Adaptive vulnerability detection model for unmanned vehicles drugs based on artificial immune systems

A V Skatkov, A A Bryukhovetskiy and D V Moiseev
Sevastopol state university, 33 Universitetskaya str., Sevastopol, Russia, 299053

E-mail: dmitriymoiseev@mail.ru

Abstract. The adaptive model and structure of the intrusion detection system (IDS) based on the analysis of traffic in the channel "UMV-dispatch center", which is based on immunological principles, is proposed. Rules classify objects belonging to several classes at the same time with different degrees of belonging. Recognition of the state of network traffic is carried out in conditions of lack of a priori information about the properties of the intrusion source and the stochastic nature of the recognized events. To increase the level of reliability of intrusion detection in the system, adaptive dynamic adjustment of decision-making rules on classification and operational States of UMV traffic is performed.

1. Introduction
Many scientific and practical studies of Russian scientists are devoted to the problem of creating the theory of analysis and synthesis of digital networks with the integration of services that ensure the transfer of various types of information in a single way in a common physical environment. Access to information resources also requires information security. Papers [1-3] are devoted to the development of models of protection mechanisms in information and computer networks and the study of their effectiveness. A special place is occupied by the problem of ensuring the security of critical infrastructures [4], which includes the "smart city" system, which is a large-scale cyber-physical system that coordinates interactions between heterogeneous physical devices and computer systems in real time. The heterogeneity of applications and wireless communications significantly complicates the security of such systems. With regard to the UMV as a cyber-physical object in the infrastructure of the "smart city", we can distinguish three zones of vulnerability:

- Traffic management systems: data stores (navigation maps, routes) and dynamic data and command streams transmitted over various networks, processed in automated systems and presented on display in graphical or textual form;
- Technical infrastructure: technologies, system software, devices with which the implementation of the main actions for the management of UMV is carried out.
- Information interaction of the subjects of the "smart city" with the use of information received from the UMV (transmitted by the UMV) and processed through the technical infrastructure.

Therefore, the development of methods to ensure the security of information interaction between UMV and other subjects of the "smart city" through software interfaces is an urgent task.

The article considers an adaptive approach based on the methods of decision trees in combination with immunological methods used to optimize the parameters of the model [6, 7].
2. Model description
A method of dynamic formation of solving rules for the problem of K-dimensional classification of assignment of vectors $x = (x_1, ..., x_n)$ with n continuous attributes to one of the $C/l=1,k$ classes. In accordance with a known arrangement for the functioning of the IMS [1] and using the production rules of the form "IF "CONDITION" THEN "ACTION" proposed a classification rule of the current state (CS) traffic from the set of all rules $R$ to imagine how:

$$R_j: \text{IF } B_{j1} \text{ AND } ... \text{ AND } B_{jn} \text{ THEN CLASS } C_j | f_j$$

(1)

where $R_j$ is the label of the $j$-th decision rule, $j=1,2,...,N$; $N$ — the number of decisive rules; $B_{ji}$ -a logical variable whose value is determined by a binary relation of the form $x_i Q t_i$, built on the input interval of each numeric attribute, $t_i$ — the threshold value for $x_i$; $Q$ — the ratio, built using the comparison operations$=,>,<$, etc.; $C_j$ — one of the classes of possible States of network traffic, provided that $f_j$ —the numerical value of the classification quality function of the rule $R_j$, which determines the degree of belonging of the vector $x$ to the class $C_j$. Evaluation of the $f_j$ function of the quality of $R_j$ rules is carried out according to the results of training using data from the KDD Cup 1999 network traffic sample database [8] and a special proposed metric:

$$f_j = w_1^j a_j / |A| - w_2^j b_j / |B|,$$

(2)

