Development a Model of a Network Attack Detection in Information and Communication Systems

Abdurakhmonov Abduaziz Abdugafforovich, Gulomov Sherzod Rajaboevich, and Azizova Zarina Ildarovna
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan
Email: {a.abduraxmanov, sherhisor30, z.i.azizova18}@gmail.com

***Abstract*—In this paper the possibility of distribution of Intrusion Detection System (IDS) functionality and Data Mining methods and tools for detecting attacks are analyzed as well variants of placement of the network attack detection system components and application of support vector machine for detecting attacks in a distributed computer network is proposed. The method of principal components which allows to form a feature space for detecting a given set of vectors (network attacks), as well as to reduce the amount of information stored in the base of decision rules necessary for classifying a network. packets, and increase the speed of formation of detection modules is presented. The scheme for applying dimension reduction methods, diagram of the application of clustering methods and its fuzzy inference mechanism is improved. Scheme of formation of detection modules, the variants of placement of functional blocks of the system for detecting network attacks in a separate node and the place of the detection module in the adaptive system are worked out.**

***Index Terms*—Support Vector Machine (SVM), data mining methods, fuzzy logic, clustering methods**

I. Introduction

The rapid development of computer networks and information technologies causes a number of problems related to the security of network resources that require new approaches. At present, the issues of building attack detection systems are a topical trend in the field of information technology. There are many works devoted to the topic of attack detection and classification using a variety of methods, which include traditional approaches based on signature pattern matching and adaptive models using data mining techniques. Most of these works were done quite a long time ago, and some of them have a limited aspect in the form of covering only a specific subject area, namely, the detection of abuse or anomalies.

Currently, the most important process is the global development of information and network technologies used in the construction of large information systems. Based on the full life cycle of information systems, the problem of ensuring the security of information and network infrastructure becomes extremely urgent. To solve this problem, firewalls, anti-virus tools, Intrusion Detection Systems (IDS), integrity control systems, cryptographic protection tools, etc. are used. The most effective tools for detecting network attacks in real time are IDS systems.

Thus, the relevance of the topic of this paper is due to the fact that currently the means and methods of detecting network attacks are based mainly on analytical techniques that allow detecting known attacks. The methodology for detecting network attacks, as a system of all those methods that are used in the field of information protection, is a rather heterogeneous description of both means and methods, and systems for their application in theory and practice. Based on this, the development of model of a network attack detection based on Data Mining methods is extremely relevant.

And so, the scientific novelty of this paper lies in the development of a network attack detection model based on the information exchange model developed for the mathematical model and the use of a forward propagation support vector machine.

II. Application of Data Mining Methods in the Problem of Detecting Network Attacks

In the context of informatization of society, the pace of development of entrepreneurial activity on the Internet becomes explosive. The prospects that the global info communication space opens up for business are so wide that at the moment, for many companies, income is directly related to the ability to access via the Internet. At the same time, the number of malicious actions aimed at blocking access to the server equipment of such companies for legitimate users is also growing. Reliable and secure functioning of modern information systems based on the use of the global network is impossible without the implementation of effective protection mechanisms, including prevention and prevention of network attacks, their detection, tracking the source and countering them.

The traditional model of the IDS intrusion detection system consists of four blocks: data collection, decision rule base, analyzer and reaction block, Fig. 1. The purpose of the study is to organize the protection of distributed computer networks, therefore, unlike host and network

Manuscript received November 24, 2021; revised April 27, 2022.

doi: 10.12720/jait.13.4.312-319

IDS, it is necessary to distribute the components of the detection system network attacks and their parts on various network nodes [1]. Most of the existing IDSs have a monolithic structure, which does not allow efficient distribution of the system's computing load. Fig. 1 highlights the components of the traditional IDS model that can be distributed across the nodes of a computer network.

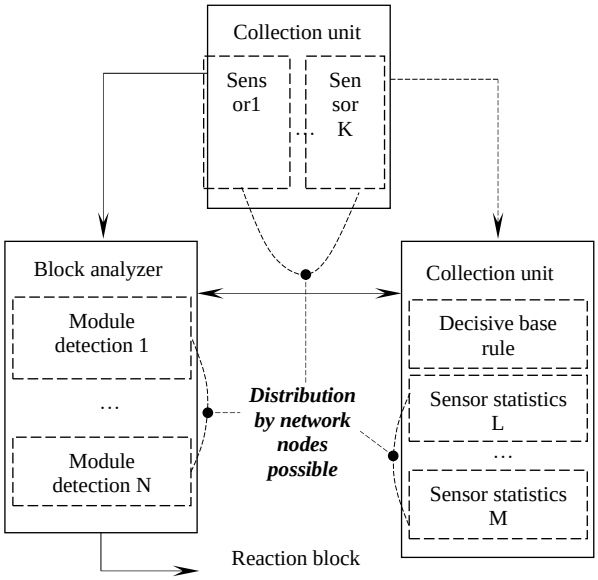

Figure 1. Possibility of distribution of IDS functionality.

