MINISTRY OF HIGHER EDUCATION, SCIENCE AND INNOVATION OF THE REPUBLIC OF UZBEKISTAN

TASHKENT STATE UNIVERSITY OF ECONOMICS

DIGITAL TRANSFORMATION AND ARTIFICIAL INTELLIGENCE: PROBLEMS, INNOVATIONS AND TRENDS

1st International Scientific - Practical Conference

CONFERENCE PROCEEDINGS

SEPTEMBER 11, TASHKENT 2024

ORGANIZING COMMITTEE

CHAIRMAN AND DEPUTIES

Tulkin Teshabayev

Rector of Tashkent State University of Economics

Gulnora Abduraxmanova

Vice-rector for scientific affairs and innovations

Sultonali Mehmonov

Vice-rector for academic affairs

Komila Karimova

First rector for youth issues and spiritual and educational affairs

Sherzod Sindarov

Vice-rector for international cooperation

Ulug‘bek Xalikov

Vice-rector for international cooperation

Nodir Akbarov

Dean of the Faculty of Digital Economy

Sanjar Mirzaliyev

Head of the Department of Scientific Research and Innovation

MEMBERS OF THE SCIENTIFIC AND TECHNICAL COMMITTEE

Bahodir Muminov

Head of the Department of Artificial Intelligence, Professor

Dilshod Mirzaev

Head of the Department of Information Systems and Technologies

Diyora Khashimova

Faculty of Digital economy, deputy dean

Rashid Nasimov

Associate professor of the Department of Artificial Intelligence

Sharofutdin Xoshimxodjaev

Associate professor of the Department of Artificial Intelligence

Guzal Belalova

Associate professor of the Department of Artificial Intelligence

Dilmurod Mirzaaxmedov

Senior teacher of the Department of Artificial Intelligence

Sanjar Muhammadiev

Senior teacher of the Department of Artificial Intelligence

Muminbek Khayrullayev

Assistant of the Department of Artificial Intelligence

Mamur Shuhratov

Assistant of the Department of Artificial Intelligence

Elyor Egamberdiev

Assistant of the Department of Artificial Intelligence

Ziyoda Norqulova

Assistant of the Department of Artificial Intelligence

ALGORITHM FOR CONSTRUCTING AND CONFIGURING PARAMETERS OF A MODEL FOR SEARCHING FOR TRACES OF ATTACKS IN AN INFORMATION SYSTEM

Obidjon Bekmirzaev,
Tashkent State University of Economics,
Uzbekistan,
bekmirzayevobidjon1989@gmail.com

A B S T R A C T **In the information system, it is not possible to constantly monitor the activity of users, that is, monitoring in real time mode is inconvenient. Therefore, it is important to form the actions of users in the system on the basis of parameters based on their role, and to create rules for searching for attack traces in the future detection of attacks, and to configure models and parameters for searching and detecting attack traces based on these rules.**

K E Y W O R D S information system, attack, user, event, algorithm, model, parameter, confidentiality, integrity and ANFIS

INTRODUCTION

In the era of rapidly developing information technology, the level of threats to data is increasing, in the process, the demand for cyber security has increased. Today, the number of attacks on information systems is increasing, and artificial intelligence systems are used to detect and eliminate these attacks. By applying machine learning and sophisticated artificial intelligence algorithms, organizations automate the critical processes of identifying, analyzing and preventing cyber security threats. These advanced algorithms sift through extensive data sets, enabling early detection of threats and empowering security teams to identify hidden threats and strengthen overall security measures. [1-3].

METHODS

The following traces are taken into account when searching for an attack trace in the information system:

– traces that do not count as an attack, $\left\{I_i^-\right\}, i=1...N, N-$ the number of occurrences;

– marks that are considered an attack, $\left\{I_j^+\right\}, j=1...M, M-$ the number of occurrences.

According to technicians and experts, there are the following classes (categories) of attacks aimed at information security during information processing in the information system [1,2]:

1. Unauthorized access attacks to the information processed in the information system;
2. Special impact attacks on the information system.

