MEST Journal
DOI 10.12709/mest.03.03.02.09

MODELS, METHODS AND INFORMATION TECHNOLOGIES OF PROTECTION OF CORPORATE SYSTEMS BASED ON INTELLECTUAL IDENTIFICATION OF THREATS

Valery Lahno
Dnipropetrovsk National University of Railway Transport named after Academician V. Lazaryan, "Technical Cybernetics" Faculty, Dnipropetrovsk, Ukraine

JEL Category: **C02**

Abstract

The article to contain results of the researches, allowing to raise level of protection of the automated and intellectual information systems of the motor transportation enterprises (AISTE) in the conditions of an intensification of transportations. The article also considers the issues of research and protection of the AISTE under the condition of several conflict data request threads. The system approach to solving problems of information security, proposed in this work provides for the integration of mathematical models of the processing and protection of information. This model connects invulnerability and flexibility for each of three aspects of security (confidentiality, availability and integrity) of information based on structural unification of these contradictions. In article results of researches on development of methods and models of intellectual recognition of threats to information systems of transport. The article to contain mathematical models and results of an estimation information systems having Internet connection through various communication channels.

Keywords: Protection of information, data processing system, security policy, mathematical models

1 INTRODUCTION

The influence of information automation systems pervades many aspects of everyday life in most parts of the world. In the shape of factory and process control systems, they enable high productivity in industrial production, transport systems they provide the backbone of technical civilization. One of the foremost transport businesses security concerns is the protection of critical information, both within their internal financial infrastructures and from external elements. Now more and more open and standardized Internet technologies (e-business, e-logistics, e-cargo etc.) are used for that purpose.

Address of the author:
Valery Lahno
lva964@gmail.com

The focus on cyber security is increasing rapidly due to many high profile and highly disruptive/damaging security breaches threatening financial and physical damage across critical national and corporate infrastructures. It also appears the nature of the threat is changing (Ahmad, Dubrovskiy, & Flinn, 2005).

The automated systems on transport vary in technologies applied, from basic management systems such as car navigation; traffic signal control systems; container management systems; variable message signs; automatic number plate recognition or speed cameras to monitor applications, such as security CCTV systems; and to more advanced applications that integrate live data and feedback from a number of other sources, such as parking guidance and information systems; weather information; and the like.

A Transportation Management System (TMS) of "Ukrzaliznytsia" (The State Administration of Railway Transport of Ukraine) is a software system designed to manage transportation operations. TMS are one of the systems managing the supply chain. They belong to a sub-group called Supply chain execution (SCE). TMS, whether it is part of an Enterprise Level ERP System and has become a critical part of any (SCE).

The modern approach to ensure the reliability of information processes (IP) and its protection from unauthorized access (UA) is supported at the international level by standard ISO/IEC 15408 (ISO/IEC 15408-1:2009, 2009). According to this approach, a reliable IP successfully counteracts to the specified threats of security at the given external conditions of its operation. This leads to continuous improvement as ways and means of information protection (MIP) as well as ways and means of implementation of threats to information security (IS), resulting that appearance of new MIP leads to its bypassing by means of attack (Trivedi, Kim, & Arpan, 2001).

The purpose of the article - description of the method and models of recognition of information security threats, which, unlike the existing permit to take a final decision on the existence of a threat to existing and new classes of attacks against information systems (Chi, Park, Jung, & Lee, 2001).

2 PREVIOUS RESEARCHES

To evaluate security of such a system, a security analyst needs to take into account the effects of interactions of local vulnerabilities and find global vulnerabilities introduced by interactions. This requires an appropriate modeling of the system. Important information such as the connectivity of elements in the system and security related attributes of each element need to be modeled so that analysis can be performed. Analysis of security vulnerabilities, the most likely attack path, probability of attack at various elements in the system, an overall security metric etc. is useful in improving the overall security and robustness of the system. Various aspects which need to be considered while deciding on an appropriate model for representation and analysis are: ease of modeling, scalability of computation, and utility of the performed analysis. The analysis of the protection of information systems and automated control systems for transport companies has yielded the following results (period 2012 -2014), fig. 1, 2 (Kolodgy, 2014).