Depending on the software and hardware components of individual nodes of the computer network and network equipment, it is possible to place a different composition of sensors, detection modules and database components. For the effective functioning of IDS on the nodes, the following conditions must be met:

- The composition of the detection modules must correspond to the set of potentially possible types of attacks, which can be determined based on the information processed on the host and the installed software;
- The composition of the sensors should correspond to the analyzed information flows, only those network protocols should be processed through which an attack is possible;
- The internal structure of the database must match the sensors and detection modules involved.

Fig. 2 shows options for placing IDS components in a distributed computer network.

The computational load of IDS on a node can also vary by creating a single database for several nodes and by placing only a block of sensors on separate nodes. This approach can significantly increase the volume of network traffic to the node containing the common database, and increase the risks of a denial-of-service attack for nodes that do not have their own analyzer box.

Based on the reviewed studies on the application of Data Mining methods, the following subtasks of the network attack detection process can be distinguished:

1) Element classification - assigning the analyzed vector to the class of normal or abnormal.
2) Search and clean up training data for noise and outliers.
3) Breaking down training data to optimize the traffic analysis process.
4) Determination of the necessary and sufficient set of parameters retrieved by the sensor to classify a specific set of network packets as an attack class.
5) Automatic formation of a modular architecture using the clustering procedure.
6) Introduction of an additional level of signal verification of individual modules to reduce the number of false positives.
7) Assigning the detected anomaly to a known attack or class of attacks.

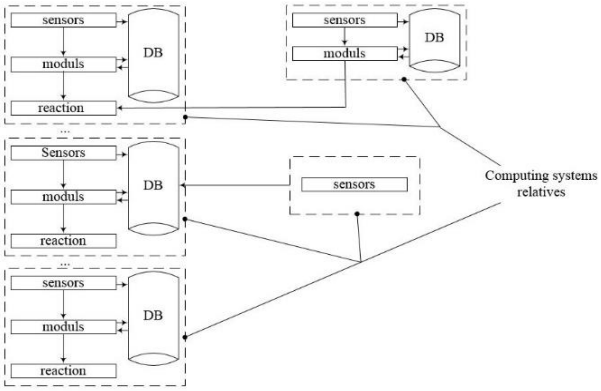

Figure 2. Options for placing IDS components in a distributed computer network.

To solve the listed subtasks, various Data Mining methods are widely used, and leads to a change in the methods of applying Data Mining methods to solve other problems [2]. In accordance with the subtasks presented above, related to the detection of network attacks, several groups of Data Mining methods can be distinguished, which are presented in the Table I with the implemented functions.

TABLE I. DATA MINING METHODS AND TOOLS FOR DETECTING ATTACKS

Method name Data mining	Functions implemented in the problem of detecting network attacks
Classification methods	Assignment of analyzed vectors to sets of normal and anomalous
Reduction methods dimensions	Increasing performance by generating an optimized feature space
Clustering methods	Building an optimized set of detection modules
Fuzzy logic	• organization of interaction of detection modules; • creation of redundant modular IDS architecture

Most IDSs are based on a classification process that infers an attack or anomalous behavior. Currently, there is a lot of research on the topic of network attack detection. These studies are based on techniques such as neural networks, decision trees, association rules, genetic algorithms, and many others.

Based on the results of the analysis of many studies, the support vector machine was selected as a classifier. This method shows one of the best indicators of attack detection and has ample opportunities for internal configuration.

Support Vector Machine (SVM) is a set of similar algorithms in the "supervised learning" category used in classification and regression analysis problems [3], [4]. This method belongs to the family of linear classifiers. A feature of the support vector machine is the constant reduction in empirical classification error and an increase in the gap between classes. Therefore, this method is often called the maximum clearance classifier method.

The method finds elements that are on the boundaries between two classes, which are called support vectors.

Fig. 3 shows the various cases that arise when using SVM for 2D data:

Examples of dividing planes (a);

Dividing plane with a penalty (b);

Linear inseparability (c).

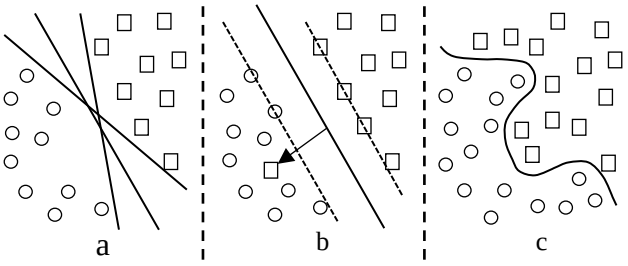

Figure 3. Support vector machine.

