Розроблено адаптивну систему виявлення кібератак, яка базується на удосконалених алгоритмах розбиття простору ознак на кластери. Удосконалена процедура розпізнавання за рахунок одночасної кластеризації та формування перевірочних допустимих відхилень для ознак аномалій та кібернападів. За допомогою імітаційних моделей, створених у MatLAB та Simulink, перевірена працездатність алгоритмів розпізнавання кібератак у критично важливих інформаційних системах

Ключові слова: система виявлення кібератак, кібербезпека, кластеризація ознак, перевірочні допустимі відхилення

Разработана адаптивная система обнаружения кибератак, основанная на усовершенствованных алгоритмах разбития пространства признаков на кластеры. Усовершенствована процедура распознавания за счет одновременной кластеризации и формирования проверочных допустимых отклонений для признаков аномалий и кибератак. С помощью имитационных моделей, созданных в MatLAB и Simulink, проверена работоспособность алгоритмов распознавания кибератак в критически важных информационных системах

Ключевые слова: система обнаружения кибератак, кибербезопасность, кластеризация признаков, проверочные допустимые отклонения

UDC 004.056
DOI: 10.15587/1729-4061.2017.102225

DEVELOPMENT OF A SYSTEM FOR THE DETECTION OF CYBER ATTACKS BASED ON THE CLUSTERING AND FORMATION OF REFERENCE DEVIATIONS OF ATTRIBUTES

V. Lakhno
Doctor of Technical Sciences, Associate Professor*
E-mail: lva964@gmail.com
V. Malyukov
Doctor of Physical and Mathematical Sciences**
E-mail: volod.malyukov@gmail.com
V. Domrachev
PhD, Associate Professor
Department of Applied Information System
Taras Shevchenko National University of Kyiv
Volodymyrska str., 60, Kyiv, Ukraine, 01033
E-mail: vlad.mipt@gmail.com
O. Stepanenko
Doctor of Economic Sciences, Associate professor
Department of Economics Information Systems
Vadym Hetman Kyiv National Economic University
Peremohy ave., 54/1, Kyiv, Ukraine, 03057
E-mail: olga_stepanenko@kneu.edu.ua
O. Kramarov
Postgraduate student**
E-mail: o.kramarov@gmail.com
*Department of Managing Information Security***
**Department of Information Systems and
Mathematical Sciences***
***European University
Akademika Vernads'koho blvd., 16 V, Kyiv, Ukraine, 03115

1. Introduction

Active expansion of computer technologies, in particular in critically important information systems (CIIS), is accompanied by the emergence of new threats to cyber security (CS). It is possible to enhance CS of CIIS by using, in particular, intelligent systems (and technologies) for the detection of cyber attacks (ISDA). Given a constant complication in the scenarios of cyber attacks, ISDA must have characteristics of adaptive systems. In other words, the ability to deliberately modify the algorithm for detecting the anomalies and cyber attacks by using the methods of clustering of attributes of the recognition objects (RO), as well as machine intelligent technologies of learning (MITL).

This makes it relevant to examine improvement of those existing and development of the new algorithms for the clustering of RO attributes, as well as the applied adaptive subsystems as a part of ISDA.

2. Literature review and problem statement

Information that is accepted as the basis for building the clusters in adaptive systems of recognition (ASR) of cyber attacks was explored in many studies, for example, in the form of complex attributes of RO in CIIS [1, 2]. These studies were mainly of theoretical character. As indicators or metrics [3] for building the classifiers, the authors investigated: threshold values of parameters of the input and output traffic [4], unpredicted addresses of packets [5], attributes of requests to databases (DB) [6, 7], etc. These articles do not take into account the possibility of parallel formation of reference deviations for the features of anomalies and cyber attacks, which increases the time of RO analysis in ASR (or ISDA) [8]. For complex targeted attacks, information attributes may be quite fuzzy [9, 10], which does not contribute to building the effective algorithms of recognition.

In papers [11, 12], it was assumed that to enhance effectiveness of recognition, it is expedient to split the set of values of each indicator into disjoint groups by certain rules. This task can be solved by using the methods and models for cluster analysis [13, 14]. However, these studies have not been brought to hardware or software implementation.

By using an information condition of functional effectiveness (ICFE) of ASR learning [15, 16], it is possible to implement adaptive algorithms for the clustering of RO attributes into ISDA.

As was shown in articles [17, 18], in case the RO attributes glossary is unchanged, it is possible to improve effectiveness of ASR learning. These studies do not take into account the possibility of increasing the degree of intersection of the RO classes.

Thus, given the potential of the ISDA application, it appears to be an important task to improve the algorithms for clustering and formation of reference deviations of the OR attributes for the timely detection of anomalies and cyber attacks in CIIS.

3. The aim and tasks of research

The aim of present research is to develop an algorithm for the partition of the feature space (FS) into clusters in the process of recognition of cyber attacks in the systems of cyber protection.

To achieve the aim of the study, the following tasks are to be solved:

– to improve algorithms for the clustering of attributes of anomalies and cyber attacks and for the simultaneous formation of verifying admissible deviations in the intelligent systems of cyber attack detection;

– to conduct simulation in order to test and verify the adequacy of the proposed algorithms.

4. Algorithms for the clustering of attributes and the formation of verifying admissible deviations in the intelligent systems of cyber attack detection

Splitting FS and further clustering, for any RO class CT_m^0, in accordance with [19, 20], was carried out by transforming FS to a hyper-spherical form. Since the main stage of clustering when splitting FS into groups is an increase in the radius (cr_m) of container (RC) at every step of ASR (or ISDA) leanring, it is possible to use the following recurrent expression:

$$cr_m(ls) = \left[cr_m(ls-1) + \xi \mid cr_m(ls) \in IS_m^{cr}\right], \tag{1}$$

where ls is the number of steps of increasing RC C_m^0; ξ are accepted for the chosen attributes of steps of increasing RC; IS_m^{cr} is the permissible value of RC.

In the process of ASR learning, we make an assumption about fuzzy compactness of the implementation of binary learning matrices (BLM) [16, 21, 22], obtained at the stage of splitting SF into relevant RO classes. Fuzzy partition $RC^{|M|}$ includes the elements that can be attributed to fuzzy RO classes, for example, when it is difficult to distinguish a DoS attack from a DDoS attack [4, 16].