where $a_j$ — the number of cases of correct classification of abnormal States (AS) traffic; $|A|$ - the total number of AS, $b_j$ -the number of erroneously classified normal States (NS) traffic, $|B|$ - the total number of NS; $w_1, w_2$ — nonnegative weights; $w_1 + w_2 = 1$. Based on the data defining the specified metric, estimates can be obtained for other metrics, namely, DR/DR* - the level of likelihood of intrusion detection in the setup and test mode/in the working mode; FAR/FAR* - the level of false positives of IDS in the setup and test mode/in the working mode. These estimates are most sensitive to the state of network traffic. The values $w_1$ and $w_2$ are assigned by the expert (LPR) on the basis of taking into account the specific features of the problem. The coefficient $w_1$ evaluates the contribution of the number of correctly detected attacks and $w_2$-the contribution of the number of false alarms.

The change domain $f_j$ for each rule is the segment $[-1; 1]$ and is defined for each rule of the set $R$, the form of formula (2) is not the only possible one. Similarly, you can design other relationships to correct estimates that take into account the characteristics of controlled events.

In order to ensure the necessary quality of traffic state recognition, it is proposed to dynamically form subsets of such rules on the basis of the set of all rules $R$, each of which has a metric value $f_j$ above the specified threshold value $f_a$, which should be determined by the LPR depending on the level of criticality of the controlled object.

$R_j$ rules with $f_j > f_a$ form a subset of effective rules - $R_a$; for $R_j: \forall j R_j \subseteq R_a : f_j > f_a$.

The rest of the rules are included in the set $R_0 (R = R_0 \cup R_a)$. Due to the use of the rules included in the $R_a$ in the classification of network States, a high value of the likelihood of intrusion detection is provided, as well as a minimum number of false alarms. This statement is true in the conditions of stationary processes occurring in the network. With qualitative changes in traffic, it is advisable to adjust the set of rules $R_a$ taking into account the new values of $f_j$. Initially, there is no a priori information about the values of the $f_j$ function. The quality of the rules included in the $R_a$ is integratedly estimated by the metric $F_{ap}$, which is constructed as the average value $f_j$ of the quality metrics of the individual classification rules:

$$F_{ap} = \frac{\text{avg}}{R_j \in R_{ap}} \{f_j\},$$

(3)
The set $R_{\alpha}$ constructed in this way corresponds to the group of conditions $B_{ji}$ included in (1). When varying $B_{ji}$, a new subset of $R_{\alpha}$ is formed. All such subsets of $R_{\alpha}$ are included as elements in the set of active rules of $R_{\alpha}$:

$$R_{\alpha} \subseteq R_{\alpha c},... R_{\alpha c} = \bigcup_{p=1,|R_{ac}|} R_{ac}.$$

(4)

Then each such subset is characterized by an integral estimate $F_{\alpha p}$, where $p=1,..,|R_{ac}|$. The most efficient subset of $R_{\alpha}$ is chosen by the condition:

$$F_{\text{max}} = \max_{R_{\alpha} \subseteq R_{ac}} \left\{ F_{\alpha p} \right\}.\quad (5)$$

A subset of $R_{\alpha}$ satisfying condition (3) is denoted as $R_{\alpha m}$. Accordingly, for a subset of $R_{\alpha m}$, we introduce the threshold $f_{\alpha m} > f_{\alpha}$, which will be used to determine the quality of traffic state classification.

The structure of IDS is shown in Fig. 1. It highlights the following main functional modules: training, testing, classification of traffic conditions, decision-making on the assessment of the current state of traffic, adaptation, decision-making on the modification of the rule base. The training module implements preprocessing procedures and training itself. With the help of the preprocessing procedure, the most informative signs of network traffic are identified. The main goal is to reduce the dimension of the processed vectors and increase the reliability of traffic state classification. The procedure is based on the estimation of a priori probability about the values of features and a posteriori probability of belonging of vectors to the given classes.

Figure 1. Intrusion detection system structure based on adaptive model.