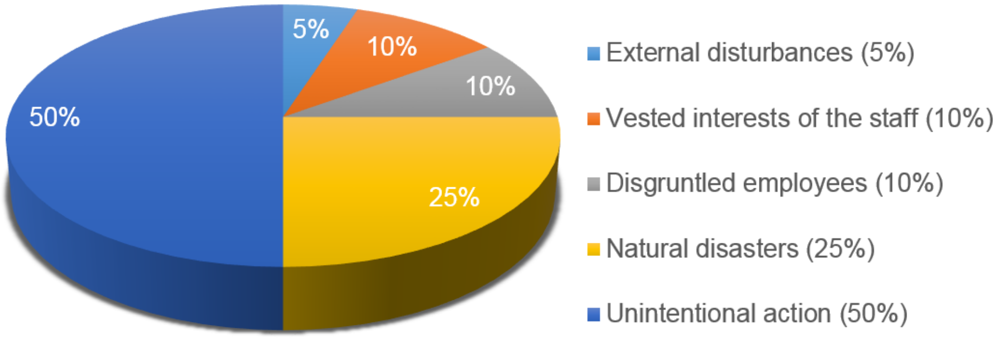

Fig 1. The distribution of sources breach AIS

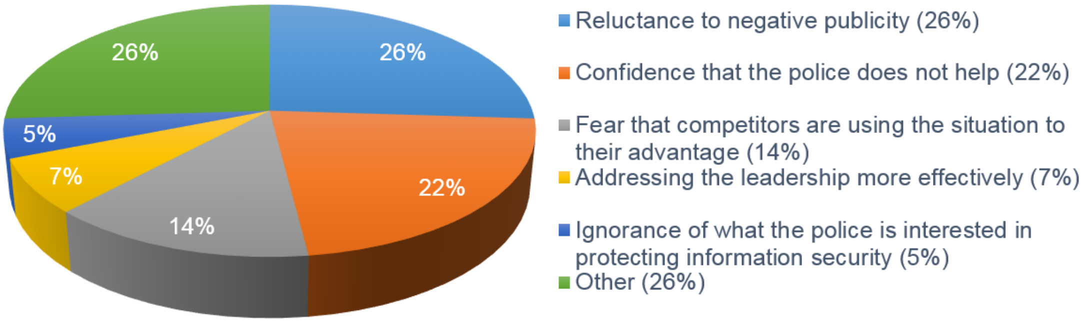

Fig. 2. The reasons for silence with information security incidents

The analysis of the threat to an automated information system must include an analysis of the vulnerabilities associated with the system environment. The goal of this step is to develop a list of system vulnerabilities (flaws or weaknesses) that could be exploited by the potential threat-sources. Vulnerability (Mirkovic, Dietrich, Dittrich, & Reiher, 2004), (Chi, Park, Jung, & Lee, 2001): A flaw or weakness in system security procedures, design, implementation, or internal controls that could be exercised (accidentally triggered or intentionally exploited) and result in a security breach or a violation of the system's security policy.

3 MODELS, METHODS AND INFORMATION TECHNOLOGIES OF PROTECTION OF CORPORATE SYSTEMS OF TRANSPORT BASED ON INTELLECTUAL IDENTIFICATION OF THREATS

The main task of discrete recognition and vulnerability search procedures (DRVSP) building is search of informative sub descriptions (or description fragments) of objects (Lahno & Petrov, 2011).

We consider informative objects to be the objects that reflect certain regularities in description of objects used for training, that is presence or, vice versa, absence of these fragments in the object, which is being considered, allows attributing it to one of classes. The fragments that are met in descriptions of one-class objects and cannot be met in descriptions of other classes' objects are considered to informative in DRVSP. The regarded fragments as a rule have a substantial description in terms of designing information safety systems (ISS).

An elementary classifier is understood as a fragment in a description of a training sample. A certain multitude of elementary classifiers with preset properties are built for each $\{KL_1,...,KL_l\}$ class. Each of such objects is not "typical" for its class, as it resembles to descriptions of objects belonging to other classes. Presence of untypical objects extends the length of fragments used to distinguish objects belonging to different classes. Long fragments are less frequent in new object, thus extending the number of unrecognized objects.

The necessity of building effective realizations for discrete recognition and vulnerability search procedures is directly connected to problems of metric (quantitative) characters of informative fragments' multitudes. The most important and technically complex are the problems of obtaining asymptotical estimates for typical number values of (impasse) covering and the length of integer matrix (impasse) covering and also the problems of obtaining analogical estimates for permissible and maximum conjunctions of a logical function, which are used for synthesis of circuit hardware-based ISS solutions.

There is, as a rule, no reliable information about the structure of *PA* (*PA* - the number of possible targets offender) multitude available while solving tasks connected with projecting an effective AIS (AIS - Automated and intellectual information systems) information safety system, that's why having built a discrete recognition and vulnerability search procedures algorithm we cannot guarantee its high performance on new objects different from

$\{sp_{a1},...,sp_{am}\}$. Nevertheless, if the training samples are quite typical for the considered multitude of objects, than the algorithm that makes infrequent mistakes in studies will show acceptable results with unknown (not included in training samples) objects also. In this connection, correctness of discerning algorithm is the problem that should be paid great attention. The algorithm is considered to be correct if it discerns all the training samples correctly.

The simplest example of a correct algorithm is the following: the considered object sp_{an} is compared to descriptions of every training sample $\{sp_{a1},...,sp_{am}\}$. In case if the sp_{an} object's description coincides with a description of a sp_{an} training sample, the sp_{an} object is attributed to the same class as the $\mathrm{sp}_{\mathrm{ai}}$ object. In other case the algorithm declines to recognize the object. There is no difficulty noticing that though the foregoing algorithm is correct, it is not able to discern any object which description does not coincide with description of any training sample.