The rules of ASR learning, according to [1, 2, 14, 23, 24], are built based on the iteration procedure of searching for the maximum boundary magnitude of an information condition of functional effectiveness (ICFE):

$$is'_k = \operatorname{Arg}\max_{IS_k}\{\max_{IS_{k-1}}\{\ldots\{\max_{IS_1 \cap IS_{CE}} \frac{1}{M}\sum_{m-1}^{M} CE_m\}\ldots\}\}, \tag{2}$$

where CE_m is the ICFE of ASR learning to recognize RO that belong to class C_m^0; IS_k is the permissible range of values of the k-th informative attribute of RO; IS_{CE} is the permissible range of ICFE in the course of ASR learning.

The following constraints are imposed on expression (2):

$$\left\{\left\lfloor CT_m^o \neq \varnothing \right\rfloor : \left(\forall CT_m^o \in RC^{|M|}\right)\right\}; \tag{3}$$

$$\left\{\left|\begin{matrix} CT_a^o \neq CT_b^o \to \\ \to CT_a^o \bigcap CT_b^o \neq \varnothing \end{matrix}\right| : \left(\exists CT_a^o \in RC^{|M|}, \exists CT_b^o \in RC^{|M|}\right)\right\}; \tag{4}$$

$$\left\{\left|\begin{matrix} CT_a^o \neq CT_b^o \to \\ \to B\,CT_a^o \bigcap BCT_b^o = \varnothing \end{matrix}\right| : \left(\forall CT_a^o \in RC^{|M|}, \forall CT_b^o \in RC^{|M|}\right)\right\}, \tag{5}$$

where BCT_a^0, BCT_b^0 are the nuclei of RO classes CT_a^0 and CT_b^0, respectively;

$$\bigcup_{CT_m^o \in RC} CT_m^o \subseteq RS_B; a \neq b; a, b, m = \overline{1, M}. \tag{6}$$

Accepted assumptions: classes CT_a^0 are CT_b^0 adjacent; the classes have a minimum distance between the centers of clusters $cr(ct_a \oplus ct_b)$ among all classes for RO; RO are described by binary learning matrices (BLM) [21–23]. We accepted that ct_a and ct_b are the reference vectors of RO classes, in particular, by the KDD Cup 1999 Data [2, 5, 7].

The ASR learning procedure is given in the form of predicate expression:

$$\left\{\left[\begin{matrix} CT_a^o \neq CT_b^o \to \\ \to \left(cr'_a < cr\left(ct_a \oplus ct_b\right)\right) \cdot \\ \cdot \left(cr'_b < cr\left(ct_a \oplus ct_b\right)\right) \end{matrix}\right] : \left(\forall CE_a^o \in RC^{|M|}, \forall CT_b^o \in RC^{|M|}\right)\right\}, \tag{7}$$

where cr'_a, cr'_b are the optimal radii of containers C_a^0 and C_b^0, respectively.

To reduce the number of cycles during a learning procedure, the sets of input signals (factors) that influence ASR were determined. These sets correlate with the dimensionality of the vector of ASR testing parameters is=<is_1,..., is_k,..., is_{RS}> in the course of recognition of the templates of attacks.

ASR (or ISDA) learning is an iteration procedure of searching for global ICFE [2, 5, 8, 20, 24]:

$$ca^* = \operatorname{Arg}\max_{IS_{ca}}\{\max_{IS_{CE} \cap IS_{cr}} \overline{CE}\}, \tag{8}$$

where IS_{ca} is the admissible range of magnitudes of reference deviation (ca) for RO class $\left\{CT_m^o\right\}$; IS_{CE} is the operation range of determining ICFE indicator $\overline{CE}$; IS_{cr} is the permissible range of RC magnitude cr.

The algorithm of OR classification is functional at the following restrictions:

$$\left\{\left\lfloor CT^{o}_{m,\xi} \neq \varnothing, m = \overline{1,M} \right\rfloor : \left(\forall CT^{o}_{m,\xi} \in RC^{|M|}\right)\right\}, \qquad (9)$$

$$\left\{\left[\begin{array}{l} CT^{o}_{m,\xi} \neq \\ \neq CT^{o}_{c,\xi} \to BCT^{o}_{m,\xi} \cap BCT^{o}_{c,\xi} = \varnothing \end{array}\right] : \left(\forall CT^{o}_{m,\xi} \in RC^{|M|}, \forall CT^{o}_{c,\xi} \in RC^{|M|}\right)\right\}, \qquad (10)$$

$$\left\{\left[\begin{array}{l} CT^{o}_{m,\xi} \\ \neq CT^{o}_{c/\xi} \to \\ \left(cr'_{m,\xi} < cr\left(ct_{m,\xi} \oplus ct_{c,\xi}\right)\right) \wedge \\ \wedge \left(cr'_{c,\xi} < cr\left(ct_{m,\xi} \oplus ct_{c,h}\right)\right) \end{array}\right] : \left(\forall CT^{o}_{m,\xi} \in RC^{|M|}, \forall CT^{o}_{c,\xi} \in RC^{|O|}\right)\right\}, \qquad (11)$$

$$\bigcup_{CT^{o}_{m,\xi} \in RC} CT^{o}_{m,\xi} \subseteq RS, \qquad (12)$$

where $BCT^{o}_{m,\xi}$, $BCT^{o}_{c,\xi}$ are the centers of the two nearest (adjacent) clusters $CT^{o}_{m,\xi}$ and $CT^{o}_{c,\xi}$, respectively; ξ is the step of increasing the radius of cluster container (RCC); $cr'_{m,\xi}$, $cr'_{c,\xi}$ are, respectively, formed RCC $CT^{o}_{m,\xi}$ and $CT^{o}_{c,\xi}$; $cr\left(ct_m \oplus ct_c\right)$ is the inter-center code distance of clusters $CT^{o}_{m,\xi}$ and $CT^{o}_{c,\xi}$.

For better visualization, the stages of splitting FS of RO into clusters in ASR are represented in tabular form in Table 1.

As a criterion of the optimization of parameters, during ASR learning, we used statistical parameters (information measures) for the variants of solutions with two alternatives [18, 25, 26] for a modified entropic indicator, as well as the Kullback-Leibler divergence (for three hypotheses) [27].