The algorithm generates and optimizes the structure of decision rules, which is based on iterative procedure of examples training set and consists of two stages: learning – based method of constructing decision trees [9], optimization – based techniques, artificial immune systems [6, 7]. At the same time, the n-dimensional numerical vectors of the training sample act as antigens, and the decisive rules act as antibodies. For each $R_{j}$ rule of a given class, the value of the quality function $f_{j}$ is calculated according to (2). The learning procedure is repeated for the rules of each class.
After the end of the training and optimization stage, a set of \( R_a \) rules of the form (1) was formed, which can be considered in terms of AIS as a population of antibodies. The functioning of the system is as follows. According to the results of the current solution of the classification problem, the current vector \( x = (x_1, ..., x_n) \) the state of the traffic that comes to the input of the analysis module. Initial values of \( f_j \) rules quality were determined at the system training stage using training samples with known responses from the KDD Cup database. For the \( R_a \) rules included in The \( R_{an} \), a binary signal \( Y \) is generated in the case when the CS is correctly classified. This binary signal is compared with the binary signal \( S \), which for each sample \( X \) a priori determines its belonging to the class \{NS, AS\}. Using these binary signals, the quality function values are corrected in the training mode.

The classifier is proposed to be implemented in the form of a structure, which includes modules: the formation of estimates \( f_j \) for each set of rules, the formation of a subset of effective rules \( R_{an} \), determining the maximum values of \( f_j \). At the first level, the input traffic \( X \) is recognized and the conclusion about its normality / abnormality is made: \( CS=\{NS, AS\} \). At the second level, rules are applied to detect the type of abnormal traffic: DoS, probing, r2l, u2r.

In order to clarify the model of the classifier, we define the classification procedure as:

\[
(CS) \overset{CLASS (R_a)}{\rightarrow} \overset{WHERE R_a= R_a(\text{NS}) \cup R_a(\text{AS})}{\rightarrow} R_a(\text{NS}) \quad \overset{\text{a subset of the rules of recognition of the national Assembly;}}{\rightarrow} R_a(\text{AS}) \quad \overset{\text{is a subset of AU recognition rules.}}{\rightarrow}
\]

Then the rules for recognizing the vehicle using the classification procedure will be defined as follows:

\[
\textbf{IF} (CS) \overset{CLASS (R_a(\text{NS}))}{\rightarrow} \overset{\text{TO} (\text{NS})}{\rightarrow} Y(\text{NS}) = Y = 1 \quad \text{for CS}=\text{NS. Similarly for AS}}
\]

With the correct classification of the current state of the CS, each rule (7) for \( \forall R_j \subseteq R_{an} \) generates a binary signal: \( Y(\text{NS}) = 1 \) For \( \text{CS}=\text{NS} \) for the first rule, \( Y(\text{AS}) = 1 \) for \( \text{CS}=\text{AS} \) for the second, respectively. Otherwise, if the condition \( CS \ll\text{NS} \) for the first rule or \( CS \ll\text{AS} \) for the second rule is satisfied, then the signals \( Y(\text{NS}) = 0 \) and \( Y(\text{AS}) = 0 \) are generated, respectively.

Decision-making under the specified conditions in the rules (7) with the correct classification of the traffic state, in particular, corresponds to the confirmation of hypotheses:

\[
H(\text{NS}|\text{NS}) : \text{CS}=\text{NS} \quad \overset{Y(\text{NS}) = 1}{\rightarrow} \text{for the first rule and } H(\text{AS}|\text{AS}) : \text{CS}=\text{AS} \quad \overset{Y(\text{AS}) = 1}{\rightarrow} \text{for the second rule.}
\]

However, when classifying the current state of traffic, errors of the first and second kind are possible. In this case, the following events occur:

\[
\text{N (NS|AS)} : \text{mistaking the current state of abnormal traffic of the vehicle}=\text{AS for the normal state,}
\]

\[
\text{N (AS|NS)} : \text{mistaking the current state of normal traffic CS}=\text{NS for an abnormal state.}
\]