It's obvious that requirement of full coincidence in descriptions of a considered object and one of the training samples is too cautious. The analysis of informational attack varieties and types of unauthorized access to informational system resources shows that the problem of $\mathrm{sp}_{\mathrm{ai}}$ objects' proximity and their class membership can be solved basing on comparison of a multitude of their sub descriptions. This brings up a problem of choosing character subsets that would generate the sub descriptions, according to which the objects should be compared. A variant of solution for such a problem is used in an of estimation algorithm (EA) model.

Let's introduce the following symbols. Let NP_{p_a} stand for a set of r_{p_a} $,r_{p_a} \le MI$ different integer-valued characters of $\{p_{aj_1},...,p_{aj_r}\}$ kind. Proximity of $sp'_a = (\alpha p'_{a1}, \alpha p'_{a2},...,\alpha p'_{aMI})$ and $sp''_a = (\alpha p''_{a1}, \alpha p''_{a2},...,\alpha p''_{aMI})$ belonging to *PA* by the NP_{p_a} set of characters we will estimate by the following value

$$BN(sp'_a, sp''_a, NP_{pa}) = \\ = \begin{cases} 1, & if \quad \alpha \mathrm{p}'_{\mathrm{j_{ti}}} = \alpha \mathrm{p}''_{\mathrm{j_{ti}}} , \\ 0 & \mathrm{otherwise}. \end{cases} \tag{1}$$

Thus, the schematic circuit of estimation algorithm building for information safety systems is the following. The whole range of different $NP_{p_a} = \{p_{aj_1},...,p_{a_{MI}}\}$, $r_{p_a} \le MI$ type sub multitudes is picked out inside the $\{p_{a_1},...,p_{a_{jMI}}\}$ character system. Later the picked sub multitudes are named reference multitudes of the algorithm, and their whole range is designated by ΩMI.

Further let us set the following parameters:

- po_{sp_a} is a parameter characterizing significance of a sp_{ai}, *i= 1, 2,..., PA* target (object);
- $po_{NP_{pa}}$ is a parameter characterizing significance of an object belonging to a reference multitude $NP_{p_a} \in \Omega MI$.

Further comes the estimation procedure. The considered object sp_{an} is compared to every training sample sp_{ai} of every reference multitude.

A $\Gamma(sp_a, KL)$ estimation of sp_a object belonging to *KL* class is calculated for each vulnerability class of AIS KL, $KL \in \{KL_1,...,KL_l\}$ in the following way:

$$\Gamma(sp_a, \mathrm{KL}) = \\ = \frac{1}{|\,\mathrm{LW}_{\mathrm{KL}}\,|} \sum_{sp_{ai} \in KL} \sum_{NP_{p_a} \in \Omega MI} po_{sp_a} \cdot po_{NP_{pa}} \cdot BN, \tag{2}$$

where $|LW_{KL}| = |KL \cap \{sp_{a1},...,sp_{aMI.}\}|$.

The sp_{an} object is attributed to the class that has the highest estimate. In case if there are several classes with the highest estimate, discerning fails. Obviously the ready-built algorithm is not always correct. Correctness of this algorithm requires compliance with a linear inequalities system of the following type:

$$\begin{aligned}
&\Gamma(sp_{a1},KL_1) > \Gamma(sp_{a1},KL_2),\\
&\Gamma(sp_{aMI_1},KL_1) > \Gamma(sp_{aMI_1},KL_2),\\
&\Gamma(sp_{aMI_{1+1}},KL_2) > \Gamma(sp_{aMI_{1+1}},KL_1).\\
&.\quad .\quad .\\
&\Gamma(sp_{aMI},KL_2) > \Gamma(sp_{aMI},KL_1).
\end{aligned}$$

The solution of the system comes up to choice of $po_{sp_{ai}}$ *i* = 1,2,..., *PA,* and $po_{NP_{pa}}$, $NP_{p_a} \in \Omega MI$ parameters. In case if the system is not combined, its maximum combined subsystem should be found and the solution of this subsystem defines the parameter points for $po_{sp_{ai}}$ and $po_{NP_{pa}}$.

Another way of assuring the algorithm correctness is choosing a “good” system of reference multitudes. This means, that the system should be chosen in a way assuring that each $sp'_a \notin KL$ training sample meets the $\Gamma(sp'_a, KL) = 0$ condition and each $sp''_a \in KL$ training sample meets the condition of $\Gamma(sp''_a, KL) > 0$. This can be achieved in the following way.

If $NP_{p_a} = \{p_{aj_1},...,p_{a_{MI}}\}$ is a reference multitude, than the NP_{p_a} character set should be named a test, in case if the $BN(sp'_a, sp''_a, NP_{pa}) = 0$ equation is valid for all sp'_a, sp''_a, training objects belonging to different classes. In other words, a test is a number of characters which allows discerning any two objects of different classes.