Table 1

Stages of splitting FS into clusters

Stage	Action	Description
1	2	3
1	Step counter (SC) of changing VAD ca_i by features of RO is set as "0":	$l := 0$
2	Calculation of the lower $A_{low_i}[l]$ and the upper $A_{up_i}[l]$ of VAD of RO features for entire FS	$A_{low_i}[l] = lm_i - ca\frac{ca_{low_i}}{100}$; $A_{up_i}[l] = lm_i + ca\frac{ca_{low_i}}{100}$, where lm_i is the i-th attribute of standard vector-realization of non-classified multi-dimensional matrix (NMLM) $\lVert lm_i^{(j)} \rVert$ [16, 23]; ca_{low_i} is the VAD for RO attributes, which are determined based on methods [2, 16, 21, 23]
3	Formation of BLM $\lVert ct_i^{(j)} \rVert$	Rule $ct_i^{(j)} = \begin{cases} 1, \text{ if } A_{low_i}[l] < lm_i^{(j)} < A_{up_i}[l]; \\ 0, \text{ else} \end{cases}$
4	Value of SC for increasing RC	$\xi := 0$
5	Initialization of SC for increasing RC	$\xi := 1$
6	Splitting NMLM $\left\{ct_i^{(j)}\right\}$ into two clusters	$\left\{CT^{o}_{m}[\xi] \mid m = \overline{1,\ 2}\right\}$
6.1	Initial original standard vectors for RO attributes $\{ct_m\}$ for C^0_m are calculated	Verification of conditions: 1) $cr\left(ct_1 \oplus ct^0\right) \to \min$, $cr\left(ct_2 \oplus ct^1\right) \to \min$; 2) $cr\left(ct_1 \oplus ct_2\right) \to \max$, where ct^0, ct^1 are zero and unity vectors.
6.2	Value of C^0_m is set as "0"	$cr_m[\xi] := 0, n_m := 0$, where n_m is the number of realizations of RO, which belong to C^0_m
6.3	RO implementations, belonging to clusters $CT^0_m[\xi]$, are defined	Rules: $ct_i \in CT^{o}_{1}[\xi]$, if $cr\left(ct_i \oplus ct_1\right) <=$ $<= cr\ \&\ cr\left(ct_i \oplus ct_1\right) < \left(ct_i \oplus ct_2\right)$; $ct_i \in CT^{o}_{2}[\xi]$, if $cr\left(ct_i \oplus ct_2\right) <=$ $<= cr\ \&\ cr\left(ct_i \oplus ct_2\right) < \left(ct_i \oplus ct_1\right)$; where $ct_i \mid i = \overline{1,N}$ are the implementations of BLM $\lVert ct_i^{(j)} \rVert$

Continuation of Table 1

1	2	3
6.4	Calculation of current ICFE [2, 5, 8, 24, 25]	$\overline{\text{CE}}^{*}=(1/\text{M})\cdot\sum_{\text{m}=1}^{\text{M}}\max_{\{\text{ls}\}}\text{CE}_{\text{c}}$, where CE_{c} is the value of ICFE of ASR learning for the realization of class of anomalies or cyber attacks – CT_{c}^{0}; {ls} is the set of steps for ASR learning as a part of ISDA
6.5	Formation of set {ct_{m}} of standard realizations for clusters $\{\text{CT}_{\text{m}}^{0}[\xi]\}$	Rule for defining coordinates: $\text{ct}_{\text{m,i}}=\begin{cases}1, \text{ if} \frac{1}{\text{n}}\sum_{\text{j}=1}^{\text{n}}\text{cr}_{\text{m,i}}^{(\text{j})}>1/2;\\0, \text{ else}\end{cases}$
6.6	Conditions verification	$\begin{cases}\text{if } \text{N}'=\sum_{\text{m}=1}^{\text{M}}\text{n}_{\text{m}}<\text{N} \text{ then}\rightarrow 6.7\ \&\ 6.3\\ \text{else } 6.9\end{cases}$
6.7	Conditions verification	$\begin{cases}\text{if } \text{cr}_{\text{m}}[\xi]<\text{cr}(\text{ct}_1\oplus\text{ct}_2) \text{ then}\rightarrow 6.8\ \&\ 6.3\\ \text{else } 6.9\end{cases}$
6.8	Increasing RC	$\text{cr}_{\text{m}}[\xi]:=\text{cr}_{\text{m}}[\xi]+1$
6.9	Calculation of ICFE and optimal radii of clusters $\{\text{CT}_{\text{m}}^{0}[\xi]\}$	Under conditions: $\text{N}'=\sum_{\text{m}=1}^{\text{M}}\text{n}_{\text{m}}<\text{N}$, where N' is the number of RO implementations that belong to RC_{ξ} and $\text{cr}_{\text{m}}[\xi]<\text{cr}(\text{ct}_1\oplus\text{ct}_2)$
7	Increasing SC	$\xi:=\xi+1$
8	Splitting a binary space of features (BSF) into 3 clusters	$\{\text{CT}_{\text{m}}^{\text{o}}[\xi]\,\vert\,\text{m}=\overline{1,3}\}$
8.1	Calculation of BLM for cluster CT_3^{0}, the standard vector-realization ct_3 of which satisfies the conditions	Verification of conditions: $\text{cr}(\text{ct}_1\oplus\text{ct}_3)\rightarrow\min\ \&\ \text{cr}(\text{ct}_2\oplus\text{ct}_3)\rightarrow\min$, where ct_1, ct_2 are the standard realizations of clusters $\{\text{CT}_{\text{m}}^{\text{o}}\,\vert\,\text{m}=\overline{1,2}\}$, restored at performing stage 6
8.2	Value of radius of cluster CT_3^{o} is set as "0"	$\text{cr}_3[\xi]:=0.$
8.3	Determining the cases of obtaining RO features implementations in cluster CT_3^{0}	Rules for determining the cases of obtaining RO features implementations in cluster CT_3^{o}: $\text{ct}_{\text{i}}\in\text{CT}_3^{\text{o}}$ if $\text{cr}(\text{ct}_{\text{i}}\oplus\text{ct}_3)<=$ $<=\text{cr}\ \&\ \text{cr}(\text{ct}_{\text{i}}\oplus\text{ct}_3)<=$ $<=\text{cr}(\text{ct}_{\text{i}}\oplus\text{ct}_1)\ \&\ \text{cr}(\text{ct}_{\text{i}}\oplus\text{ct}_3)<=$ $<=\text{cr}(\text{ct}_1\oplus\text{ct}_2)$, where $\text{ct}_{\text{i}}\,\vert\,\text{i}=\overline{1,\text{N}}$ are the implementations of BLM $\Vert\text{ct}_{\text{i}}^{(\text{j})}\Vert$
8.4	Correction of containers for clusters $\{\text{CT}_{\text{m}}^{\text{o}}\,\vert\,\text{m}=\overline{1,2}\}$ is performed	Implementations $\{\text{ss}^{(\text{j})},\text{j}=\overline{1,\text{n}}\}$, which arrived to container of category CT_3^{o}, are removed from container $\{\text{CT}_{\text{m}}^{\text{o}}\}$. Radius of container $\{\text{CT}_{\text{m}}^{\text{o}}\}$: $\text{cr}_{\text{m}}[\xi]:=\text{cr}_{\text{m}}[\xi]-1$
8.5	Calculation of current ICFE	Expression – stage 6.4
8.6	Formation of set $\{\text{ct}_{\text{m}}\}$ of standard implementations $\{\text{CT}_{\text{m}}^{\text{o}}[\xi]\}$	Rule for defining coordinates: $\text{ct}_{\text{m,i}}=\begin{cases}1, \text{ if} \frac{1}{\text{n}}\sum_{\text{j}=1}^{\text{n}}\text{cr}_{\text{m,i}}^{(\text{j})}>1/2;\\0, \text{ else}\end{cases}$
8.7	Condition verification	$\begin{cases}\text{if } \text{cr}_3[\xi]<\text{cr}(\text{ct}_1\oplus\text{ct}_3)\ \&\ \text{cr}_3[\xi]<\\<\text{cr}(\text{ct}_2\oplus\text{ct}_3) \text{ then}\rightarrow 8.8;\\ \text{else } 8.9\end{cases}$