In accordance with the adopted 2-level classification mechanism, each current state of the vehicle traffic is recognized twice on the sets of \( R_a(\text{NS}) \) and \( R_a(\text{AS}) \) rules. Taking into account this fact, we define the metapravil classification of the vehicle based on the rules of the form (7):

\[
\text{IF} (CS) \overset{CLASS (R_a(\text{NS}))}{\rightarrow} Y(\text{NS}) = 1
\]

\[
\text{AND}
\]

\[
\text{IF} (CS) \overset{CLASS (R_a(\text{AS}))}{\rightarrow} Y(\text{NS}) = 0,
\]

\[
\text{Y(AS)} = Y = 1.
\]

\[
\text{IF} (CS) \overset{CLASS (R_a(\text{HC}))}{\rightarrow} Y(\text{NS}) = 1
\]

\[
\text{AND}
\]

\[
\text{IF} (CS) \overset{CLASS (R_a(\text{AS}))}{\rightarrow} Y(\text{NS}) = 0
\]

\[
\text{Y(AS)} = Y = 1.
\]

\[
(8)
\]
\[
\begin{align*}
\textbf{IF} & \ (\text{CS}) \ \text{CLASS} \ (R_a \ (\text{NS})) = 0 \\
\quad & \text{AND} \\
\textbf{IF} & \ (\text{CS}) \ \text{CLASS} \ (R_a \ (\text{AS})) = 0 \ \text{THEN} \quad " \text{didn't work no one rule }", \ Y=Z. \\
\textbf{IF} & \ (\text{CS}) \ \text{CLASS} \ (R_a \ (\text{NS})) = 1 \\
\quad & \text{AND} \\
\textbf{IF} & \ (\text{CS}) \ \text{CLASS} \ (R_a \ (\text{AS})) = 1
\end{align*}
\]

then " to generate the value of the signal y to perform a special procedure of classification of the vehicle CS".

The first rule system metapravil (8) recognizes the current state of traffic CS as normal, i.e. CS=NS, and generates an output signal \(Y=0\). The second metapravilo - state of CS anomalous, camping on E. CS=NS and \(Y=1\). The third describes a situation when neither of the two rules recognizes the traffic CS, i.e. \(Y(\text{NS})=0\) and \(Y(\text{AS})=0\). In this case, the output signal \(Y\) takes a special value \(Z\). The Appearance of such a situation requires the intervention of an expert who makes a decision on the classification of the vehicle. The last metapravil captures the situation when recognizing the current state of traffic, when both the rules from the set of \(R_a(\text{NS})\) and the rules from the set of \(R_a(\text{AS})\) worked. In this case, a special basic procedure for the classification of traffic vehicles is applied, which is given below.

We introduce notations for the following qualitative estimates of the rules \(R_j \subseteq R_{as}\) classification CS traffic:

- \(dR(\text{AS})/ \ dR(\text{NS})\) — relative number of correctly classified AS/NS border

Then the criterion for assessing the current state of the vehicle traffic is formulated as follows:

\[
F(\text{CS}) = \frac{\sum_{R_j|\text{AS} \in R_{as}} dR_j(\text{AS}) \cdot \left(\frac{f_j(\text{AS})}{f_{\text{as}}(\text{AS})}\right)}{\sum_{R_j|\text{NS} \in R_{as}} dR_j(\text{NS}) \cdot \left(\frac{f_j(\text{NS})}{f_{\text{as}}(\text{NS})}\right)}
\]

(9)

where \(f(\text{AS}), f(\text{NS})\) — quality assessment of traffic state classification rule \(R(\text{AS})\) and \(R(\text{NS})\) respectively;

- \(f_{\text{as}}(\text{AS}), f_{\text{as}}(\text{NS})\) — threshold values for subsets of \(R(\text{AS})\) and \(R(\text{NS})\) — rules-each of which belongs to \(R_{as}(\text{AS})\), \(R_{as}(\text{NS})\), respectively;

- \(dR(\text{AS}), dR(\text{NS})\) — the relative number of correctly classified AS and NS (the level of likelihood in the classification of abnormal traffic rule \(R(\text{AS})\) and normal traffic rule \(R(\text{NS})\)).