It is appropriate to mention here that presently the most aggressive way of checking an AIS information protection system effectiveness for unauthorized access is a penetration test. While doing such check, a test applies every possible way of bypassing the mechanisms of AIS protection, which can be used by transgressors of safety policy. Results of penetration tests are analyzed, allowing to raise effectiveness of the information protection system and also to eliminate all the vulnerabilities that were found. Carrying out penetration tests is one of the most important procedures for raising general information safety of an enterprise or corporation in the countries of European Union or the USA. The penetration test model is regulated, in a number of states, by the organs responsible for licensing and attesting in the sphere of information protection.

Let ΩMI_T stand for some range of tests. If the range of reference multitudes for the algorithm consists of tests, then it’s obvious, that such algorithm is correct in all cases when the $po_{sp_{ai}}$ where *i* = 1, 2,..., *PA,* and $po_{NP_{pa}}$, $NP_{p_a} \in \Omega MI$ parameters have positive values.

If the $NP_{p_{a1}}$ character set is a test, then any $NP_{p_{a2}}$ character set corresponding to $NP_{p_{a1}} \subset NP_{p_{a2}}$ is also a test. At the same time, if the objects are close in $NP_{p_{a2}}$, they would be close in $NP_{p_{a1}}$ also. If the objects are close in $NP_{p_{a1}}$ set of columns, they will always be close in $NP_{p_{a2}}$. The shorter tests are more informative in this respect and it’s reasonable to restrict the test length (that is character sets) or to build terminal tests.

The NP_{p_a} character set can be named a terminal test in case if it meets the following two conditions:

1. NP_{p_a} is a test (it is a set of characters, that allows to reveal vulnerabilities of a system);
2. any own sub multitude of the NP_{p_a} set is not a test itself.

In other words, a terminal test is an unshortenable set of characters, which discerns any two training samples belonging to different classes of information safety threats $B_{p_a k1}, B_{p_a k2}$.

Let each $p_{axj}, j = 1,2,...,n$ character have a terminal *PA* multitude of legitimate values.

Let $NP_{p_a} = \{\, p_{ax_{j1}},..., p_{ax_{jr}} \,\}$ stand for some character set, and let $sp_a = (\alpha p_{a1}, \alpha p_{a2},..., \alpha p_{an})$ be an object of a training sample. Let’s designate the $(\alpha p_{aj1},...,\alpha p_{ajr})$ fragment in the object's description by (sp_a, NP_{p_a}).

Each NP_{p_a} test causes numerous description fragments of the following type $(sp_{ai}, NP_{p_a}),\ i = 1,2,...,PA$, where $\mathrm{sp_{ai}}$ is a training sample, though each of these fragments is met in only class, and is not met in other classes. Thus, if we turn from consideration of reference multitudes to analysis of objects' fragments description, while building algorithms of discrete recognition and vulnerability search procedures; we will be able to build less cautious, but at the same time more correct procedures.

Let NP_{p_a} be a certain set of r_{p_a} different characters of $NP_{p_a} = \{ p_{ax_{j1}},...,p_{ax_{jr}} \}$ type, $\sigma_{DOP} = (\sigma_{DOP_1},...,\sigma_{DOP_r})$, σ_{DOP_i} is a legitimate value of $p_{ax_i}, i = 1,2,...,r_{p_a}$ character. The σ_{DOP_i} set is an elementary classifier, caused by characters from NP_{p_a}. Proximity of the $sp_{an} = (\alpha p_{a1}, \alpha p_{a2},...,\alpha p_{aMI})$ object of *PA* and the $\sigma_{DOP} = (\sigma_{DOP_1},...,\sigma_{DOP_r})$ elementary classifier, caused by a set of characters from NP_{p_a} should be estimated by the following value:

$$BN(\sigma_{\mathrm{DOP}}, sp_a, NP_{pa}) = \\ = \begin{cases} 1, & if\ \alpha \mathrm{p}_{\mathrm{j_{ti}}} = \sigma_{DOP\,ti} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

A multitude of all elementary classifiers, caused by character sets from $\{p_{ax1},...,p_{axn}\}$, should be designated by *MC*. Thus, $MC = \{(\sigma_{DOP}, NP_{pa})\}$,

where $NP_{pa} \subseteq \{p_{ax1},...,p_{axn}\}$, $NP_{p_a} = \{p_{ax_{j1}},...,p_{ax_{jr}}\}$, $\sigma_{DOP} = (\sigma_{DOP_1},...,\sigma_{DOP_r})$, $\sigma_{DOP_i} \in NP_{p_{aj}},\quad npu\ i = 1,2,...,r_{p_a}$.

Each *AL* algorithm of information threat recognition builds a certain $MC^{AL}(KL)$ submultitude of *MC* multitude for each $KL,\quad KL \in \{KL_1,...,KL_l\}$ class.