Continuation of Table 1

1	2	3
8.8	Increasing radius	$cr_3[\xi] := cr_3[\xi]+1$
8.9	Optimal radius of cluster container CT_3^0 is calculated	At conditions: $cr_3[\xi] < cr(ct_1 \oplus ct_3)\,\&$ $cr_3[\xi] < cr(ct_2 \oplus ct_3)$
9	Condition verification	$\begin{cases} \text{if } \; ca[1] \le 0{,}5 \cdot ca_{low} \text{ then} \to 2 \\ \text{else } 10 \end{cases}$ ca_{low_i} is the VAD for RO attributes, which are determined based on [5, 8, 25]
10	Condition verification	$\begin{cases} \text{if } \; \overline{CE}[1] \notin IS_{CE} \text{ then} \to 11 \\ \text{else } 2 \end{cases}$
11	Search for global maximal (GMAX) value $\overline{CE}$ in the operating range of RO attributes	$ca^* = \arg\max\limits_{IS_{ca}}\{\max\limits_{IS_{CE} \cap IS_{cr}} \overline{CE}\}\,\&$ $\overline{CE}^*[1] := \text{extrem}CE_m[l]$
12	Based on methods [5, 8, 25] and others, optimal parameter of fields ca of RO attributes for the container is defined	$A_{low_i}^{op} = lm_i - ca^{op}\dfrac{ca_{low_i}}{100};$ $A_{up_i}^{op} = lm_i + ca^{op}\dfrac{ca_{up_i}}{100}$
13	Procedure of splitting BSF of RO into 4 clusters:	$\{CT_m^o[\xi] \mid m = \overline{1,4}\}$
13.1	Binary matrix of cluster is defined $\{CT_4^o\}$	Under conditions: $cr(ct_1 \oplus ct_4) \to \min$, $cr(ct_2 \oplus ct_4) \to \min\ \&$ $cr(ct_3 \oplus ct_4) \to \min$, where ct_1, ct_2, ct_3 are the standard implementations of clusters $\{CT_m^o \mid m = \overline{1,3}\}$, restored when performing stage 8
13.2	Value of radii of cluster CT_4^o is set as "0"	$cr_4[\xi] := 0$
13.3	Determining RO realizations, which arrived to cluster CT_4^o	Rule: $ct_i \in CT_4^o$, if $cr(ct_i \oplus ct_4) <= cr_4[\xi]$, where $ct_i \mid i = \overline{1,N_4}$ are the implementations of BLM $\lVert ct_i^{(j)} \rVert$
13.4	Calculation of current ICFE	Expression – stage 6.4.
13.5	Formation $\{ct_m\}$ of standard implementations for clusters $\{CT_m^o[\xi]\}$	Rule for defining coordinates: $ct_{m,i} = \begin{cases} 1, \text{ if } \dfrac{1}{n}\sum\limits_{j=1}^{n} cr_{m,i}^{(j)} > 1/2; \\ 0, \text{ else} \end{cases}$
13.6	Conditions verification	$\begin{cases} \text{if } cr_4[\xi] < cr(ct_1 \oplus ct_4), \\ cr_4[\xi] < cr(ct_2 \oplus ct_4), \\ cr_4[\xi] < cr(ct_3 \oplus ct_4) \quad \text{then} \to 8.3\,\&\,8.8; \\ \text{else } 8.9 \end{cases}$
13.7	The next RO attribute in cluster CT_4^o is added	$ct_4 := ct_4 + 1$.
13.8	Optimal radius of container CT_4^o is determined	At conditions: $cr_4[\xi] < cr(ct_1 \oplus ct_4)$, $cr_4[\xi] < cr(ct_2 \oplus ct_4)$, $cr_4[\xi] < cr(ct_3 \oplus ct_4)$
14	Adding results to a knowledge base (KB). End of algorithm operation.	

We developed the algorithm that allows us to perform parallel formation of reference tolerances during an analysis of attributes of anomalies and cyber attacks, which are difficult to explain [1, 7, 16, 18]. This approach, when a parallel formation of VAD – ($\{ca_{K,i}\}$) is performed, makes it possible to change VAD for all attributes at every step of learning simultaneously. The algorithm enables in the course of learning to update optimal parameters of containers for the recognition classes CT_m^0. The stages of splitting FS of RO into clusters are presented in tabular form in Table 2.

Table 2

Stages of algorithm of VAD formation for the attributes of recognition of cyber attacks, anomalies or threats

