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 increase 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 CTm, 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 (rCm) of container (RC) at every step of ASR (or ISDA) learning, it is possible to use the following recurrent expression:

$$\text{cr}_m(\text{ls}) = \left[ \text{cr}_m(\text{ls}-1) + \xi \right] \text{cr}_m(\text{ls}) \in IS_{m}^c,$$

where ls is the number of steps of increasing RC Cm; \(\xi\) are accepted for the chosen attributes of steps of increasing RC; IS_{m}^c 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):

$$\text{is}_k = \text{Argmax}_{\text{ls}}\{\max_{\text{ls}} \sum_{m=1}^{M} \text{CE}_m \},$$

where \(\text{CE}_m\) is the ICFE of ASR learning to recognize RO that belong to class \(C^m\); IS\(\text{CE}\) 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):

$$\left\{ \left[ CT_m \neq \emptyset \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \right\}$$

$$\left\{ \left[ CT_m \neq \emptyset \right] \rightarrow \left[ CT_m \cap CT_m \neq \emptyset \right] \rightarrow \left[ \exists CT_m \in RC^p \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \right\}$$

$$\left\{ CT_m \neq \emptyset \rightarrow \left[ \forall CT_m \in RC^p \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \right\} \rightarrow \left[ \forall CT_m \in RC^p \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \rightarrow \left[ \forall CT_m \in RC^p \right],$$

where \(BCT^a\), \(BCT^b\) are the nuclei of RO classes \(CT^a\) and \(CT^b\), respectively:

$$\bigcup_{CT \in RC} CT_m \subseteq RS_{a,b;m}$$

Accepted assumptions: classes \(CT^a\) are \(CT^b\) adjacent; the classes have a minimum distance between the centers of clusters \(er(\text{ct}_a \oplus \text{ct}_b)\) among all classes for RO: RO are described by binary learning matrices (BLM) [21–23]. We accepted that \(\text{ct}_a\) and \(\text{ct}_b\) 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:

$$\left\{ \left[ CT_m \neq \emptyset \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \rightarrow \left[ \forall CT_m \in RC^p \right] \right\},$$

where \(\text{cr}_m \), \(\text{ct}_m \) are the optimal radii of containers \(C^a\) and \(C^b\), 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 dimensional attributes of ASR 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):
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:

\[
\{CT^i_{m,i} \neq \varnothing, m = 1,M \} \land \{\forall CT^i_{m,i} \in RCH\} ,
\]

where $BCT^i_{m,i}$, $BCT^i_{m,i}$ are the centers of the two nearest (adjacent) clusters $CT^i_{m,i}$ and $CT^i_{m,i}$, respectively; $\xi$ is the step of increasing the radius of cluster container (RCC).