Let's designate

$$MC^{AL} = \bigcup_{j=1}^{l} MC^{AL}(KL_j).$$

Discerning of a sp_{an} object is carried out on the basis of calculating $BN(\sigma_{DOP}, sp_a, NP_{pa})$ value for each (σ_{DOP}, NP_{pa}) element of the $MC^{AL}(KL), KL \in \{KL_1,...,KL_l\}$ multitude. That means that the procedure of $\Gamma(sp_a, KL)$ value estimation of sp_a object's belonging to *KL* class is carried out for each element of the multitude. Thus, each discerning *AL* algorithm of the regarded family is determined by a $MC^{AL}(KL)$ multitude of elementary classifiers and by the $\Gamma(sp_a, KL)$ way of value estimation.

Generally, a $\sigma_{DOP} = (\sigma_{DOP_1},...,\sigma_{DOP_r})$ elementary classifier, caused by characters of NP_{pa}, can have one of the following three properties:

1. each fragment of (sp'_a, NP_{pa}), type, where $sp'_a \in KL$, coincides with $\sigma_{DOP} = (\sigma_{DOP_1},...,\sigma_{DOP_r})$;
2. only some fragments of (sp'_a, NP_{pa}) type, where $sp'_a \in KL$, coincides with $\sigma_{DOP} = (\sigma_{DOP_1},...,\sigma_{DOP_r})$;
3. neither of (sp'_a, NP_{pa}) type fragments, where $sp'_a \in KL$, coincides with $\sigma_{DOP} = (\sigma_{DOP_1},...,\sigma_{DOP_r})$.

The first situation is rather uncommon and meets seldom, that is why working with character value sets, which meet the first characteristic, is considered to be impossible. Considerable difference in self-descriptiveness of the following two features consists in the fact that the second feature characterizes only a training samples sub multitude of *KL*, and the third feature characterizes all the *KL* objects. Therefore, in case, when it's important to regard the *KL* class separately from other classes, there automatically comes a conclusion that the character sets, which comply with the third feature, are more informative. In the stated case it's more natural to consider the situation, when the set of character value is not present in all objects of *KL* class and is not also present in the spa object, as an argument for referring the discerned sp_a object to *KL* class.

The methods of building elementary classifiers σ_{DOP_i} for *KL* class in classic models is based on building a σ_{DOP_i} matrix covering, created by training samples' descriptions of each *KL* class. Usage of such models [7] allows to reduce the calculation expenditure in case if $|KL|<|\overline{KL}|$, for example when there is a large number of information threat classes – $\{KL_1,...,KL_l\}$ and $\{B_{p_{a1}},...,B_{p_{al}}\}$.

We suggest using the method of typical sp_a objects isolation basing on the procedure of sliding control, which is the following.

One sp_{a_i}, $i\in\{1,2,...,PA\}$ object should be excluded from the training samples. A discerning algorithm is built for the rest of the $\{sp_{a_1},...,sp_{a_{PA}}\}\backslash sp_{a_i}$ samples. Later this algorithm is used for discerning the sp_a object. The sp_a object should be considered typical for its class, if the algorithm refers it to another class or declined recognizing it. The described procedure should be repeated for all the training sample objects.

Let the training samples be divided into basic and control subsamples. A multitude of representative sets should be built for the basic subsamples. Later some weight, which is calculated with the help of the control subsamples, should be compared for each representative set.

Let $p\omega$ stand for the representative set of the $KL, KL\in\{KL_1,...,KL_l\}$ class, caused by the (sp'_a,NP_{pa}) pair, where sp'_a is an object of the basic samples. And let $\delta n(\mathrm{KL},\mathrm{p}\omega)$ be the number of sp_{a_i} objects (which are the malefactor's targets) in the control samples, for which the representative set "votes correctly", while $\delta n(\overline{KL},p\omega)$ is the number of the control samples objects, for which the representative set "votes incorrectly". Then the following functions can be regarded as functions of the $vop_{(sp'_a,NP_{pa})}$ elementary classifier's significance:

$$vop_1(sp'_a,NP_{pa})=\delta n(KL,p\omega),$$
$$vop_2(sp'_a,NP_{pa})=\frac{1+\delta n(KL,p\omega)}{1-\delta n(\overline{KL},p\omega)}. \tag{4}$$

The sp_{a_i} object's belonging to *KL* class will be estimated by the following value:

$$\Gamma(sp_a,\mathrm{KL})=\frac{1}{|\mathrm{MC}^{\mathrm{AL}}(KL)|}\cdot \sum_{(sp'_a,NP_{pa})\in\mathrm{MC}^{\mathrm{AL}}(KL)}\mathrm{vop}_{(sp'_a,NP_{pa})}\cdot(1-BN)). \tag{5}$$