Stageee	Action	Clustering algorithm for a mathematical description of RO attributes
1	Value of meter of steps of VAD change ca_i for RO attribute «0»	$l := 0$
2	Calculation of $A_{low_i}[l]$ and $A_{up_i}[l]$ of VAD of RO attribute for entire FS	$A_{low_i}[l] = lm_{1,i} - ca\frac{ca_{low_I}}{100};$ $A_{up_i}[l] = lm_{1,i} + ca\frac{ca_{low_I}}{100},$ where $lm_{l,i}$ is the i-th attribute of vector-standard of implementation lm_l for basic class CT_1^0. (It was accepted that CT_1^0 characterizes the most acceptable states of IB).
3	Formation of BLM $\left\Vert ct_i^{(j)} \right\Vert$	Rule: $ct_{m,i}^{(j)} = \begin{cases} 1, \text{ if } A_{low_I}[l] < lm_i^{(j)} < A_{up_i}[l]; \\ 0, \text{ else} \end{cases}$
4	Formation of set $\{ct_m\}$ for vectors-standards of implementation of RO CT_m^0	$ct_{m,i} = \begin{cases} 1, \text{ if } {}^1\!/_n \sum_{j=1}^{n} ct_{m,i}^{(j)} > {}^1\!/_2; \\ 0, \text{ else,} \end{cases}$ where n is the number of implementation of RO (attributes), which belong to the cluster of correspondent class $CT^0{}_m$
5	Splitting $\{ct_m\}$ into pairs of the nearest adjacent vectors-standards	Methods and models [8, 10, 12, 14, 23, 25] are used
6	Restoration of container for CT_m^0	
6.1	Values of meter of recognition classes "0»	m:=0
6.2	Increasing the value of meter	m:=m+1
6.3	Value of meter of steps of RC change "0»	cr:=0
6.4	Increasing the value of meter	cr:= cr +1
6.5	Calculation of current ICFE	Expression – stage 6.4 Table 1
6.6	Condition verification	$\begin{cases} \text{if } CE_m \notin IS_{CE} \quad \text{then} \to 6.4 \\ \text{else } 6.7. \end{cases}$
6.7	Calculation of current ICFE	Expression – stage 6.4 Table 1
6.8	Calculation of GMAX of ICFE	$CE_m^*[l] := \underset{\{cr\}}{\text{extrem}}\, CE_m[l, cr]$
6.9	Calculation of optimal RC of RO class CT_m^0	$cr_m^*[l] := \arg \underset{\{cr\}}{\text{extrem}}\, CE_m[l, cr]$
7	Condition verification	$\begin{cases} \text{if } m \notin M \quad \text{then} \to 6.2 \\ \text{else } 8 \end{cases}$
8	Calculation of averaged ICFE value	$\overline{CE}_{cp} = (1/M) \cdot \sum_{m=1}^{M} \max_{\{ls\}} CE_c$
9	Condition verification	$\begin{cases} \text{if } ca[l] \leq ca_{low} / 2 \text{ then} \to 2 \\ \text{else } 10 \end{cases}$
10	Condition verification	$\begin{cases} \text{if } \overline{CE} \notin IS_{CE} \text{ then} \to 11 \\ \text{else } 6.8 \,\&\, 6.9 \end{cases}$
11	Calculation of GMAX ICFE in admissible function determination range	$ca^* = \arg \max_{IS_{ca}} \{ \max_{IS_{CE} \cap IS_{cr}} \overline{CE} \}$
12	Adding results to a knowledge base (KB). End of algorithm operation.	

Input data for ASR are an array of learning samples, obtained based on data from Tables 1, 2, as well as results of [10, 16]:

$$LM[kl][implementation][j], \tag{13}$$

where kl is the number of learning matrix for RO class; implementation is the number of implementation in BLM [10, 16]; j is the number of recognition attribute for RO.

To assess ASR effectiveness and optimality of defined VAD for RO classes of ISDA, the Pareto method was used

[5, 8, 28]. The membership degree of the best, from the standpoint of ARS or an expert, variant of Pareto-optimal solution in terms of strategies for providing cyber protection was determined by formula:

$$\max\left[\sum_{j=1}^{h}\sum_{l=1}^{d}\tilde{z}_{ij}\otimes\tilde{p}_j\otimes\tilde{p}_l^{SS}\right]=\max_{W_i\in W}CE\big(W_i(x)\big), \qquad (14)$$

where ⊗ is the triangular norm (T-norm) [5, 28]; $W_i(x)$ is the final choice of the solution option of ASR (or an expert); $\tilde{z}_{ij}$ is the fuzzy assessment of usefulness of the i-th option of solving the problem of recognition of anomaly or cyber attack, which is determined by value of ICFE; $\tilde{p}_j$ is the assessment of CIIS states in the process of RO recognition; $\tilde{p}_l^{SS}$ are the assessments of ASR states in the process of anomalies or cyber attack recognition.

Membership degree of the best variant of Pareto-optimal fuzzy solution for the formation of KB for ASR was defined using the modified Wald criterion and the Savage criterion [5].

5. Simulation of the clustering algorithm and the formation of VAD for the attributes of anomalies and cyber attacks

The algorithms were implemented in the MATLAB 7/2009 and Simulink programming environments in order to subsequently study the operation modes of ASR of anomalies and cyber attacks in CIIS (under conditions of countering the targeted cyber attacks [1, 7, 10, 16, 18]).

In accordance with recommendations of [8, 20, 21, 25], multidimensional binary learning matrices (MBLM) of RO classes had from 50 to 65 implementations. For the classes of network attacks [7, 8] (DoS/DDoS, Probe, R2L, U2R), the number of recognition attributes made up 12–41 [13, 23, 15], for virus attacks, 7–15 [5, 7] attributes. Fig. 1, *a–e* shows dependences of ICFE learning of simulation model (SM) of ASR [23] on RC of RO – cr. In Fig. 1, *a–e*, the middle section (marked in blue) corresponds to the operation area of the selected recognition attributes that have the highest informativeness indicator (ICFE) [23].

After formation of MBLM for the normal behavior of a system, according to the proposed algorithm, binary trees of traffic are constructed for network attacks, as well as error-free decisive rules, by the appropriate learning matrix of attributes [16, 18, 23]. Next, MBLM are determined and registered for the system, which allows us to form controlling commands for responding to the deviations of parameters from the estimated values, please refer to Fig. 2, *a, b*.

Fig. 3 shows results, obtained in the course of simulation modeling and testing of algorithms of parallel clustering and formation of reference deviations for the recognition attributes, on the example of a DoS class of attacks. Results of the clustering of attack attributes in the process of testing the improved algorithm and the formation of VAD are shown in blue color. Similar results were also obtained for other classes of anomalies and cyber attacks.