Taking into account the entered event designations \(H(\text{H1}|\text{H2})\) in the classification of traffic vehicle estimates \(dR(\text{AS})\) and \(dR(\text{AS})\) are calculated as follows:

\[
dR_j(\text{AS}) = \frac{\sum_{R_j|\text{AS} \in R_{as}} H(\text{AS}|\text{AS}) \cdot j}{\sum_{R_j|\text{AS} \in R_{as}} H(\text{AS}|\text{AS}) \cdot j + \sum_{R_j|\text{AS} \in R_{as}} H(\text{AS}|\text{NS}) \cdot j},
\]

(10)

\[
dR_j(\text{NS}) = \frac{\sum_{R_j|\text{NS} \in R_{as}} H(\text{NS}|\text{NS}) \cdot j}{\sum_{R_j|\text{NS} \in R_{as}} H(\text{NS}|\text{NS}) \cdot j + \sum_{R_j|\text{NS} \in R_{as}} H(\text{NS}|\text{AS}) \cdot j}.
\]

Criterion (9) allows to estimate integrally the level of likelihood of the worked rules \(R_j \subseteq R_{as}\) (NS) and \(R_j \subseteq R_{as}\) (AS). Possible cases in the classification of the vehicle are as follows:

- if condition \(F(\text{CS})>1\) is satisfied, the current state is classified as abnormal AS and \(Y=Y(\text{AS})=1\); if the condition \(F(\text{CS})<1\) is satisfied, then the current state is classified as normal NS and \(Y=0, Y(\text{NS})=1\); if \(F(\text{CS})=1\), expert intervention and additional traffic status information are required to classify the current state.
Correction of rules is carried out as a result of recalculation of values of quality functions of the activated rules by the formula (2). The correction of $f_j$ values for $R_j \subseteq R_{as}$ rules is proposed to be performed in accordance with the following rule:

$$\text{IF } Y \text{ AND } S \text{ THEN CORRECT ( } R_j \subseteq R_{as} \text{ ),}$$

where $\text{CORRECT ( } R_j \subseteq R_{as} \text{ )}$ — is a procedure for correcting the values of the quality functions $f_j$ and $R_j \subseteq R_{as}$ based on the analysis of the values of the signals $Y$ and $S$.

We define the possible cases when correcting the values of the quality functions, which depend on the values of the signals $Y$ and $S$, as well as on the activated rules $R_j$ subsets $R_{as}$:

$$\text{IF } Y=1 \text{ AND } S=1 \text{ THEN CORRECT ( } R_j \subseteq R_{as} \text{ (AS)) AND CORRECT ( } R_j \subseteq R_{as} \text{ (NS)).}$$

Procedure $\text{CORRECT ( } R_j \subseteq R_{as} \text{ (AS))}$ will increase the values of the $f_j$ rules $R_j \subseteq R_{as}$ (AS) in accordance with expression (2) and confirm the hypothesis $H(\text{AS|AS})$ for rules many $R_{as}$ (AS), a — $\text{CORRECT ( } R_j \subseteq R_{as} \text{ (NS))}$ will reduce the values of the $f_j$ rules $R_j \subseteq R_{as}$ (NS), if $R_{as}$ (NS) $\neq \emptyset$, and thus the rules for many $R_{as}$ (NS) were adopted the erroneous hypothesis of $H(\text{NS|AS})$.

Similar reasoning can be given for other correction rules, which are determined by the values of the $Y$ and $S$ signals.