Support vector machine searches for a linear function that allows the elements of a dataset to be classified into one of two classes. The problem of binary classification can be formulated as a search for a linear function f (x) that takes values less than zero for elements of one class and greater than zero for elements of another.

The dividing hyperplane is as follows:

$$(x) = w \cdot x - b = 0$$

where w is a vector perpendicular to the dividing hyperplane, the parameter b determines the distance of the hyperplane from the origin. Hyperplanes parallel to the optimal hyperplane and nearest to the support vectors of two classes can be described by the following equations:

$$\begin{bmatrix} wx - b = 1 \\ wx - b = -1 \end{bmatrix}$$

If the training data set is linearly separable, then you can choose the hyperplanes so that not a single point of the training sample falls into the strip between them and then maximize the distance between the hyperplanes. The strip width in this case is $\frac{2}{\|w\|}$, therefore one should minimize $\|w\|$. To exclude all points from the strip, the condition [5], [6] must be met:

$$c_i(wx_i - b) \geq 1, 1 \leq i \leq n$$

where c_i – class label taking values -1 и +1, x_i – vector working sample with class label c_i.

This problem of quadratic optimization is equivalent to the problem of finding the saddle point of the Lagrange function [6]:

$$\begin{cases} -\mathrm{L}(\lambda) = \sum_{i=1}^{n} \lambda_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j c_i c_j (x_i x_j) \rightarrow min_{\lambda} \\ \lambda_i \geq 0, 1 \leq i \leq n \\ \sum_{i=1}^{n} \lambda_i c_i = 0 \end{cases}$$

where L – Lagrange function, λ – Lagrange multipliers.

The solution to this problem is the solution to the quadratic programming problem. As a result of this solution, one can find the vector w:

$$w = \sum_{i=1}^{n} \lambda_i c_i x_i$$

As a result, the classification algorithm can be written as:

$$\alpha(x) = sign(\sum_{i=1}^{n} \lambda_i c_i x_i \cdot x - b)$$

To generalize the SVM to the case of linear inseparability (Fig. 4c), a constant C is introduced - an internal parameter of the method that allows you to adjust the relationship between maximizing the width of the dividing strip and minimizing the total error [5].

The main problem of using the support vector machine in the problem of binary classification is the complexity of finding a linear boundary between two classes. If such a boundary cannot be constructed, one of the solutions is to increase the dimension (transfer of data to another space of higher dimension), where it is possible to construct a plane dividing the set of elements into two classes.

In practice, in view of the linear inseparability of data, instead of a linear kernel (u×v), the support vector machine with one of the kernels is usually used [5]:

polynomial $(\gamma \times u \times v + coef_0)^{degree}$;

radial baseline $exp(-\gamma \times (u - v)^2)$;

sigmoidal $tanh(\gamma \times u \times v + coef_0)$.

For the functioning of the network attack detection system in a distributed computer network, the elements responsible for the interaction between the components of the detection system are also important.

The main element of the system is the detection module - an indivisible part within a distributed computer network that is responsible for detecting certain attacks or anomalous characteristics.

Classification block is an obligatory element of the detection module. With its help, the analyzed multi-bit vectors are marked as normal or abnormal, Fig. 4.

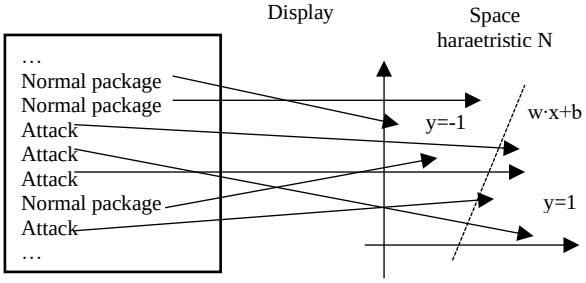

Figure 4. Classification block task.

The application of this method strongly depends on the nature of the data being processed. In particular, there are a number of settings that need to be made before teaching this method. In this regard, to build a high-quality classifier, it is necessary to perform not only training, but also testing of the support vector machine. A simplified scheme for using support vector methods is shown in Fig. 5. To be able to generate the detection module systematically, it is necessary to have an automatic tuning unit that analyzes the constructed SVM model (the number of support vectors) and the results of testing the support vector machine operation (the number of correctly classified packets, errors of the first and second kind) and makes decisions about changing the internal settings of the support vector method. vectors.

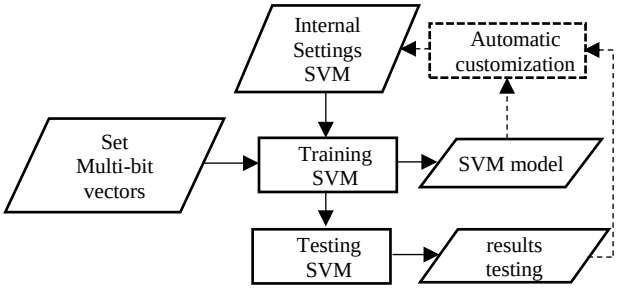

Figure 5. Application of Support Vector Machine (SVM).