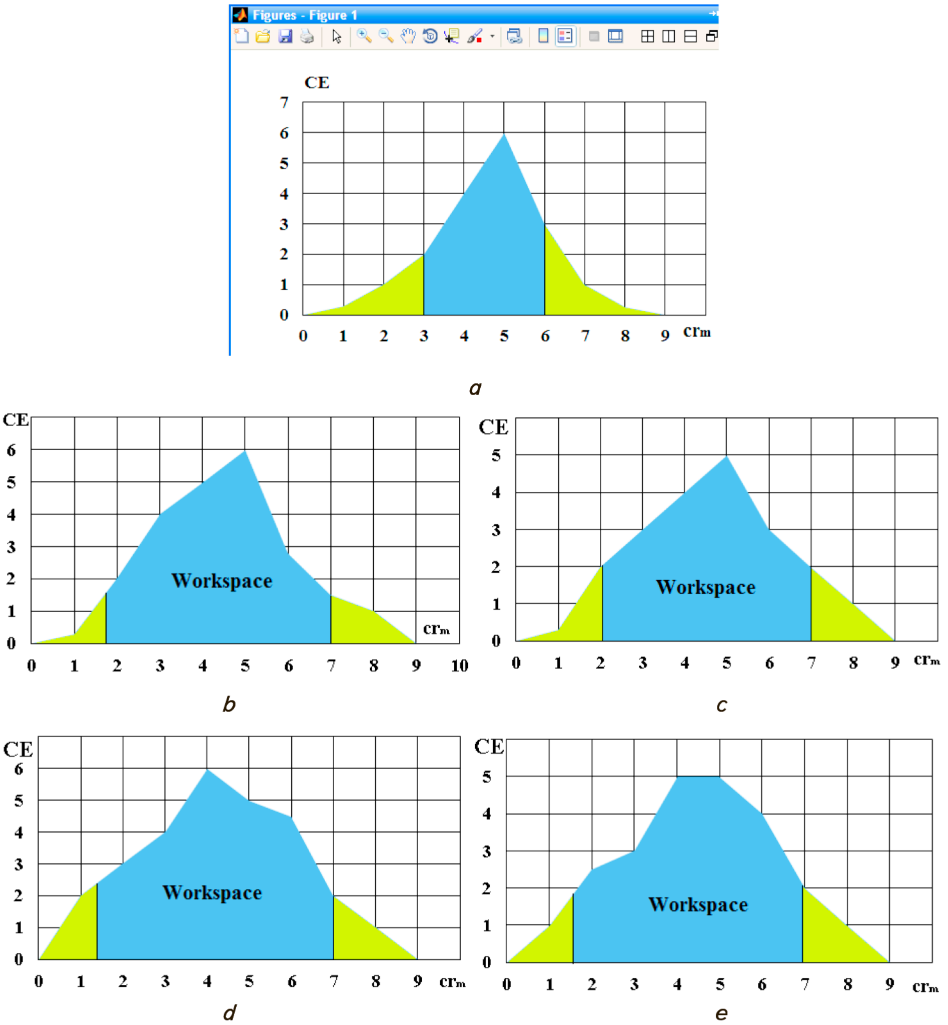


Fig. 1. Dependences of ICFE of learning of simulation model ASR on RC of RO: *a* – ICFE for the DoS/DDoS attacks; *b* – ICFE for the Probe attacks; *c* – ICFE for the R2L attacks; *d* – ICFE for the U2R attacks; *e* – ICFE for virus attacks

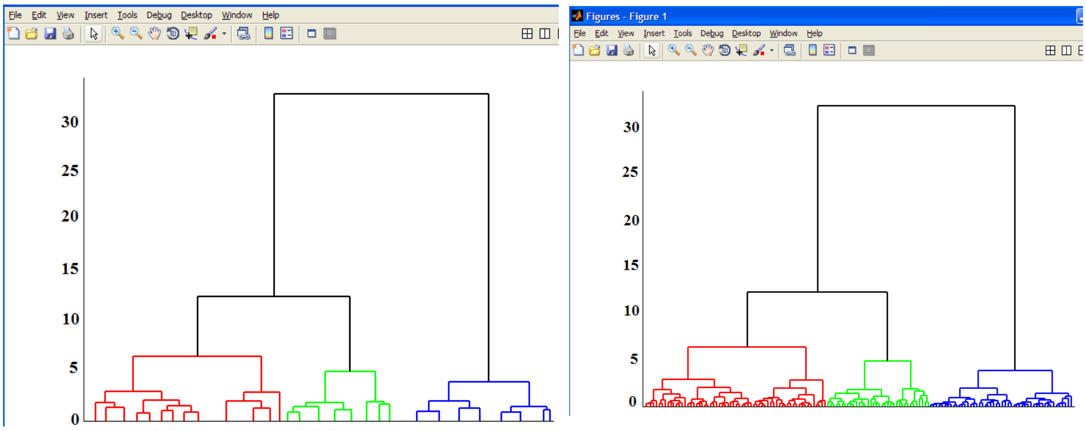
An analysis of results of the simulation experiment (Fig. 3) on determining the dependence of ICFE of ASR learning allows us to draw the following conclusions:

– the averaged maximum value of ICFE of ASR learning is equal to: for attacks of the DoS/DDoS class $\overline{CE}$=3.19; for attacks of the Probe class $\overline{CE}$=3.15; for attacks of the R2L class $\overline{CE}$=2.84; for attacks of the U2R class $\overline{CE}$=3.27; for virus attacks (VA)=2.56;

– the averaged value of optimal radius cr equals in code units for RO classes, given in Table 3, respectively: for $hy_{\gamma 1}$ class: DoS/DDoS – $cr_1^* = 4$; Probe – $cr_1^* = 3$; R2L – $cr_1^* = 4$; U2R – $cr_1^* = 4$; BA – $cr_1^* = 5$; for $hy_{\gamma 2}$ class: DoS/DDoS –

$cr_2^* = 2$; Probe – $cr_2^* = 1$; R2L – $cr_2^* = 1$; U2R – $cr_2^* = 1$; BA – $cr_2^* = 2$; for $hy_{\gamma 3}$ class: DoS/DDoS – $cr_3^* = 3$; Probe – $cr_3^* = 3$; R2L – $cr_3^* = 2$; U2R – $cr_3^* = 2$; BA – $cr_3^* = 3$.

The values of optimal RC cr, taking into consideration additional hypotheses for the examined simulation models of ASR learning, are given in Table 3.



a *b*

Fig. 2. Structural characteristics of anomalous and normal traffic: *a* — normal traffic for simulation model; *b* — traffic for the case of recognition of a network attack

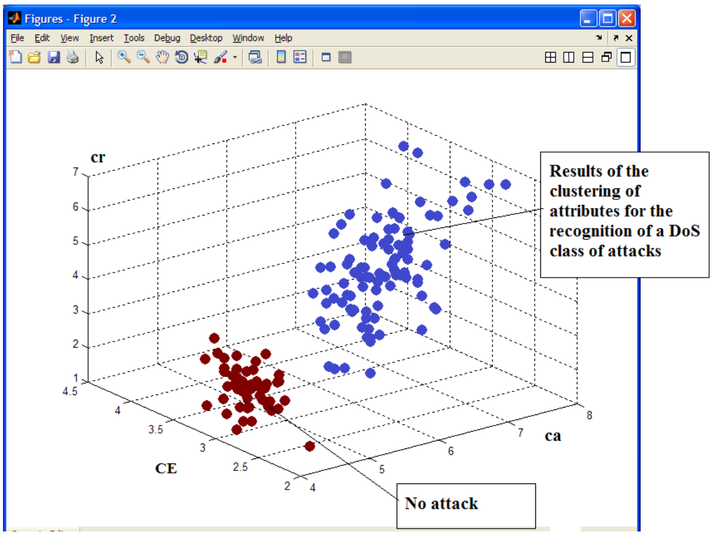


Fig. 3. Results of the stages of parallel clustering and formation of VAD for the recognition of attributes (on the example of DoS attacks)