Unauthorized access attacks are processes that carry out attacks on external public communication networks and international information exchange related to the actions of intruders who do not have access to the information system, including those who directly carry out attacks on the information system and those who do not have access to the information system.

There are the following types of information security violations caused by an attack directed at the system by an attacker for unauthorized use of information [3]:

- breach of confidentiality;
- breach of integrity;
- break usability;
- breach of authenticity;
- non-repudiation violation;
- compromise accountability, authenticity and credibility.

Taking into account the above considerations, the information system architecture and attack models proposed in the previous paragraphs, the procedure for forming the list of attack sources, and the ANFIS system for the classification of attack traces can be used to build an attack trace search model for the information system.

Also, taking into account the users in the information system, their roles and the events they perform in the system, the following are determined:

User in the information system $\left\{U_i\right\}, i=1...N, N-$ number of users;

Roles in the information system $\left\{R_j\right\}, j=1...M, M-$ number of roles;

Events in the information system $\left\{E_e\right\}, e=1...L, L-$ number of events;

Features available in the information system $\left\{P_k\right\}, k=1...K, K-$ number of features.

Information is processed based on the logical connections of the main elements of the information system. In this case, the main purpose of building a set of events is that a template file is created on the basis of the laws introduced into each system, that is, a log (log file) in which events are recorded. For this, it is necessary to determine the parameters that determine the events. In the general case, four parameters are needed: the sequence $\left\{U_i\right\} \to \left\{R_j\right\} \to \left\{E_e\right\} \to \left\{P_k\right\}$ is fixed based on the given logical binary relations.

This is the main criterion of the search model for the attack trace in the information system. Using Sugeno-Takagi, Takagi-Sugeno-Kang, Wang-Mendel and Mamdani fuzzy inference algorithms of neuro-fuzzy systems, an algorithm for construction, training and testing of ANFIS adaptive neuro-fuzzy systems was proposed. In this case, the dependence on the number of rules for the ANFIS system in validation based on the DataSet sample has the Takagi–Sugeno–Kanga fuzzy inference algorithm. Therefore, ANFIS, an adaptive neuro-fuzzy generation system based on the Takagi–Sugeno–Kang (TSK) fuzzy inference system, is chosen to search for information system attack traces.

The rules for implementing a neuro-fuzzy system model based on the rules for searching for an attack trace in an information system are implemented as follows:

$$R_i : Agar\ (x_i == A_i, \cap \ldots \cap\ (x_{ij} == A_{ij}) \cap \ldots \cap\ (x_{im} == A_{im})$$

$$y = c_{io} + \sum_{j=1}^{m} c_{ij} x_j, j = 1 \ldots n \qquad (1)$$

$$or\ R_i := \bigcap_{j=1}^{m} (x_{ij} == A_{ij})\ \ if,\ in\ it$$

$$y_i = c_{io} + \sum_{j=1}^{m} c_{ij} x_j, j = 1 \ldots r, i = 1 \ldots n \qquad (2)$$

$$R = \sum_{i=1}^{m} R_i - \ a\ set\ of\ general\ rules.$$

In expression (1), the rule base determines the methodology of searching for traces of attacks in the information system.

For cases of searching for traces of an attack on an information system based on the database of an attack in the information system, N - the number of rules is presented as a single one, in fact, it is a set of rules that are separately formed by the type of intruder, the type of information protection tool (for example, SecretNet, Dallas Lock, etc.) is expressed. The object of influence is obtained from the database of information system attack traces, data sets and information about information system infrastructure, information system attack models and design solutions for information.

The final result of determining the scope of compatibility and compatibility of attack traces in the information system is calculated by the sum of the indicators of the described information system attack traces search methodology.

The ANFIS neuro-fuzzy system is based on the following rules [4,5]:

– accuracy of input variables;

– that the membership functions are defined by the Gaussian function:

$$\mu_{ij}(y_j) = \exp\left(-\frac{1}{2} \left(\frac{x_j - a_{ij}}{b_{ij}} \right)^2 \right)$$

in this: y_i – access networks, a_{ij} ва b_{ij} – are adjustable TF parameters.