To be able to train support vector machines, there are a number of training data requirements. The data retrieved from network traffic is high-dimensional and very large arrays containing a lot of noise and emissions. In this regard, it is necessary to have a data preprocessing block. Dimension reduction methods are the best for this task.

Dimension reduction methods are a wide range of different algorithms, the main task of which is to find a space of lower dimension, in which the internal properties of the original data are preserved [6]. The reduction in dimension can be due to a number of reasons:

- The need to visualize the initial data;
- Simplification of calculations and interpretation of the obtained statistical conclusions;
- The need to compress the volumes of stored statistical information.

From a mathematical point of view, the dimensionality reduction problem can be represented in the following form: given a p-dimensional variable $x = (x_1, x_2, \dots, x_p)^T$ and it is necessary to find a space of lower dimension, in which the variable $s = (s_1, s_2, \dots, s_k)^T$, $k \le p$ reflects the content of the original data in accordance with some criterion. The s components are sometimes referred to as hidden (or latent) components. Different names for p-dimensional variables are used in different cases: the term "variable" is mainly used in statistics, while the terms "feature" and "attribute" are widely used in computer science and machine learning.

There are two global classes of dimensionality reduction methods: linear methods and non-linear methods. For linear methods, the result of each $k \le p$ component will be a linear combination of the original variables:

$$s_i = w_{i;1}x_1 + \cdots + w_{i;p}x_p$$

where $i = 1, \dots, k$. $s = Wx$.

W_{kxp}–matrix of weights of linear transformations. The same relationship can be represented as $x = As$, where A_{pxk} – matrix, and s are the so-called hidden or latent indicators. Then in the notation of observations X_{pxn}, these ratios are as follows:

$$s_{i,j} = w_{i,1}x_{1,j} + \dots + w_{i,p}x_{p,j}, i = 1, \dots, k, j = 1, \dots, n$$

where j corresponds to the sample.

$$S_{kxn} = W_{kxp}X_{pxn}; X_{pxn} = A_{pxk}S_{kxn}$$

Linear methods are more intuitive and much easier to use than more modern methods based on nonlinear transformations. Methods that are limited to the consideration of second-order statistical moments can be distinguished into a separate class. Such methods are computationally simple and involve only classical operations with matrices, without requiring the development of a search procedure in the space of transformation parameters [8]. These are historically the first and most developed methods, among which the most famous are the classical methods: Principal Component Analysis (PCA) and Factor Analysis (FA). Naturally, for a general solution to the problem of representing multidimensional data, it is necessary to overcome two restrictions: the restriction on the linearity of the transformation and the restriction on the consideration of second-order moments.

The choice of a particular method can be based on a priori knowledge of the probabilistic or geometric nature of the data being processed or on the basis of the specifics of the problem being solved. To solve data visualization problems, it is recommended to use methods that use different projections, curves and multidimensional scaling. If rendering is not required, it is better to use principal component analysis or independent component analysis. An alternative solution to any of these problems is to use neural networks.

For the tasks of determining the necessary and sufficient set of traffic parameters for detecting a specific attack, as well as for cleaning the training data from noise and emissions, the principal component method was chosen. The result of his work is the construction of a matrix of weights for calculating new parameters. For this method, the choice of the analyzed matrix is important: the

correlation matrix, covariance matrix, or the matrix of the sum of squares and mixed products.

The dimension reduction unit solves two main problems: it determines a subset of the initial parameters (let's call them basic) and forms a set of parameters in the calculated space (let's call them new). For the principal component analysis, the rule for translating basic parameters into new ones is a linear transformation. The scheme for applying dimension reduction methods is shown in Fig. 6.

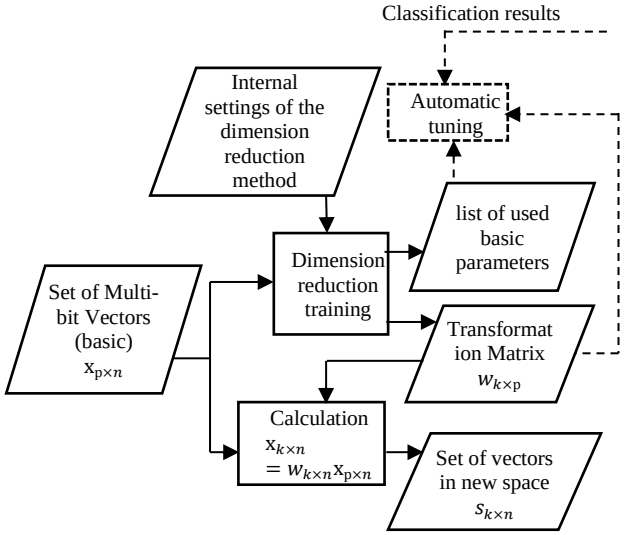

Figure 6. The scheme for applying dimension reduction methods.