Table 3

Values of optimal RC cr for the examined simulation models of ASR learning

No.	Accepted hypotheses for RO	Values of optimal RC cr				
		DoS/ DDoS	Probe	R2L	U2R	BA
Basic hypotheses						
1	Basic working hypothesis – $hy_{\gamma 1}$: attribute (attributes) rc_i of RO and indicator IE (characterizes stability of CIIS functioning [18, 23]) is within the normal state of CIIS	$cr_1^{opt} = 4-5$	$cr_1^{opt} = 3-4$	$cr_1^{opt} = 4-5$	$cr_1^{opt} = 4-5$	$cr_1^{opt} = 5-6$
2	Hypothesis $hy_{\gamma 2}$ – attribute (attributes) allows drawing a conclusion that indicator IE is lower than the norm	$cr_2^{opt} = 2-3$	$cr_2^{opt} = 1-2$	$cr_2^{opt} = 1-2$	$cr_2^{opt} = 1-2$	$cr_2^{opt} = 2-3$
3	Hypothesis $hy_{\gamma 3}$ allows drawing a conclusion that indicator IE is higher than the norm	$cr_3^{opt} = 3-4$	$cr_3^{opt} = 3-4$	$cr_3^{opt} = 2-3$	$cr_3^{opt} = 2-3$	$cr_3^{opt} = 3-4$
Additional hypotheses for simulation model						
4	Hypothesis $hy_{\gamma 1}^{D}$ – node of CIIS demonstrates increased network activity	$cr_{D1}^{opt} = 4$	$cr_{D1}^{opt} = 4$	$cr_{D1}^{opt} = 3$	$cr_{D1}^{opt} = 3$	–
5	Hypothesis $hy_{\gamma 2}^{D}$ – node of CIIS demonstrates increased activity during external traffic	$cr_{D2}^{opt} = 3$	$cr_{D2}^{opt} = 3$	$cr_{D2}^{opt} = 3$	$cr_{D2}^{opt} = 2$	–

As was shown by data analysis, for IM, Fig. 1–3, quasi-optimal value of parameter $ca_{n,i}$ of VAD equals VAD=8–16 % at maximum value of CE_{max}=6.16.

Thus, it was proved in the course of the simulation experiment that the proposed algorithms for the clustering of RO attributes enable us to obtain efficient learning matrices for ASR as a part of ISDA.

6. Discussion of results of testing the algorithms and prospects of further research

Scientific and practical results of research in the form of software applications were implemented in ASR and adaptive expert systems (AES) of cyber protection, implemented at the state enterprise "Design and engineering technological bureau of automation of control systems on railway transport of Ukraine" of the Ministry of Infrastructure of Ukraine, as well as in the information security services of computing centers at the industrial and transportation enterprises in the cities of Kyiv, Dnipro and Chernihiv.

The proposed algorithms differ from the existing ones by the possibility of simultaneous formation of reference tolerances in the course of analysis of complex attributes of anomalies and cyber attacks. This allows changing VAD for all attributes simultaneously during the procedure of training the existing and promising ISDA. The improved algorithms are also focused on the possibility of processing a large amount of specialized data during procedures of the recognition and analysis of various types of attributes of anomalies and targeted cyber attacks in CIIS.

The effectiveness of using the proposed algorithms depends on the number of informative attributes, which are used for the formation of BLM. In addition, efficiency of algorithms is determined by the input data for ASR or AES, formed at each step of clustering. When the number of attributes is insignificant, the effect of using the modified algorithm is negligible.

The results presented are a continuation of the research, results of which were described earlier in articles [10, 18, 23]. The prospects of further research include the enlargement of attributes knowledge base and the formation of BLM of ASR.

7. Conclusions

1. We proposed to refine the algorithm of splitting the feature space into clusters in the course of implementation of procedure for the recognition of anomalies and cyber attacks, which differs from the existing algorithms by the simultaneous formation of reference tolerances during analysis of complex RO attributes, and allows simultaneous changing of VAD for all attributes at every step of learning. The proposed refinements make it possible to prevent possible cases of the absorption of one RO class of basic attributes of anomalies and cyber attacks by another class. In this case, predicate expressions were obtained for ASR that is capable of self-learning.

2. We examined the devised algorithms on the simulation models in MatLab. It was proved that the proposed algorithms for the clustering of RO attributes enable to obtain effective learning matrices for ASR as a part of ISDA.

References

1. Khan, L. A new intrusion detection system using support vector machines and hierarchical clustering [Text] / L. Khan, M. Awad, B. Thuraisingham // The VLDB Journal. – 2006. – Vol. 16, Issue 4. – P. 507–521. doi: 10.1007/s00778-006-0002-5
2. Ranjan, R. A New Clutering Approach for Anomaly Intrusion Detection [Text] / R. Ranjan, G. Sahoo // International Journal of Data Mining & Knowledge Management Process. – 2014. – Vol. 4, Issue 2. – P. 29–38. doi: 10.5121/ijdkp.2014.4203
3. Feily, M. A Survey of Botnet and Botnet Detection [Text] / M. Feily, A. Shahrestani, S. Ramadass // 2009 Third International Conference on Emerging Security Information, Systems and Technologies. – 2009. doi: 10.1109/securware.2009.48
4. Mahmood, T. Security Analytics: Big Data Analytics for cybersecurity: A review of trends, techniques and tools [Text] / T. Mahmood, U. Afzal // 2013 2nd National Conference on Information Assurance (NCIA). – 2013. doi: 10.1109/ncia.2013.6725337
5. Dua, S. Data Mining and Machine Learning in Cybersecurity [Text] / S. Dua, X. Du. – UK, CRC press, 2016. – 256 p.
6. Zhang, S. An Empirical Study on Using the National Vulnerability Database to Predict Software Vulnerabilities [Text] / S. Zhang, D. Caragea, X. Ou // Lecture Notes in Computer Science. – 2011. – P. 217–231. doi: 10.1007/978-3-642-23088-2_15
7. Lee, K.-C. Sec-Buzzer: cyber security emerging topic mining with open threat intelligence retrieval and timeline event annotation [Text] / K.-C. Lee, C.-H. Hsieh, L.-J. Wei, C.-H. Mao, J.-H. Dai, Y.-T. Kuang // Soft Computing. – 2016. – Vol. 21, Issue 11. – P. 2883–2896. doi: 10.1007/s00500-016-2265-0
8. Buczak, A. L. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection [Text] / A. L. Buczak, E. Guven // IEEE Communications Surveys & Tutorials. – 2016. – Vol. 18, Issue 2. – P. 1153–1176. doi: 10.1109/comst.2015.2494502
9. Petit, J. Potential Cyberattacks on Automated Vehicles [Text] / J. Petit, S. E. Shladover // IEEE Transactions on Intelligent Transportation Systems. – 2015. – Vol. 16, Issue 2. – P. 546–556. doi: 10.1109/tits.2014.2342271
10. Lakhno, V. A. Applying the functional effectiveness information index in cybersecurity adaptive expert system of information and communication transport systems [Text] / V. A. Lakhno, P. U. Kravchuk, V. L. Pleskach, O. P. Stepanenko, R. V. Tishchenko, V. A. Chernyshov // Journal of Theoretical and Applied Information Technology. – 2017. – Vol. 95, Issue 8. – P. 1705–1714.
11. Dovbysh, A. S. Information-extreme Algorithm for Recognizing Current Distribution Maps in Magnetocardiography [Text] / A. S. Dovbysh, S. S. Martynenko, A. S. Kovalenko, N. N. Budnyk // Journal of Automation and Information Sciences. – 2011. – Vol. 43, Issue 2. – P. 63–70. doi: 10.1615/jautomatinfscien.v43.i2.60
12. Ameer Ali, M. Review on Fuzzy Clustering Algorithms [Text] / M. Ameer Ali, G. C. Karmakar, L. S. Dooley // IETECH Journal of Advanced Computations. – 2008. – Vol. 2, Issue 3. – P. 169–181.

