\psi_{01} \geq \min_{x \in X} x^T\psi_{01} - \delta_1, \qquad (19)$$

$$\overline{x}_*^T\psi_{01} \leq \max_{\overline{x} \in \overline{X}} \overline{x}^T\psi_{01} + \delta_2, \qquad (20)$$

where in the right members minimum and maximum are computed only by vectors of teaching sequence, δ_1 and δ_2 are algorithm parameters, vector ψ_{01}, and number $c_0 = \frac{\min_x x^T\psi_{01} + \max_{\overline{x}} \overline{x}^T\psi_{01}}{2}$ assign optimal separation of hyperarea. If only one of the inequities (19), (20) is not performed, we form a vector $z_{**} = x_* - \overline{x}_*$, add it to the distinguished group Z_2, and execute a new iteration (i.e. form a new quadratic form $W(\alpha)$, find its maximum point α_{02}^* and determine vector ψ_{02}).

We will keep on until either both inequities are resolved (19), (20), or it turns out that separation is impossible. ($W(\alpha) > W_0$, W_0 — assigned value).

Using this analyzed basic algorithm separating hyperarea is built, which minimizes the number of incorrectly classified vectors. Principally this problem can be solved in a precise manner but requires much more enumeration. That is why to form "close to the optimal" hyperarea, a standard method of "sequential minimization" is used.

If accurate separation by hyperarea is impossible, one element that "hinders mostly to separation" is excluded from the learning sequence. If separation is still impossible, another element is excluded from the remaining multitude. All in all, having excluded m vectors that hinder separation, it is possible to separate a multitude of the remaining vectors.

According to the admitted assessment, chances of incorrect classification using formed hyperarea are assessed from above:

$$P \leq \frac{d\left(\ln\frac{l}{d}+1\right) - \ln\eta}{2l}\left(1 + \sqrt{1 + \frac{4m}{d\left(\ln\frac{l}{d}+1\right) - \ln\eta}}\right) + \frac{m}{l},$$

where $d = \min\left(n, \left[\frac{D^2}{\rho^2}\right] + 1, r\right)$ is the "internal dimension of the problem" (n is space dimension, D is multitude diameter, $\rho_\varepsilon = \frac{1}{\sqrt{2W(\alpha_\varepsilon)}}$, r is several informative vectors x_i^*, $\overline{x}_j^*$).

Therefore, algorithm singularity is based on the vector of learning sequence determination x_* or $\overline{x}_*$, which "hinders separation the most". In the capacity of such vector, the vector x_* or $\overline{x}_*$ is selected which is, at the time of the halt, according to the condition $W(\alpha_*) > W_0$ contributing the most to the dimension $W(\alpha^*) = \sum_{i=1}^{a}\sum_{j=1}^{b}\alpha_{ij}^* - \frac{1}{2}\psi_*^T\psi_*$, where $\psi_* = \sum_{i=1}^{a}\sum_{j=1}^{b}\alpha_{ij}^*(x_i - \overline{x}_j)$.

4. Conclusions

Unlike known methods of managing critical infrastructure IS security, which does not envisage separation of order assessment

system including operational characteristics of IS components, and as a result, use the sequential algorithm of the assessment process, which leads to the big mathematical complexity of calculation. The proposed IS security management method is based on internal cyberattack data. As a basis, a method of support vector modification based on parameters, which are specific for internal cyberattacks on IS. This method is designed to distribute and identify the input parameters of internal cyberattacks and increase the reliability of making a management decision to assess the state of IP security, provided that the time for making a management decision will not exceed the time required by existing methods. The employed mathematical apparatus reduces the amount of input data for managing the state of the information system security and increases the reliability of making a management decision on assessing the state of security of information resources of critical infrastructure.

References

[1] V. Grechaninov, et al., Formation of Dependability and Cyber Protection Model in Information Systems of Situational Center, in: Workshop on Emerging Technology Trends on the Smart Industry and the Internet of Things, vol. 3149 (2022) 107–117.

[2] P. Anakhov, et al., Increasing the Functional Network Stability in the Depression Zone of the Hydroelectric Power Station Reservoir, in: Workshop on Emerging Technology Trends on the Smart Industry and the Internet of Things, vol. 3149 (2022) 169–176.

[3] V. Grechaninov, et al., Decentralized Access Demarcation System Construction in Situational Center Network, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems II, vol. 3188, no. 2 (2022) 197–206.

[4] D. Biryukov, Protection of Critical Infrastructure: Problems and Prospects of Implementation in Ukraine,NISD, Kyiv (2012).

[5] O. Dovgan, Critical Infrastructure as an Object of Protection Against Cybernetic Attacks, Information Security: Challenges and Threats of Modernity: Materials of a Scientific and Practical Conference (2013) 17–20.

[6] S. Hnatiuk, Criteria for Determining the Elements of the Critical Infrastructure of the State, Innovative Potential of World Science—21st Century (2013) 55–57.

[7] S. Toliupa, Intrusion Detection Systems and Functional Stability of Distributed Information Systems Against Cyber Threats, Brailovsky: Monograph, Format (2021).

[8] S. Salnyk, A. Storchak, A. Mykytyuk, Model of Violation of the Security of Information Resources of Communication Systems, Inf. Technol. Secur. 7(1) (2019) 25–34.

[9] V. Buriachok, et al., Invasion Detection Model using Two-Stage Criterion of Detection of Network Anomalies, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, vol. 2746 (2020) 23–32.

[10] I. Kuzminykh, et al., Investigation of the IoT Device Lifetime with Secure Data Transmission, Internet of Things, Smart Spaces, and Next Generation Networks and Systems, vol. 11660 (2019) 16–27. doi: 10.1007/978-3-030-30859-9_2.

[11] S. Toliupa, O. Pliushch, I. Parkhomenko, Construction of Attack Detection Systems in Information Networks Based on Neural Network Structures, Cybersecur. Educ. Sci. Technol. 2(10) (2020) 169–183. doi: 10.28925/2663-4023.2020.10.169183.

[12] A. Storchak, S. Salnyk, A Method of Assessing the Level of Security of the Network Part of a Special Purpose Communication System Against Cyber Threats, Inf. Proces. Syst. 3(158) (2019) 98–109.

[13] L. Slipachuk, S. Toliupa, V. Nakonechnyi, The Process of the Critical Infrastructure Cyber Security Management using the Integrated System of the National Cyber Security Sector Management in Ukraine, 3rd Int. Conf. Adv. Inf. Commun. Technol. (2019) 451–454.

[14] S. Toliupa, I. Parkhomenko, H. Shvedova, Security and Regulatory Aspects of the Critical Infrastructure Objects Functioning and Cyberpower Level

Assesment, 3rd Int. Conf. Adv. Inf. Commun. Technol. (2019) 463–468.
[15] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press (2000).
[16] V. Kecman, Learning and Soft Computing — Support Vector Machines, Neural Networks, Fuzzy Logic Systems, The MIT Press, Cambridge, MA (2001).
[17] A. Ben-Hur, et al., Support Vector Clustering, J. Mach. Learn. Res. 2 (2001) 125–137.