Similarly, to the support vector machine for the dimension reduction block, it is necessary to place the automatic parameter selection block.

Clustering methods are a set of algorithms, the purpose of which is to divide a set of objects into groups so that similar objects fall into the same group, and different groups contain objects with dissimilar characteristics. The quality of clustering is characterized by high similarity of objects within each group and high differences between groups. The key requirements for the source data for performing the clustering procedure are data homogeneity and completeness. Uniformity requires that all clustered records be described by a similar set of characteristics.

Cluster analysis is used to solve the following problems:

- Dividing objects into groups in order to optimize further calculations and analysis;
- Reduction of data volume by analyzing only selected cluster representatives instead of processing all data;
- Selection of atypical objects that cannot be attributed to known categories (classes).

A quantitative measure of the proximity of objects is the distance between a pair of objects in space. This value is usually in the range [-1; 1] or normalizes to the interval [0; one].

Similarities between a pair of objects (O_i, O_j) denoted S_{O_i,O_j}, and can be measured in several ways, depending on the distance metric used.

Normalized measure of the difference between a pair of objects (O_i, O_j) is calculated as follows:

$d_{O_i,O_j} = 1 - S_{O_i,O_j}$ of the multitude of clustering methods, two most common categories can be distinguished: hierarchical and iterative methods.

In the course of an experimental study, the following problem was identified: complex distributed attacks consist of many network packets located at a great distance from each other in the space of basic parameters [9].

At the same time, many packets with a normal traffic label are at a short distance from the attack packets. As a result, training of the support vector machine on such a training set in a reasonable time is impossible and the reduction in dimension does not affect the situation in any way. At the same time, using the visualization unit, it was found that similar SVM models are often formed for attacks of similar type. In connection with these observations, it became necessary to redistribute training packages between detection modules:

- Complex attack packets are split into several groups - clusters and processed independently of each other;
- Similar fragments of similar attacks are placed in single detection modules.

Dividing the training set into groups is the task of clustering methods.

Fig. 7 shows scheme of the application of clustering methods.

Clustering is possible both on the entire training set and on a variety of attacks. In the first case, it is possible to select subsets consisting of some attacks or only from vectors of normal traffic, which makes it possible to exclude from the training set stand-alone clusters that do not require training of a classifier.

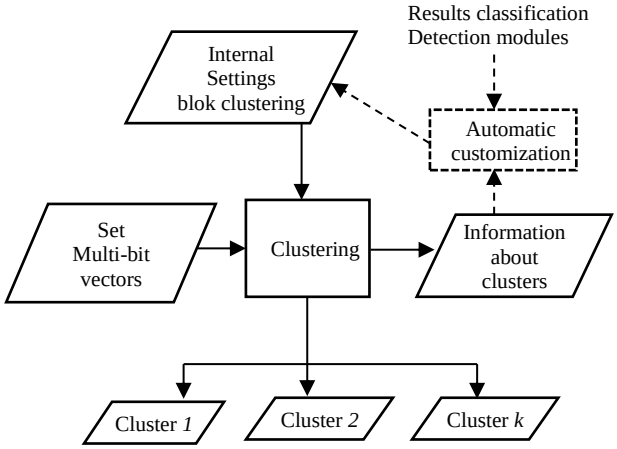

Figure 7. Scheme of the application of clustering methods.

In the second case, a set of clusters is formed that describe the centroids for vectors labeled "attack", which allows constructing relatively simple local SVM-models that allow classifying vectors located near these centers with minimal computational complexity.

III. Fuzzy Logic Apparatus

Fuzzy logic is a superstructure over classical formal logic and set theory. Fuzzy logic is based on the concept of a fuzzy set, built on the membership function.

Element degree x to a fuzzy set C expressed $MF_C(x)$. The membership function of an element to a set takes values in the interval [0; 1], instead of extreme values {0,1} - typical for formal logic: $MF_C(x) \in [0;\ 1]$. A fuzzy set C can be represented as a set of ordered pairs $C = \{MF_C(x);\ x\}, MF_C(x) \in [0;\ 1]$. Expression $MF_C(x) = 0$ means that the element does not belong to the set, 1 - the element belongs to the set.

The most frequently used membership functions in practice are triangular, trapezoidal and Gaussian.