13. Guan, Y. Y-means: a clustering method for intrusion detection [Text] / Y. Guan, A. A. Ghorbani, N. Belacel // CCECE 2003 – Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No.03CH37436). – 2003. doi: 10.1109/ccece.2003.1226084
14. Halkidi, M. On Clustering Validation Techniques [Text] / M. Halkidi, Y. Batistakis, M. Vazirgiannis // Journal of Intelligent Information Systems. – 2001. – Vol. 17, Issue 2/3. – P. 107–145. doi: 10.1023/a:1012801612483
15. Gamal, M. M. A Security Analysis Framework Powered by an Expert System [Text] / M. M. Gamal, B. Hasan, A. F. Hegazy // International Journal of Computer Science and Security (IJCSS). – 2011. – Vol. 4, Issue 6. – P. 505–527.
16. Lakhno, V. A model developed for teaching an adaptive system of recognizing cyberattacks among non-uniform queries in information systems [Text] / V. Lakhno, H. Mohylnyi, V. Donchenko, O. Smahina, M. Pyroh // Eastern-European Journal of Enterprise Technologies. – 2016. – Vol. 4, Issue 9 (82). – P. 27–36. doi: 10.15587/1729-4061.2016.73315
17. Riadi, I. Log Analysis Techniques using Clustering in Network Forensics [Text] / I. Riadi, J. E. Istiyanto, A. Ashari, N. Subanar // (IJCSIS) I International Journal of Computer Science and Information Security. – 2012. – Vol. 10, Issue 7.
18. Lakhno, V. Development of adaptive expert system of information security using a procedure of clustering the attributes of anomalies and cyber attacks [Text] / V. Lakhno, Y. Tkach, T. Petrenko, S. Zaitsev, V. Bazylevych // Eastern-European Journal of Enterprise Technologies. – 2016. – Vol. 6, Issue 9 (84). – P. 32–44. doi: 10.15587/1729-4061.2016.85600
19. Kiss, I. A clustering-based approach to detect cyber attacks in process control systems [Text] / I. Kiss, B. Genge, P. Haller // 2015 IEEE 13th International Conference on Industrial Informatics (INDIN). – 2015. doi: 10.1109/indin.2015.7281725
20. Dovbysh, A. S. Informatsionno-ekstremalnyy algoritm optimizatsii parametrov giperellipsoidnykh konteynerov klassov raspoznavaniya [Text] / A. S. Dovbysh, N. N. Budnik, V. V. Moskalenko // Problemy upravleniya i informatiki. – 2012. – Issue 5. – P. 111–119.
21. Lee, S. M. Detection of DDoS attacks using optimized traffic matrix [Text] / S. M. Lee, D. S. Kim, J. H. Lee, J. S. Park // Computers & Mathematics with Applications. – 2012. – Vol. 63, Issue 2. – 501–510. doi: 10.1016/j.camwa.2011.08.020
22. Gao, P. Identification of Successive “Unobservable” Cyber Data Attacks in Power Systems Through Matrix Decomposition [Text] / P. Gao, M. Wang, J. H. Chow, S. G. Ghiocel, B. Fardanesh, G. Stefopoulos, M. P. Razanousky // IEEE Transactions on Signal Processing. – 2016. – Vol. 64, Issue 21. – P. 5557–5570. doi: 10.1109/tsp.2016.2597131
23. Lakhno, V. Design of adaptive system of detection of cyber-attacks, based on the model of logical procedures and the coverage matrices of features [Text] / V. Lakhno, S. Kazmirchuk, Y. Kovalenko, L. Myrutenko, T. Zhmurko // Eastern-European Journal of Enterprise Technologies. – 2016. – Vol. 3, Issue 9 (81). – P. 30–38. doi: 10.15587/1729-4061.2016.71769
24. Dovbysh, A. S. Optimization of the parameters of learning intellectual system of human signature verification [Text] / A. S. Dovbysh, D. V. Velikodnyi, J. V. Simonovski // Radioelectronic and computer systems. – 2015. – Issue 2. – P. 44–49.
25. Akhmetov, B. Designing a decision support system for the weakly formalized problems in the provision of cybersecurity [Text] / B. Akhmetov, V. Lakhno, Y. Boiko, A. Mishchenko // Eastern-European Journal of Enterprise Technologies. – 2017. – Vol. 1, Issue 2 (85). – P. 4–15. doi: 10.15587/1729-4061.2017.90506
26. Callegari, C. Improving PCA-based anomaly detection by using multiple time scale analysis and Kullback-Leibler divergence [Text] / C. Callegari, L. Gazzarrini, S. Giordano, M. Pagano, T. Pepe // International Journal of Communication Systems. – 2012. – Vol. 27, Issue 10. – P. 1731–1751. doi: 10.1002/dac.2432
27. Chinh, H. N. Fast Detection of Ddos Attacks Using Non-Adaptive Group Testing [Text] / H. N. Chinh, T. Hanh, N. D. Thuc // International Journal of Network Security & Its Applications. – 2013. – Vol. 5, Issue 5. – P. 63–71. doi: 10.5121/ijnsa.2013.5505