To assess the quality of classification of traffic States in the operating mode of adaptive IDS, the system is also self-tested. For this purpose, the current values of the output parameters of the system $DR^*$, $FAR^*$ are formed during the operation of the SOV, which are compared with the values of $DR$, $FAR$ obtained earlier at the stage of testing and optimization. The decision to modify the rule base is based on the criterion:

$$F_{SOV}=(DR^*/DR)(FAR/FAR^*)$$

If $F_{SOV}<1$, it means that the quality of the classification of IDS has deteriorated and a modification of the rule base is required. In this case, a decision is made to switch the SOV to the training mode on the updated sample $X^*$ and then to the optimization mode, after which the testing is repeated and updated values of the parameters DR, FAR are formed on the basis of the $f_j$ estimates.

The proposed model is locally stable within the selected classes of attacks, as well as sensitive to the configurable parameters of binary relations and the threshold $f_a$ of forming a subset of effective rules $R_{as}$. DR [10], which is most sensitive to the state of network traffic, was used as an assessment of the classification quality. After completing the training and optimization stages for all classes, experiments were conducted on the same mixed sample of 49,198 vectors. As a result, 79 rules describing normal traffic and 92 abnormal abnormal traffic were formed. The length of the rules (number of conditions) for the HC and AC classes varied in the range 6-7, and in the rules for detecting attack types — 9-12. The average likelihood level DR (in descending order) for the classes in the experiments was: $\text{Normal} — 0.986$, $\text{DoS} — 0.979$, $\text{Probe} — 0.925$, $\text{R2L} — 0.889$, $\text{U2R} — 0.728$. The obtained results allow the expert to take into account the specific features of network administration, including in terms of early detection of attacks of these classes.

The offered adaptive model of possible vulnerabilities in UMV on the basis of methods of artificial immune systems can be a basis for IT-technologies of ensuring computer security in the conditions of adaptation at fast change of a condition of network traffic in the channel "UMV-dispatching center". The use of the adaptive model of the decision-making system in combination with the AIS allows to increase the level of likelihood of recognition of events, to minimize the number of false alarms, as well as to ensure high reactivity of the system, which is especially important for the stages of early detection.

**Acknowledgments**
This work was supported by the Russian Foundation for Basic Research (projects № 19-29-06015/19, 19-29-06023/19).
References

[1] Basharin G P and Samouylov K E 2001 Modern stage of development of the teletraffic theory, *Math information* 1 153-66

[2] Zhdanov A A 2012 *Autonomous artificial intelligence* (Moscow: BINOM)

[3] Zegzhda P D, Poltavtseva M A and Lavrov D S 2017 Systematization of cyberphysical systems and assessment of their security? Problems of information security, *Computer system* 2 127-38

[4] Skatkov A V 2012 *Information technologies for critical infrastructures* (Sevastopol: SevNTU)

[5] Pokusaev N, Suprunovskiy V, Katzin V and Namiot D 2019 The ontology of security and Autonomous (unmanned) vehicles, *International journal of open information technology* 7 230-7

[6] Dasgupta D 2006 Artificial immune systems and their application, *TRANS* (Moscow: Fizmatlit)

[7] Roit A, Brostoff J and Mail D 2000 Immunology *TRANS.from the English* (Moscow: Mir)

[8] KDD cup 99 intrusion detection data set [Electronic resource]. - Yeah. text data (752 MB). - DARPA: Irvine, CA 92697-3425, 1999. Retrieved from: / http: // kdd.ics.uci.edu / databases / kddcup99 /

[9] Bryukhovetskiy A A 2013 Detection of vulnerabilities in critical applications based on decision trees, *Electronic and computer systems* (Kharkiv: Khai) pp 18-23

[10] Ji Z and Dasgupta D A 2004 Real negative selection algorithm with variable detectors, *Proceedings of genetic and evolutionary computation, Springer–Verlag*, (Seattle, WA, USA) pp 287-98