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

<table>
<thead>
<tr>
<th>Stage</th>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Step counter (SC) of changing VAD $ca_i$ by features of RO is set as “0”: $1 := 0$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Calculation of the lower $A_{min}[1]$ and the upper $A_{up}[1]$ of VAD of RO features for entire FS</td>
<td>$A_{min}[1] = \text{Im}<em>i - \frac{ca</em>{min}}{100}$; $A_{up}[1] = \text{Im}<em>i + \frac{ca</em>{max}}{100}$ where $\text{Im}_i$ is the $i$-th attribute of standard vector-realization of non-classified multi-dimensional matrix (NMLM) $\text{Im}<em>i$ [16, 23]; $ca</em>{min}$ is the VAD for RO attributes, which are determined based on methods [2, 16, 21, 23]</td>
</tr>
<tr>
<td>3</td>
<td>Formation of BLM $[ct^i]$</td>
<td>Rule $ct^i_{\xi} = 1$, if $A_{min}[1] &lt; \text{Im}<em>i[0] &lt; A</em>{min}[1]$; $0$, else</td>
</tr>
<tr>
<td>4</td>
<td>Value of SC for increasing RC</td>
<td>$\xi := 0$</td>
</tr>
<tr>
<td>5</td>
<td>Initialization of SC for increasing RC</td>
<td>$\xi := 1$</td>
</tr>
<tr>
<td>6</td>
<td>Splitting NMLM ${ct^0}$ into two clusters</td>
<td>Verification of conditions: 1) $cr(ct_{\xi} \oplus ct^1) \rightarrow \text{min}$, $cr(ct_{\xi} \oplus ct^1) \rightarrow \text{min}$; 2) $cr(ct_{\xi} \oplus ct_{\xi}) \rightarrow \text{max}$, where $ct^0$, $ct^1$ are zero and unity vectors.</td>
</tr>
<tr>
<td>6.1</td>
<td>Initial original standard vectors for RO attributes $ct_{m,i}$ for $C^0_m$ are calculated</td>
<td>Rules: $ct_{\xi} \in CT^i_{\xi}[\xi]$, if $cr(ct_{\xi} \oplus ct_{\xi}) \Leftarrow$ $\Leftarrow cr &amp; cr(ct_{\xi} \oplus ct_{\xi}) &lt; (ct_{\xi} \oplus ct_{\xi})$; $ct_{\xi} \in CT^i_{\xi}[\xi]$, if $cr(ct_{\xi} \oplus ct_{\xi}) \Leftarrow$ $\Leftarrow cr &amp; cr(ct_{\xi} \oplus ct_{\xi}) &lt; (ct_{\xi} \oplus ct_{\xi})$; where $ct^i = \frac{1}{N}$ are the implementations of BLM $ct^i$</td>
</tr>
<tr>
<td>6.2</td>
<td>Value of $C^0_m$ is set as “0”</td>
<td>$cr_{\xi}[\xi] = 0$, $n_m := 0$, where $n_m$ is the number of realizations of RO, which belong to $C^0_m$</td>
</tr>
<tr>
<td>6.3</td>
<td>RO implementations, belonging to clusters $CT^i_{\xi}[\xi]$, are defined</td>
<td></td>
</tr>
</tbody>
</table>

\[
\bigcup_{CT^i_{m,i} \in RS,} CT^i_{m,i} \subseteq RS, \tag{12}
\]
Continuation of Table 1