We will consider the following value as an informative significance of the p_{axj} character

$$\mathrm{IZ}_{\mathrm{p_{axj}}}=\frac{\sum\limits_{\substack{(sp'_a,NP_{pa})\in\mathrm{MC}^{\mathrm{AL}}(KL)\\ p_{axj}\in NP_{pa}}}\mathrm{vop}_{(sp'_a,NP_{pa})}}{\sum\limits_{\substack{(sp'_a,NP_{pa})\in\mathrm{MC}^{\mathrm{AL}}(KL)\\ p_{axj}\in NP_{pa}}}\mathrm{vop}_{(sp'_a,NP_{pa})}}. \tag{6}$$

This part of the work sets forth the basic principles of discrete recognition and vulnerability search procedures construction, using the apparatus of logic functions that allows bringing to practice effective circuit solutions of information protection for automatic systems.

Let's regard the situation, when the objects of the considered PA multitude are described by the characters, each possessing values of the {0, 1,..., k_{p_a} *- 1*} multitude.

Let's associate the (σ_{DOP}, NP_{pa}) elementary classifier, where $\sigma_{DOP}=(\sigma_{DOP_1},...,\sigma_{DOP_r})$, NP_{pa} is a set of characters numbered $j_1,...j_{r_{pa}}$, with an elementary conjunction $\mathfrak{R}=p_{axj_1}^{\sigma_{DOP_1}}...p_{axj_{r_{pa}}}^{\sigma_{DOP_{r_{pa}}}}$.

If $sp_a=(\alpha p_{a1},...,\alpha p_{aMI})$ is an object of the *PA* multitude, then obviously $BN(\sigma_{DOP},sp_a,NP_{pa})=1$ only in case when $(\alpha p_{a1},...,\alpha p_{aMI})\in NI_{\mathfrak{R}}$, where $NI_{\mathfrak{R}}$ is a truth interval for the elementary conjunction $\mathfrak{R}$.

Let's show that building a multitude of $(KL_l)=(B_{p_{al}})$ class elementary classifiers for the models previously considered in the article adds up to finding permissible and maximum conjunctions of the characteristic $(KL_l)=(B_{p_{al}})$ class function, which is a double-valued logical function possessing different values for training samples of KL_l and $\overline{KL_l}$.

Table 1. The knowledge base for the intelligent recognition of threats to information systems

Attributes (Signs Class threats)	Signs Class threats	The importance of sign	The universum	Terms for the linguistic evaluation $\phi_u,\dots,\phi_v$
The set of classes of information security threats $KL=\{KL_1,...,KL_n\}$, The set targets for attack $PA=\{PA_1,...,PA_z\}$, The set of information security $N_j^{p_a}=\{n_1^{p_{a1}},...,n_j^{p_{au}}\}$, The mathematical sets of possible attackers $U=\{u_1,...,u_g\}$, The sets of incidents $NIS=\{nis_1,...,nis_f\}$, The sets of variants attack on the system $AT=\{AT_1,...,AT_q\}$, and others.	$p_{ax}=$ $=\{p_{ax1},...,$ $p_{ax_{MI}}\}$.	based on *NIS* $-1\le IZ_{p_{axj}}$ ≤ 1	$[0,N_a]$ or [0,1], c. u.	Critical and uncritical *or* Identified, partially identified threats, undiag-nosed
The state systems (AIS) $S_{IK}=\{S_{IK_1},...,S_{IK_m}\}$				
Methods and means of protection of information systems $D_{33i}=\{D_{33i_1},...,D_{33i_r}\}$				
The rules for result output *IF* $(KL_1\vee...\vee KL_n\vee S_{IK_J}\vee...\vee S_{IK_m})$ *THEN* D_{33i_r} and $\mu^{d_j}(S_{IK_i})=\overset{h_j}{\underset{p=1}{\vee}}[\mu^{y_1}(y_1)\wedge...\wedge\mu^{\phi_v}(\phi_v)]$, $p=\overline{1,h_j}$, $j=\overline{1,MI}$, де $\mu^{y_1}(y_1),\dots,$ $\mu^{\phi}(\phi_u)$, $\mu^{\phi}(\phi_v)$ – membership function *y₁* , $\phi_u,\dots,\phi_v$ of the fuzzy variables to terms; *y₁* – the state of information security {below critical, critical, above the critical, high}; $\vee$ – logical **OR**, $\wedge$ - Logical **AND** as operations max and min, respectively.				