After defuzzification to obtain the output variable, the functional relationship is expressed as:

$$y' = \frac{\sum_{i}^{n} ((c_{io} + \sum_{j=1}^{m} c_{ij} x_j) \prod_{j=1}^{m} \mu_{A_{ij}}(x_j'))}{\sum_{i=1}^{n} \prod_{j=1}^{m} \mu_{A_{ij}}(x_j')} = \frac{\sum_{i}^{n} (c_{io} + \sum_{j=1}^{m} c_{ij} x_j) \prod_{j}^{m} \exp\left[-(\frac{x_j' - a_{i,j}}{b_{i,j}})^2 \right]}{\sum_{i=1}^{n} \prod_{j}^{m} \exp\left[-(\frac{x_j' - a_{i,j}}{b_{i,j}})^2 \right]}$$

(3)

Expression (3) is based on the ANFIS system using the TSK algorithm, which includes five steps:

RESULTS

Step 1. Fuzzification of input discrete variables follows the following procedure $x_j' \left(j = 1, 2, ..., n \right)$.

Step 2. The elements $a_{i,j}$ and $b_{i,j}$ are calculated with parameters of relevance functions $\mu_{A_{ij}}(x_j')$, weights and values given by Gaussian functions.

Step 3. is multiplied by the results on the elements of the second step and produces the function values, that is:

$$y_i = (c_{jo} + \sum_{j=1}^{m} c_{ij} x_j').$$

(4)

Step 4. The first element of the fourth step is necessary to activate the inferences of rules according to the values collected in step 3, the degrees of relevance of the prerequisites of the rules. The second element of the fourth step is to perform additional steps for the subsequent phaseification of the result of the ANFIS system.

This step consists of a single normalization element, which defines the results of the ANFIS system.

ANFIS TSK neuro-fuzzy systems include 2 parametric steps (step 1 and 3). Parameters that can be adjusted during training of the ANFIS system:

– in the first step–non-linear parameters a_{ij}, b_{ij} are relevance functions of the fuzzifier;

– the third step–parameters c_{io} and c_{ij} of linear functions use conclusions from rule base

$$y_i = (c_{io} + \sum_{j=1}^{m} c_{ij} x_j')$$

.

When there are *n* rules and *m* – input variables, the number of parameters of the first step is *2nm*, and the number of parameters of the second step is equal to *2-n(m+1)*. The total number of adjustable parameters is *n(3m+1)*.

At the next stage of the proposed model, the parameters c_{io} and c_{ij} of the linear functions are calculated under the condition of the specified values of the parameters a_{ij} and b_{ij} .

At the next stage of the proposed model, the parameters c_{io} and c_{ij} of the linear functions are calculated under the

condition of the specified values of the parameters a_{ij} and b_{ij} .

The parameters c_{io} and c_{ij} are found by solving a system of linear equations. And this number represents the output variable in expression (3) in the following form:

$$y' = \sum_{i=1}^{m} w_i'(c_{io} + \sum_{j=1}^{m} c_{io} x_j) \quad (5)$$

in this:

Algorithm for training ANFIS system using TSK algorithm.

With K training examples $x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, ..., x_m^{(k)} y^{(k)}$, where: Substituting the values of the output variables $k = 1,...,K$ and the values of the variables $y'^{(k)}$ the values of the reference variables $y^{(k)}$ a system of linear equations of the form K is obtained [5]:

$$W_i' = \frac{\prod_{j=1}^{m} \mu_{A_{ij}}(x_j')}{\sum_{i=1}^{n} \prod_{j=1}^{m} \mu_{A_{ij}}(x_j')} = \frac{\prod_{j}^{m} \exp\left[-(\frac{x_j' - a_{i,j}}{b_{i,j}})^2\right]}{\sum_{i=1}^{n} \prod_{j}^{m} \exp\left[-(\frac{x_j' - a_{i,j}}{b_{i,j}})^2\right]} = const \quad (6)$$