The triangular membership function is as follows:

$$MF(x)\begin{cases} 1-\dfrac{b-x}{b-a}, a \le x \le b \\ 1-\dfrac{x-b}{c-b}, b \le x \le c \\ 0, x \notin (a;c) \end{cases}$$

The trapezoidal membership function is as follows:

$$MF(x)\begin{cases} 1-\dfrac{b-x}{b-a}, a \le x \le b \\ 1, b \le x \le c \\ 1-\dfrac{x-c}{d-c}, c \le x \le d \\ 0, x \notin (a;d) \end{cases}$$

The Gaussian membership function is as follows:

$$MF(x) = e^{-(\frac{x-c}{\sigma})^2}$$

The basic operations of fuzzy logic are intersection and union.

Intersection of fuzzy sets:

$A \cap B: MF_{A\cap B}\ (x) = min(MF_A\ (x), MF_B(x))$;

Union of fuzzy sets:

$A \cup B: MF_{A\cup B}\ (x) = max(MF_A\ (x), MF_B(x))$.

Fuzzy inference is based on a rule base containing membership functions for the corresponding linguistic terms and fuzzy statements:

$$\begin{cases} R_1: if\ x_1 \in A_{11}\ and\ \dots and\ x_n \in A_{1n}\ then\ y \in B_1 \\ \dots \\ R_i: if\ x_1 \in A_{i1}\ and\ \dots and\ x_n \in A_{in}\ then\ y \in B_i \\ \dots \\ R_m: if\ x_1 \in A_{m1}\ and\ \dots and\ x_n \in A_{mn}\ then\ y \in B_m \end{cases}$$

where $\{x_k | k = 1, \dots, n\}$ – input variables; y – output variable; A_{ik} – given fuzzy sets with membership functions.

As a result of fuzzy inference, it is calculated y^* – clear meaning, derived from clear values $\{x_k | k = 1, \dots, n\}$.

The fuzzy inference mechanism consists of four stages, Fig. 8:

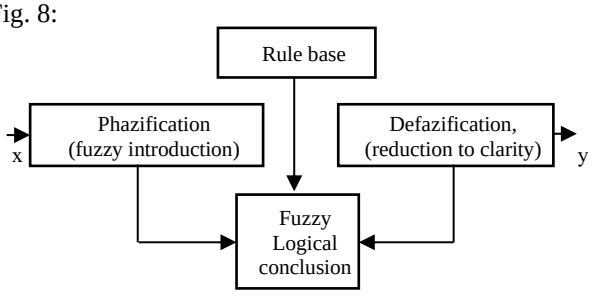

Figure 8. Fuzzy inference system.

- fuzzification (introduction of fuzziness);
- fuzzy conclusion;
- composition;
- dephasing (clarification).

The differences in fuzzy inference algorithms are in the form of the rules used, logical operations and the defuzzification method [10].

When preparing the base of decision rules (training the system), the clustering block is the first functional component in the chain of blocks of the system. Therefore, the time required to train the detection system primarily depends on the quality of the clustering procedure performed. As a result of the analysis, clustering methods for the problem being solved, the k-means method was chosen for clustering large amounts of data and the agglomerative hierarchical clustering method for building an optimal set of clusters on small training sets, for example, on a set of packets of a particular type of attack.

Fig. 9 shows the scheme of formation of detection modules.

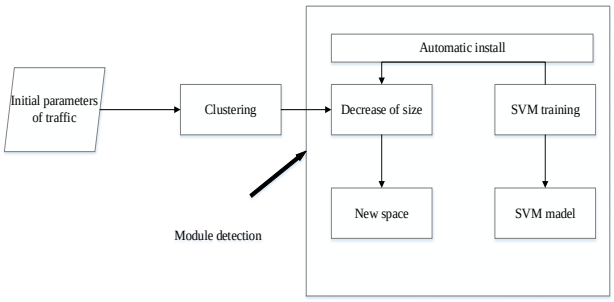

Figure 9. Scheme of formation of detection modules.

The use of clustering methods allows not only to detect complex distributed attacks with a high probability, but also to significantly increase the system performance. By forming several simple detection modules instead of one complex one containing hundreds of support vectors, training time is reduced and the speed of traffic analysis increases.

During the formation of each detection module, the dataset composed of the base traffic parameters labeled normal/abnormal is trained using dimensionality reduction methods. As a result, the least significant basic parameters are discarded and the most important parameters are determined in the new space. In this new space, the dataset is trained on the support vector machine and a dividing hyperplane - the SVM model - is formed [11]. The auto-tuning unit selects the internal settings of the remaining units. The result of the creation of modules are data sets that are placed in the base of decision rules. The scheme of network packet analysis is shown in Fig. 10.

After extraction by multiple sensors, the general list of basic parameters is transmitted to the detection modules. During the operation of the detection system, the module's signals about the detection of potentially dangerous traffic are accumulated in the database. The intrusion decision is made using fuzzy inference rules applied to the current signals of all modules.