For example, a system of logical equations for intelligent recognition of DDoS-attacks Application layer ("slow" HTTP GET flood and "slow" HTTP POST flood), we can write this:

$$\mu^{d_j}(\mathrm{S}) = \bigvee_{p=1}^{h_j}\left[\mu^{y_1^{jp}}(y_1) \wedge \mu^{\phi_{\text{я}}^{jp}}(\phi_{13}) \wedge \mu^{\phi_{\text{я}}^{jp}}(\phi_{14})\right],$$

$$p = \overline{1, h_j}\ ,\ j = \overline{1, M}\ , \qquad (7)$$

where $\mu^{y_1^{jp}}(y_1)$, $\mu^{\phi_{\text{я}}^{jp}}(\phi_{13})$, $\mu^{\phi_{\text{я}}^{jp}}(\phi_{14})$ – membership function variables *y₁* , ϕ_{13}, ϕ_{14} their fuzzy terms y_1^{jp} , ϕ_{13}^{jp} , ϕ_{14}^{jp} , respectively;

S – the state protection of information systems against DoS / DDoS (Xiang, Zhou, & Chowdhury, 2004);

y₁ – the state of information {below the critical (*bc*), critical (*cr*), above the critical (*ac*), high (*h*) (Lahno & Petrov, 2010);

$\vee$ – logical OR, $\wedge$ – logical AND, like max and min, respectively.

The main objective is to search DRVSP building fragments describing objects, see. Table 1.

The probability of detection of various attacks on the IP is based on Bayes' theorem and the knowledge base (see. table 1). As an evaluation criterion, used parameter changes in the state system (see. Equation 7).

Bayes' theorem is stated mathematically as the following equation (Daston, 1988):

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)},$$

where *A* and *B* are events:

- *A* - The threats: identified, partially identified, undiagnosed;
- *B* - A change in the system state - $\mathrm{S}_{\mathrm{IK}} = \{\mathrm{S}_{\mathrm{IK}_1}, \ldots, \mathrm{S}_{\mathrm{IK}_\mathrm{m}}\}$);

P(*A*) and *P*(*B*) are the probabilities of *A* and *B* independent of each other.

P(*A*|*B*), a conditional probability, is the probability of *A* given that *B* is true.

P(*B*|*A*), is the probability of *B* given that *A* is true.

Fig. 3 shows the main results obtained during the test simulation recognition DoS / Ddos attacks.

According to the results of the experiment, the DRVSP DoS/DDoS - attacks, following results were obtained for the errors of the first kind (false positives) - 10.2% for the error of the second kind (the number of detected attacks) - 2.9%.

Thus, building a multitude of elementary classifiers for the simulated class of information treats adds up to the following:

1. specifying a characteristic function;
2. building a disjunctive normal form, which realizes this function. The biggest difficulty is building disjunctive normal forms from maximum conjunctions (shortened disjunctive normal forms) of a characteristic function;
3. calculating a permissible (maximum) conjunction $\mathfrak{R}$, which determines of the object belongs to a certain class of threats (Lahno & Petrov, 2012).
4. For each class, the number of threats to information security signs ranged from 3 to 9. Informational content of a sign can change in the range from -1 to +1. To assess the DRVSP used method of cross-validation. The results of validation of the method DRVSP shown in Fig. 4 -6.

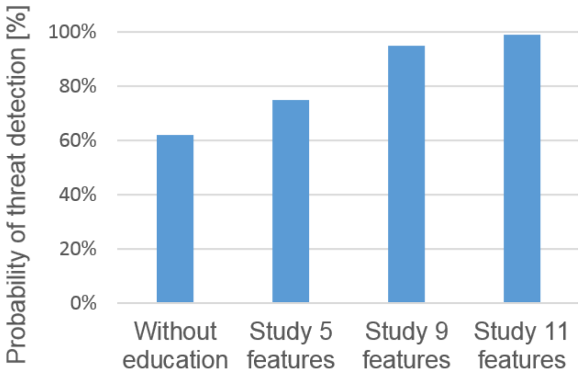

Fig 3. The probability of detecting DDoS attacks

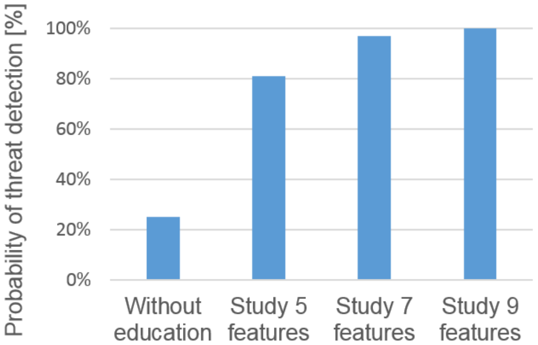

Fig. 4. The probability of recognizing the threat of “Unauthorized access to the user's password”

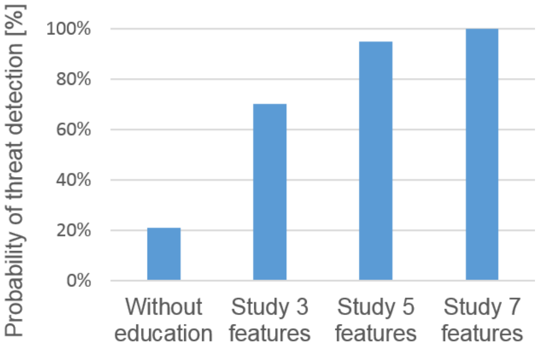

Fig. 6. The probability of recognizing the threat of “Unauthorized access to the navigation system”