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4</td>
<td>Calculation of current ICFE [2, 5, 8, 24, 25]</td>
<td>$CE^* = (1/M) \sum_{i=1}^{M} CE_i$, where $CE_i$ is the value of ICFE of ASR learning for the realization of class of anomalies or cyber attacks – $CT_{i}^*$ is the set of steps for ASR learning as a part of ISDA</td>
</tr>
<tr>
<td>6.5</td>
<td>Formation of set ${ct_m}$ of standard realizations for clusters ${CT_{m}^*[\xi]}$</td>
<td>Rule for defining coordinates: $ct_{m} = \begin{cases} 1, &amp; \text{if } \frac{1}{n} \sum_{i=1}^{n} ct_{i}^{(i)} &gt; \frac{1}{2}; \ 0, &amp; \text{else} \end{cases}$</td>
</tr>
<tr>
<td>6.6</td>
<td>Conditions verification</td>
<td>$\begin{cases} \text{if } \sum_{n=1}^{N} n_n &lt; N \text{ then } \rightarrow 6.7 &amp; 6.3 \ \text{else } 6.9 \end{cases}$</td>
</tr>
<tr>
<td>6.7</td>
<td>Conditions verification</td>
<td>$\begin{cases} \text{if } cr_{n}[\xi] &lt; cr_{(ct_{1} \oplus ct_{2})} \text{ then } \rightarrow 6.8 &amp; 6.3 \ \text{else } 6.9 \end{cases}$</td>
</tr>
<tr>
<td>6.8</td>
<td>Increasing RC</td>
<td>$cr_{n}[\xi] = cr_{n}[\xi] + 1$</td>
</tr>
<tr>
<td>6.9</td>
<td>Calculation of ICFE and optimal radii of clusters ${CT_{m}^*[\xi]}$</td>
<td>Under conditions: $N' = \sum_{n=1}^{M} n_n &lt; N$, where $N'$ is the number of RO implementations that belong to $RC_{3}$ and $cr_{n}[\xi] &lt; cr_{(ct_{1} \oplus ct_{2})}$</td>
</tr>
<tr>
<td>7</td>
<td>Increasing SC</td>
<td>$\xi := \xi + 1$</td>
</tr>
<tr>
<td>8</td>
<td>Splitting a binary space of features (BSF) into 3 clusters</td>
<td>${CT_{m}^*[\xi]</td>
</tr>
<tr>
<td>8.1</td>
<td>Calculation of BLM for cluster $CT_{3}^*$, the standard vector-realization $ct_{3}$ of which satisfies the conditions</td>
<td>Verification of conditions: $cr_{(ct_{1} \oplus ct_{2})} \rightarrow \min$ &amp; $cr_{(ct_{3} \oplus ct_{2})} \rightarrow \min$, where $ct_{1}$, $ct_{2}$ are the standard realizations of clusters ${CT_{m}^*</td>
</tr>
<tr>
<td>8.2</td>
<td>Value of radius of cluster $CT_{3}^*$ is set as “0”</td>
<td>$cr_{[\xi]} = 0.$</td>
</tr>
<tr>
<td>8.3</td>
<td>Determining the cases of obtaining RO features implementations in cluster $CT_{3}^*$</td>
<td>Rules for determining the cases of obtaining RO features implementations in cluster $CT_{3}^<em>$: $ct_{i} \in CT_{3}^</em>$ if $cr_{(ct_{i} \oplus ct_{2})} &lt;=$ $&lt; cr_{(ct_{1} \oplus ct_{3})} &lt;=$ $&lt; cr_{(ct_{3} \oplus ct_{2})}$ &amp; $cr_{(ct_{1} \oplus ct_{3})} &lt;=$ $&lt; cr_{(ct_{3} \oplus ct_{1})}$. where $ct_{i} = \overline{1,N}$ are the implementations of BLM ${ct_{i}^{(i)}}$</td>
</tr>
<tr>
<td>8.4</td>
<td>Correction of containers for clusters ${CT_{m}^*</td>
<td>m = \overline{1,2}}$ is performed</td>
</tr>
<tr>
<td>8.5</td>
<td>Calculation of current ICFE</td>
<td>Expression – stage 6.4</td>
</tr>
<tr>
<td>8.6</td>
<td>Formation of set ${ct_m}$ of standard implementations ${CT_{m}^*[\xi]}$</td>
<td>Rule for defining coordinates: $ct_{m} = \begin{cases} 1, &amp; \text{if } \frac{1}{n} \sum_{i=1}^{n} ct_{i}^{(i)} &gt; \frac{1}{2}; \ 0, &amp; \text{else} \end{cases}$</td>
</tr>
<tr>
<td>8.7</td>
<td>Condition verification</td>
<td>$\begin{cases} \text{if } cr_{[\xi]} &lt; cr_{(ct_{1} \oplus ct_{2})} &amp; cr_{[\xi]} &lt; cr_{(ct_{2} \oplus ct_{3})} \text{ then } \rightarrow 8.8; \ \text{else } 8.9 \end{cases}$</td>
</tr>
</tbody>
</table>
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 – \( \{vad_{k,i}\} \) 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 \( \{c_{k,i}\} \). The stages of splitting FS of RO into clusters are presented in tabular form in Table 2.

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 – \( \{vad_{k,i}\} \) 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 \( \{c_{k,i}\} \). The stages of splitting FS of RO into clusters are presented in tabular form in Table 2.