$$\begin{bmatrix} w_1'^{(1)} & w_1'^{(1)} & x_1^{(1)} & ... & w_1'^{(1)} & x_m^{(1)} & ... & w_n'^{(1)} & w_n'^{(1)} & x_1^{(1)} & ... & w_n'^{(1)} & x_m^{(1)} \\ w_1'^{(2)} & w_1'^{(2)} & x_1^{(2)} & ... & w_1'^{(2)} & x_m^{(2)} & ... & w_n'^{(2)} & w_n'^{(2)} & x_1^{(2)} & ... & w_n'^{(2)} & x_m^{(2)} \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ w_1'^{(k)} & w_1'^{(k)} & x_1^{(k)} & ... & w_1'^{(k)} & x_m^{(k)} & ... & w_n'^{(k)} & w_n'^{(k)} & x_1^{(k)} & ... & w_n'^{(k)} & x_m^{(k)} \end{bmatrix} x \begin{bmatrix} c_{10} \\ ... \\ c_{1m} \\ ... \\ c_{n0} \\ ... \\ c_{nm} \end{bmatrix} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ ... \\ y^{(k)} \end{bmatrix} \quad (7)$$

in this: The total degree of conditions under rule $i-$ when presenting an input vector $w_i'^{(k)}\, k - чи\,(x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, ..., x_m^{(k)})$.

Thus, the abbreviation of (7) is: $w \times c = y$

W is a matrix of size $K \times (m+1)n$, the number of rows is much larger than the number of columns *k*. The solution of this system of equations can be done in one step using the pseudo-inversion of the matrix. *W*:

$$c = w^+ y = (w^T \bullet w)^{-1} w^T y$$

After determining the selected *i,j* - linear parameters, correcting the actual output data of the network for all training samples, calculating and using linear relations for them:

$$y' = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ ... \\ y^{(k)} \end{bmatrix} = w \bullet c$$

where the error vector is defined as: $e = y' - y$, from which the current parameters are determined:

$$a_{ij}^{(k)}(t+1) = a_{ij}^{(k)}(t) + c \frac{dE^{(k)}(t)}{da_{ij}^{(k)}},$$

$$b_{ij}^{(k)}(t+1) = b_{ij}^{(k)}(t) + c \frac{dE^{(k)}(t)}{db_{ij}^{(k)}}$$

An attack trace search model based on the neuro-fuzzy ANFIS system using the TSK algorithm is presented in Figure 1 [6].

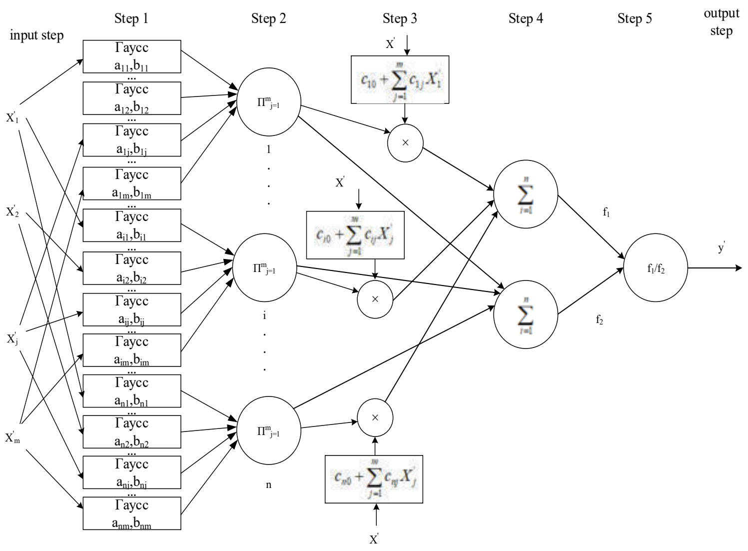

Fig. 1. An attack trace search model based on the ANFIS system

To implement the steps of the neuro-fuzzy system and to search and identify the attack trace of the information system, a software package was developed in a programming language and the calculations were performed on the computer. MATLAB environment was used to compare and describe the studies.

There is an error in training the network with the initial data and parameters of the ANFIS system. During the experiments, it was found that the training sampling error is reduced when refining and changing the initial data set with certain parameters of the ANFIS system.