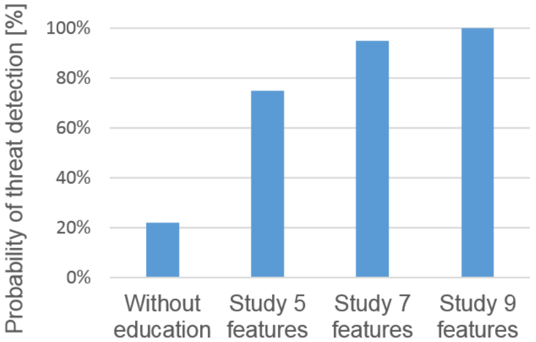

Fig. 5. The probability of recognizing the threat of “Unauthorized access to software and databases”

4 CONCLUSIONS

Operation is devoted to research and development of theoretical methods, models and software products for support of information security on transport.

The method of intellectual recognition of threats based on logic functions and indistinct sets is developed. The method allows increasing efficiency of recognition of threats for information security to 85-98% (depending on a threat class). It is possible, also to use a method for creation of new systems of information security on transport.

WORKS CITED

Ahmad, D., Dubrovskiy, A., & Flinn, X. (2005). *Defense from the hackers of corporate networks.* Moscow: Companies AyTi; DMK - Press.

Chi, S., Park, J., Jung, K., & Lee, J. (2001). *Network Security Modeling and Cyber At-tack Simulation Methodology* (Vol. 2119). LNCS. Vol. 2119. Retrieved from http://link.springer.com/chapter/10.1007%2F3-540-47719-5_26

Daston, L. (1988). *Classical Probability in the Enlightenment.* Princeton Univ Press.

ISO/IEC 15408-1:2009. (2009). *Information technology - Security techniques - Evaluation criteria for IT security - Part 1: Introduction and general model.* ISO.

Kolodgy, C. J. (2014). *Worldwide Security and Vulnerability Management 2004-2014.* Manchester: National Computer Center Publications.

Lahno, V., & Petrov, A. (2010). *Modelling of discrete recognition and information vulnerability search procedures* (Vol. XI A). TEKA.

Lahno, V., & Petrov, A. (2011). *Ensuring security of automated information systems, transportation companies with the intensification of traffic.* Lugansk.

Lahno, V., & Petrov, A. (2012). Modeling information security system of transport enterprises. In H. H. Marek Dudek (Ed.), *Management and production engineering* (pp. 221-248). ATH. Academia Techniczno – Humanistyczna of Bielsko-Biala.

Mirkovic, J., Dietrich, S., Dittrich, D., & Reiher, P. (2004). *Internet Denial of Service: Attack and Defense Mechanisms.* Prentice Hall PTR.

Trivedi, K. S., Kim, D. S., & Arpan, R. (2001). *Dependability and Security Models.* Durham, NC, USA: Duke University Department of Electrical and Computer Engineering .

Xiang, Y., Zhou, W., & Chowdhury, M. (2004). *A Survey of Active and Passive Defence Mechanisms against DDoS Attacks. Technical Report, TR C04/02.* Australia: School of Information Technology, Deakin University.

Received for publication:	14.02.2015
Revision received:	03.03.2015
Accepted for publication:	10.05.2015

How to cite this article?

Style – **APA** *Sixth Edition:*

Lahno, V. (2015, July 15). Models, methods and information technologies of protection of corporate systems Based on intellectual identification of threats. (Z. Čekerevac, Ed.) *MEST Journal, 3*(2), 79-89. doi:10.12709/mest.03.03.02.09

Style – **Chicago** *Sixteenth Edition:*

Lahno, Valery. 2015. "Models, methods and information technologies of protection of corporate systems Based on intellectual identification of threats." Edited by Zoran Čekerevac. *MEST Journal* (MESTE) 3 (2): 79-89. doi:10.12709/mest.03.03.02.09.

Style – **GOST** *Name Sort:*

Lahno Valery Models, methods and information technologies of protection of corporate systems Based on intellectual identification of threats [Journal] // MEST Journal / ed. Čekerevac Zoran. - Belgrade : MESTE, July 15, 2015. - 2 : Vol. 3. - pp. 79-89.

Style – **Harvard** *Anglia:*

Lahno, V., 2015. Models, methods and information technologies of protection of corporate systems Based on intellectual identification of threats. *MEST Journal,* 15 July, 3(2), pp. 79-89.

Style – **ISO 690** *Numerical Reference:*

Models, methods and information technologies of protection of corporate systems Based on intellectual identification of threats. **Lahno, Valery.** [ed.] Zoran Čekerevac. 2, Belgrade : MESTE, July 15, 2015, MEST Journal, Vol. 3, pp. 79-89.