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
<th>Formula/Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.8</td>
<td>Increasing radius</td>
<td>( cr_i[\xi] = cr_i[\xi] + 1 )</td>
</tr>
<tr>
<td>8.9</td>
<td>Optimal radius of cluster container ( CT_i ) is calculated</td>
<td>At conditions: ( cr_i[\xi] &lt; cr(ct_i \oplus ct_j) ) &amp; ( cr_i[\xi] &lt; cr(ct_i \oplus ct_j) )</td>
</tr>
</tbody>
</table>
| 9   | Condition verification                                                       | \[
\begin{align*}
\text{if} & \quad ca[i] \leq 0.5 \cdot ca_{\text{min}} \rightarrow 2 \\
\text{else} & \quad 10
\end{align*}
\]
\( ca_{\text{min}} \) is the VAD for RO attributes, which are determined based on [5, 8, 25]|
| 10  | Condition verification                                                       | \[
\begin{align*}
\text{if} & \quad CE[i] \notin IS \rightarrow 11 \\
\text{else} & \quad 2
\end{align*}
\]
| 11  | Search for global maximal (GMAX) value \( CE \) in the operating range of RO attributes | \( ca^* = \max \{ \max_{ct_i} CE_{ct_i} \} \) and \( CE[i] = \text{extrem}CE_{ct_i}[i] \) |
| 12  | Based on methods [5, 8, 25] and others, optimal parameter of fields \( ca \) of RO attributes for the container is defined | \[ A_{ax}^n = \text{lim}_n - ca_n \cdot ca_{\text{min}} \]
\[ A_{ax}^n = \text{lim}_n + ca_n \cdot ca_{\text{min}} \]
| 13  | Procedure of splitting BSF of RO into 4 clusters:                           | \( \{CT_{m}[\xi]|m = 1, 4\} \)
| 13.1| Binary matrix of cluster is defined \( CT_i \)                              | Under conditions: \( cr(\text{ct}_i \oplus \text{ct}_j) \rightarrow \text{min} \) & \( cr(\text{ct}_3 \oplus \text{ct}_j) \rightarrow \text{min} \) where \( \text{ct}_1, \text{ct}_2, \text{ct}_3 \) are the standard implementations of clusters \( \{CT_{m}[|m = 1, 3\} \)
| 13.2| Value of radii of cluster \( CT^m_i \) is set as \( 0^* \)                  | \( cr_i[\xi] = 0 \)
| 13.3| Determining RO realizations, which arrived to cluster \( CT^m_i \)          | Rule: \( \text{ct}_i \in CT^m_i \) if \( cr(\text{ct}_i \oplus \text{ct}_j) \leq \text{ct}_i[\xi] \) where \( \text{ct}_i \in \{\text{ct}_{i,j}\} \) are the implementations of BLM \( \|ct_i\| \)
| 13.4| Calculation of current ICFE                                                | Expression – stage 6.4. |
| 13.5| Formation \( \{ct_{ax}\} \) of standard implementations for clusters \( \{CT_{m}[\xi]\} \) | Rule for defining coordinates:
\[ ct_{ax} = \begin{cases} 
1, & \text{if} \quad \frac{1}{n} \sum_{j=1}^{n} cr_j[\xi] > \frac{1}{2}; \\
0, & \text{else}
\end{cases} \]
| 13.6| Conditions verification                                                      | \[
\begin{align*}
\text{if} & \quad cr_i[\xi] < cr(\text{ct}_i \oplus \text{ct}_j), \\
\text{cr}_i[\xi] & < cr(\text{ct}_i \oplus \text{ct}_j), \\
\text{cr}_i[\xi] & < cr(\text{ct}_i \oplus \text{ct}_j) \rightarrow 8.3 \& 8.8; \\
\text{else} & \quad 8.9
\end{align*}
\]
| 13.7| The next RO attribute in cluster \( CT^m_i \) is added                      | \( \text{ct}_i = \text{ct}_i + 1 \)
| 13.8| Optimal radius of container \( CT^m_i \) is determined                     | At conditions: \( cr_i[\xi] < cr(\text{ct}_i \oplus \text{ct}_j), \\
\text{cr}_i[\xi] & < cr(\text{ct}_i \oplus \text{ct}_j), \text{ct}_i[\xi] < cr(\text{ct}_i \oplus \text{ct}_j) \)

