Combined Neural Network Model for Diagnosing Computer Incidents

Vladimir Avramenko[a], Igor Kotenko[b], Albert Malikov[a] and Igor Saenko[b]

[a]*Military Communication Academy, Saint-Petersburg, Russia*
[a]*Saint-Petersburg Institute for Information and Automation of the Russian Academy of Sciences, Saint-Petersburg, Russia*

Abstract

The basis for making a decision on responding to computer incidents is information about characteristics of the identified security breach, the values of which are determined during the diagnosis procedure. To reduce the time spent on determining the values of characteristics and increase the reliability of the diagnosis results, it is proposed to use machine learning methods implemented on the basis of artificial neural networks. The paper considers a combined artificial neural network as the basis of a model for diagnosing computer incidents. In this model, through the use of deep learning, the drawback of the classical multilayer perceptron associated with the need to form a massive base of training examples for the specific structure of the neural network is eliminated. The results of the experiments showed that the proposed model allows one to reduce the duration of training, while not reducing the values of the quality indicators of the functioning of the neural network.

1. Introduction

Existing methods for protecting information from unauthorized influences in information and communication systems suggests a whole arsenal of information protection tools. The improvement of information protection technologies is primarily aimed at the early detection of masked and unknown scenarios for the implementation of security breaches. After detecting a computer incident, an analysis of the available information is carried out to identify the goals, causes, possible consequences and other characteristics of a security breach. Based on the analysis, response measures are taken. At the same time, standard means of information protection do not allow directly answering key questions of the analysis stage. For this, it is necessary to collect data from almost all the basic elements of the information and communication system and characterize the violation that has occurred, i.e. carry out a procedure of diagnosis. It should be noted that, despite the need to cover the maximum amount of information necessary for making a decision, diagnosis is required to be carried out in the shortest possible time. Prompt and accurate diagnosis allows one to provide a high security and thereby the steadiness

Russian Advances in Artificial Intelligence: selected contributions to the Russian Conference on Artificial intelligence (RCAI 2020), October 10-16, 2020, Moscow, Russia

vsavr@yandex.ru (V. Avramenko); ivkote@comsec.spb.ru (I. Kotenko); mkv.vas@yandex.ru (A. Malikov); ibsaen@comsec.spb.ru (I. Saenko)

of information and communication systems. Thus, the diagnosis of computer incidents is an important stage in managing information protection.

The need to use a variety of information protection tools in the information security system entails the complexity of management processes. Continuous improvement of methods for implementing security breaches and hiding information “traces” requires the use of diagnostic tools, not limited to a set of predefined rules and template, but taking into account the dynamics of events in the information and communication system, resistant to the appearance of external noises and that operate in close to real time. Existing diagnosis methods consisting in studying the contents of information sources (reports of protection tools, event logs of servers and workstations, communication equipment) take a lot of time, and the accuracy of the result depends on the knowledge and experience of specialists involved. There is a contradiction between the existing approach to diagnosing computer incidents and modern requirements for this process.

This paper proposes a model for diagnosing computer incidents based on a combined neural network, which is an extension of the approach proposed in the paper [1]. In the process of functioning of the protection system, new training examples for the diagnostic subsystem appear; accordingly, correction of combined artificial neural networks is required. Learning with new examples should be carried out in the shortest possible time. In order to determine the optimal structure of the neural network, computational experiments were conducted.

The main theoretical contribution of this work is the formulation and solution of a new problem of diagnosing computer incidents based on an optimized combined neural network that provides minimal training time for given requirements for the quality indicators of the neural network.

The paper has the following structure. Section 2 provides an overview of relevant work. Section 3 defines the statement of the problem of diagnosis computer incidents. Sections 4 and 5 consider a combined neural network model for diagnosis computer incidents. Section 6 discusses the experimental results. Conclusions and directions for further research are presented in Section 7.

2. Related work

To carry out the procedure for diagnosing computer incidents, it is necessary to organize effective processing of data from various sources. In works [2, 3, 4, 5], methods of searching, collecting and processing events occurring in the information system, the appearance of which is associated with computer incidents, are presented. It is emphasized that it becomes possible to cope with the analysis of steadily increasing volumes of service information reflected in audit logs only by organizing effective procedures for automating the processing of event logs. Traditional approaches for manual investigation of event logs by a security administrator (without specialized automation tools) are currently extremely inefficient.