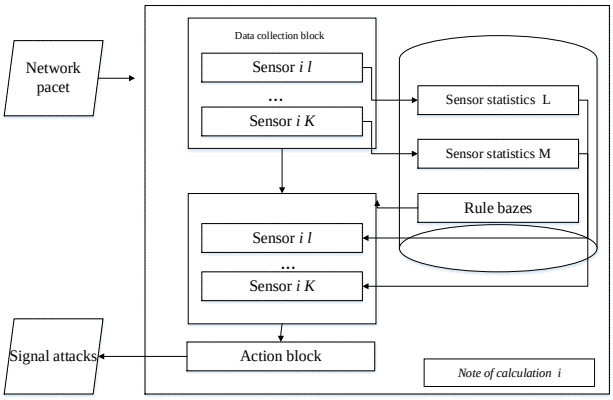

Figure 10. Scheme of network packet analysis.

Distributed computer network nodes have a different hardware platform, operating system, many installed services and user programs, as well as different purposes throughout the network. The potential goals of the offender, his capabilities and the means used depend on these factors. The performance of the node can also be a critical factor, which requires minimizing any auxiliary load, including information security means.

In this regard, it is necessary to build a flexible architecture of the detection system capable of changing the composition of functional blocks and detection modules. An obligatory functional block for the monitored node is a data extraction block - a set of sensors. The reaction unit, analyzer unit and database unit may not be available depending on the requirements for the operation of the node.

Variants of placement of functional blocks of the system and detection of network attacks in a separate node are shown in Fig. 11.

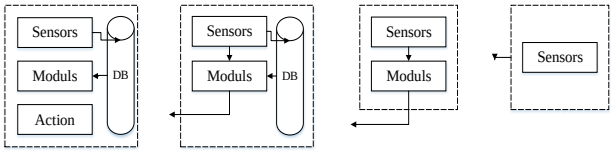

Figure 11. Variants of placement of functional blocks of the system; detection of network attacks in a separate node.

The formation of a common database for several nodes makes it possible to organize a network of simple variants of the detection subsystem, consisting only of a plurality of sensors.

The main adaptation of the detection system is organized through modular architecture. Separate detection modules can be configured to accurately detect a narrow range of attacks or anomalies.

All the possibilities of adapting the functional blocks of the detection system are shown in Fig. 12.

Each detection module is associated with many classes of attacks, to which it responds with a certain probability. In the training set for the attacks under consideration, parameters such as types of potential targets, attack category, and any other customizable characteristics are filled in [12], [13]. To dynamically change the data collection unit, a list of sensors is associated with each detection module, ensuring the extraction of the necessary data from the traffic.

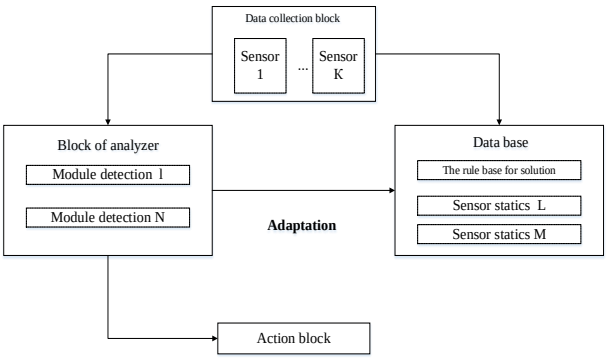

Figure 12. Possibilities of adaptation of functional blocks of the detection system to the hardware and software structure of the node.

The dependence of the detection modules with the components of other functional blocks and the principles for assessing the need for their use are shown in Fig. 13.

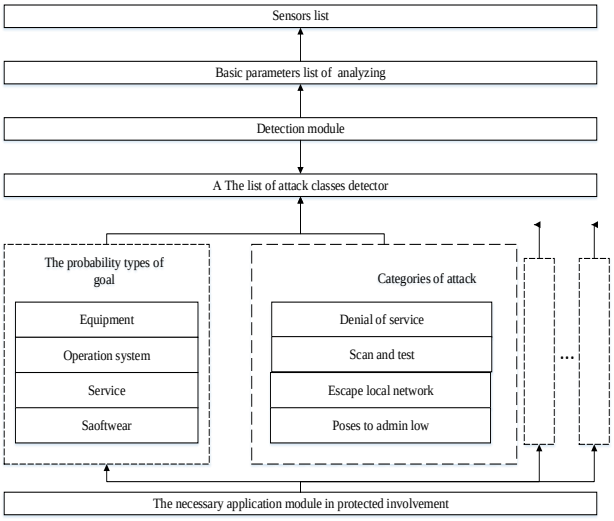

Figure 13. Place of the detection module in the adaptive system.

The adaptability of the system is realized by the ability to automatically change the set of sensors and detection modules used, depending on the structure of the protected software and hardware environment and the set of potential attacks.