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 – \( \{vad_{k,i}\} \) 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 \( \{c_{k,i}\} \). The stages of splitting FS of RO into clusters are presented in tabular form in Table 2.
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]:

\[ \text{LM}_{kl}[\text{implementation}][j], \quad (13) \]

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.
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:

$$\max \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} s_{ij} \ominus p_{ij} \ominus p_{ij}^{*} \right\} = \max CE \{ W(\xi) \}$$

where $\ominus$ is the triangular norm (T-norm) [5, 28]; $W(\xi)$ is the fuzzy assessment of usefulness of the $i$-th option of solving the problem of recognition of anomalies or cyber attack, which is determined by value of ICFE. $p_{ij}$ is the assessment of CHS states in the process of RO recognition; $p_{ij}^{*}$ 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 CHS (under conditions of countering the targeted cyber attacks [1, 7, 10, 16, 18]).

In accordance with recommendations of [8, 20, 21, 25], multilayered 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 [23, 24, 15], for virus attacks, 7–15 [5, 7] attributes. Fig. 1, a–e, shows dependencies of ICFE learning of simulation model (SM) of ASR [23] on 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 decision trees, 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.

An analysis of results of the simulation experiments (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 CE = 3.19; for attacks of the Probe class CE = 3.15; for attacks of the R2L class CE = 2.84; for attacks of the U2R class 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 class: DoS/DDoS – cr_{hy} = 4; Probe – cr_{hy} = 3; R2L – cr_{hy} = 4; U2R – cr_{hy} = 4; BA – cr_{hy} = 5; for hy class: DoS/DDoS –
cr_1^2 = 2; Probe – cr_1^2 = 1; R2L – cr_1^2 = 1; U2R – cr_1^2 = 1; BA – cr_1^2 = 2; for hy_3^2 class: DoS/DDoS – cr_1^2 = 3; Probe – cr_1^2 = 2; R2L – cr_1^2 = 2; U2R – cr_1^2 = 2; BA – cr_1^2 = 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.

---

**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**

<table>
<thead>
<tr>
<th>No.</th>
<th>Accepted hypotheses for RO</th>
<th>Values of optimal RC cr</th>
<th>DoS/DDoS</th>
<th>Probe</th>
<th>R2L</th>
<th>U2R</th>
<th>BA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic hypotheses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Basic working hypothesis – hy_1^2</td>
<td>cr_1^2 = 4 – 5 cr_1^2 = 3 – 4 cr_1^2 = 4 – 5 cr_1^2 = 4 – 5 cr_1^2 = 5 – 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Hypothesis hy_2^2 – attribute (attributes) allows drawing a conclusion that indicator IE is lower than the norm</td>
<td>cr_2^1 = 2 – 3 cr_2^1 = 1 – 2 cr_2^1 = 1 – 2 cr_2^1 = 1 – 2 cr_2^1 = 2 – 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Hypothesis hy_3^2 – attribute (attributes) allows drawing a conclusion that indicator IE is higher than the norm</td>
<td>cr_3^1 = 3 – 4 cr_3^1 = 3 – 4 cr_3^1 = 2 – 3 cr_3^1 = 2 – 3 cr_3^1 = 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Additional hypotheses for simulation model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Hypothesis hy_4^2 – node of CIIS demonstrates increased network activity</td>
<td>cr_4^2 = 4 cr_4^2 = 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Hypothesis hy_5^2 – node of CIIS demonstrates increased activity during external traffic</td>
<td>cr_5^2 = 3 cr_5^2 = 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As was shown by data analysis, for IM, Fig. 1–3, quasi-optimal value of parameter $c_{\text{ca},i}$ of VAD equals VAD = 8–16% at maximum value of $C_{\text{max}}=6.16$.

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 CIHS.

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