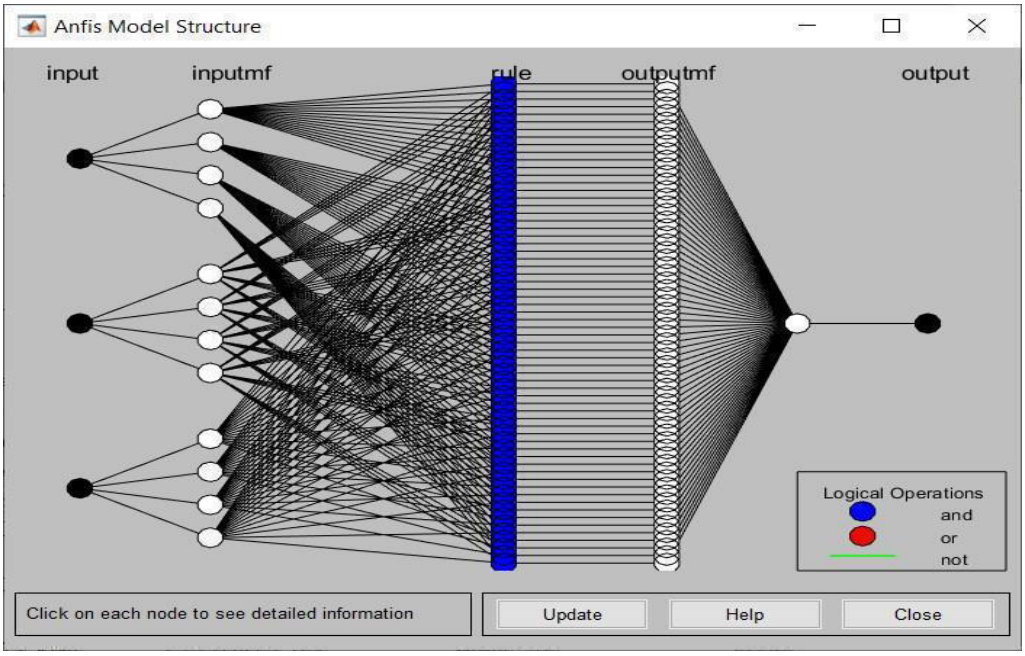

Fig. 2. ANFIS system network structure

The training error of the network was reduced as a result of training based on the dataset generated with the parameters given in the figure above.

DISCUSSION

As a result of training the ANFIS system on the basis of the data sets created with the parameters of the network structure shown in Figures 2, the training error of the network reached a certain range of values, which gave the best result compared to the existing methods.

CONCLUSION

It is inconvenient to monitor users in the information system in real time mode. Therefore, it was proposed to create rules for searching for attack traces in the detection of attacks by shaping users' actions in the system based on their role, and to adjust the model and parameters for searching and identifying attack traces based on these rules.

References

[1] Nuralievich, B. O., & Boltaevich, M. B. (2021, November). Method of Detection and Elimination of Tracks of Attacks in the Information System. In 2021 International Conference on Information Science and Communications Technologies (ICISCT) (pp. 1-2). IEEE.

[2] Nuralievich, B. O., Boltaevich, M. B., & Ugli, B. U. B. (2022, September). The Procedure for Forming a List of Sources of Attack in the Information System. In 2022 International Conference on Information Science and Communications Technologies (ICISCT) (pp. 1-4). IEEE.

[3] Bekmirzaev O., Shirinov B. An Algorithm for Viewing Node State Events Under Attack for Information Systems // AIP Conference Proceedings., 2024, 3147(1), 050003. DOI: 10.1063/5.0210404

[4] Bekmirzaev O., Samarov H. A Method of Evaluating the Effectiveness of Information System Protection // AIP Conference Proceedings., 2024, 3147(1), 050004. DOI: 10.1063/5.0210405

[5] Muminov, B., & Bekmirzaev, O. (2022). Classification and analysis of network attacks in the category of "denial of service" information system. central asian journal of education and computer sciences (CAJECS), 1(1), 7-15.