IV. Conclusion

In conclusion, it should be noted that improved Data Mining methods for detecting attacks in automatic mode allow to form the necessary and sufficient subset of the parameters of the new space for detecting a specific attack and for the training set a set of attack detection modules based on classifiers with simple SVM models. Moreover, the use of fuzzy logic and fuzzy rules presented in the paper allow improving the performance of classifiers (an extension of the support vector machine), constructing a set of overlapping clusters to increase the probability of detecting network attacks, and classifying the detected attack as attacks known during training.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Gulomov Sherzod Rajaboevich, Abdurakhmonov Abduaziz Abdugafforovich and Azizova Zarina Ildarovna participated in research and wrote the paper. The authors contributed to the analyze and verification of the results obtained and the approval of the final version of this paper.

REFERENCES

[1] S. K. Dey, M. M. Rahman, and M. R. Uddin, "Detection of flow-based anomaly in openflow controller: Machine learning approach in software defined networking," in *Proc. 4th International Conference on Electrical Engineering and Information Communication Technology*, 2018, pp. 416-421

[2] R. K. Malaiya, D. Kwon, S. C. Suh, H. Kim, I. Kim, and J. Kim, "An empirical evaluation of deep learning for network anomaly detection," *IEEE Access*, vol. 7, pp. 140806-140817, 2019.

[3] M. Gao, L. Ma, H. Liu, Z. Zhang, Z. Ning, and J. Xu, "Malicious network traffic detection based on deep neural networks and association analysis," *Sensors*, vol. 20, p. 1452, 2020.

[4] Y. Zhang, P. Li, and X. Wang, "Intrusion detection for IoT based on improved genetic algorithm and deep belief network," *IEEE Access*, vol. 7, pp. 31711-31722, 2019.

[5] J. Ye, X. Cheng, J. Zhu, L. Feng, and L. Song, "A DDoS attack detection method based on SVM in software defined network," *Secur. Commun. Netw.*, p. 9, 2018.

[6] S. K. Dey and M. M. Rahman, "Effects of machine learning approach in flow-based anomaly detection on software-defined networking," *Symmetry*, p. 19, 2020.

[7] Y. Gan, "Application analysis of data mining in the field of computer network security," *China New Communications*, p. 159, 2018.

[8] B. Ghaddar and J. Naoum-Sawaya, "High dimensional data classification and feature selection using support vector machines," *Eur. J. Oper. Res.*, vol. 265, pp. 993-1004, 2018.

[9] S. K. Biswas, "Intrusion detection using machine learning a comparison study," *International Journal of Pure and Applied Mathematics*, vol. 118, no. 19, pp. 101-114, 2018.

[10] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on SDN based network intrusion detection system using machine learning approaches," *Peer Netw. Appl.*, vol. 12, pp. 493-501, 2019.

[11] C. J. Ugochukwu and E. O Bennett, "An intrusion detection system using machine learning algorithm," *International Journal of Computer Science and Mathematical Theory*, vol. 4, no. 1, 2018.

[12] N. Ryabchuk, *et al.*, "Artificial intelligence technologies using in social engineering attacks," in *Proc. CEUR Workshop Proceedings. Vol-2654: Proceedings of the International Workshop on Cyber Hygiene*, Kyiv, Ukraine, November 30, 2019, pp. 546-555.

[13] K. Molodetska, Y. Brodskiy, and S. Fedushko, "Model of assessment of information-psychological influence in social networking services based on information insurance," in *Proc. CEUR Workshop Proceedings. Vol 2616: Proceedings of the 2nd International Workshop on Control, Optimisation and Analytical Processing of Social Networks*, Lviv, Ukraine, May 21, 2020, pp. 187-198.

Gulomov Sherzod (PhD) was born on February 26, 1983 in Shakhrisabz city, the Republic of Uzbekistan. In 2009 he graduated «Information technology» faculty of Tashkent University of Information Technologies. He has more than 160 published scientific works in the form of articles, journals, theses and tutorials in the field Computer networks and Cyber Security. Currently he works as head of the department «Providing Information Security» at the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.

Abdurakhmanov Abduaziz was born on February 20, 1983 in Tashkent, the Republic of Uzbekistan. In 2011 he graduated «Information technology» faculty of Tashkent University of Information Technologies. Currently he works as Vice director in Nurafshan branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi. As the result of long-term research and development work, more than 50 scientific papers, patents for fundamental and applied projects have been published.

Zarina Azizova was born in December 15, 1993 in Fergana city, the Republic of Uzbekistan. He graduated from Tashkent University of Information Technologies named after Muhammad al-Khwarizmi with a master's in Computer Engineering in 2018. He has more than 40 published scientific works in the form of journals articles, conference theses and tutorials in the field of Computer Sciences and Cyber Security. Currently he works as senior teacher of the department «Information Security» at the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.