A promising way is the automation of routine diagnosis functions. The procedure for collecting data for diagnosis has been automated for a long time and is represented by various solutions [6, 7, 8]. And with the procedure for analyzing the collected data, difficulties arise associated with the high dynamics of the fixed actions of users and events of the information system, the uncertainty of the relationships between events. Moreover, the analysis is carried out against a rather pronounced noisy background formed by events that are not related to a security breach. In order to fulfill the requirements to diagnosis, it is advisable to shift the functions of studying event logs to the "shoulders" of computers. Given the existing experience of diagnosis, this problem can be solved using one of the methods of machine learning supervised learning. Training is carried out by finding patterns in the existing already solved examples of diagnosis, when there is a set of source data and the corresponding correct answer. To implement this method in the tasks of determining categorical output values, machine learning algorithms based on decision trees and artificial neural networks are most widely used. However, the use of decision trees in the problem of diagnosing computer incidents is limited due to the difficulty of adjusting the separating parameters, which can lead to conflicting answers. Artificial neural networks (namely, multilayer direct distribution neural networks) are more preferable from this point of view. However, the quality of the classification result is affected by the dependence of the structure of an artificial neural network on the number of available training examples. At the same time, the dimension of the input and output layers are determined by the conditions of the problem being solved, and the dimension of the hidden layer is selected taking into account the maximization of the quality index of the artificial neural network.

It should be noted that the dimension of the input data set (to cover the events that occurred in a typical information and communication system and affect the outcome of the diagnosis) is quite large and may amount to several tens of thousands. Learning such a neural network requires a large amount of training examples. In this regard, it is necessary to carry out optimization of the structure of the neural network in order to prevent the negative effect of retraining the network on a limited number of examples. To this end, it is proposed to build a combined neural network consisting of the coding part of an auto-encoder and a multilayer perceptron [1]. Passing through the encoding part of the auto-encoder, the input data is converted into group diagnostic signs of a much smaller dimension. Next, the classification of the obtained set of group diagnostic features is carried out on the set of values of one characteristic of a security violation.

This combined neural network allows you to use a well-known and proven approach to solving classification problems for a new subject area, in the analysis of computer incidents.

The use of combined neural networks in the control system of an information security system allows one to: (a) provide systematic information for making decisions on responding to computer incidents; (b) significantly accelerate the processing and analysis of incoming information; (c) increase the accuracy of the determined values of information security violation characteristics; (d) identify the prerequisites for computer incidents. In general, the use of combined neural networks in the problems of diagnosis provides a significant increase in the effectiveness of information protection management. Consideration of previous experience in the application of artificial neural networks is successfully used in many fields, for example, in technical and medical diagnostics [9, 10]. On a time scale close to real, artificial neural networks form the values of the desired characteristics, without requiring human participation in the process of their work.

Examples of the successful use of artificial neural networks in diagnosing failure of information systems equipment are given in papers [11, 12]. Event logs contain data on system failures, which are preprocessed and transmitted for trouble search to the input of artificial neural networks.

Papers [13, 14, 15, 16] propose an approach that uses artificial neural networks to search for attacks and abnormal actions. High accuracy of the obtained experimental forecasts and the ability of the forecasting system to function in a time mode close to real are noted.

The laboriousness of the process of forming training examples for artificial neural networks that are based on supervised learning served as an incentive for the development of deep learning, during which the neural network independently generates distinctive features for input data sets. Effective optimization procedures using unsupervised learning were presented in [17], which demonstrated high performance for deep neural architectures.

Despite the fact that artificial neural networks are a powerful tool for processing data, as evidenced by their widespread implementation in various fields of activity, the issues of automatic diagnosis of computer incidents based on artificial neural networks have not been properly addressed in the known literature. In papers [1, 18], the authors proposed a combined artificial neural network, which allows determining the values of the characteristic of a security violation with high accuracy in close to real time. However, the issue of optimizing the network structure to account for new training examples and when introducing restrictions on the duration of training remained unexplored.

In this paper, we propose an improved model that will eliminate these shortcomings.

3. Statement of the computer incident diagnostic task

Having discovered a computer incident, the information protection system collects data from various sources, which include the event logs of information protection tools, servers and workstations. Of the total volume of this data, only those that are potentially relevant to a breach of information security are of interest. The rest are considered as information noise. In other words, the transformation is carried out from the initial set of events X registered in the information system to the set of informative diagnostic signs $X^{'}$,

$$F : X \longrightarrow X^{'},$$

where $X = \{x_{tsob}^{ist}\}$, $ist = \overline{1, N_s}$, $tsob = \overline{1, N_{ts}}$, N_s - number of information sources, N_{ts} - number of event types, $X^{'} = \{x_{tpr}^{'}\}$, $tpr = \overline{1, N_{tp}}$, N_{tp} - number of types of diagnostic features.

Further, it is necessary to determine the values of the security violation characteristics by the available set of diagnostic signs $hn_i^j \in HN$, $i = \overline{1, N_{ch}}$, $j = \overline{1, N_{zn}}$, where N_{ch} - the number of security violation characteristics, the values of which must be determined during the diagnosis, N_{zn} - the number of possible values of the i-th security violation characteristic. Thus, the solution to the diagnostic problem is to find the mapping of the set of informative events $X^{'}$ to the set of values of the security violation characteristics HN, $F : X^{'} \longrightarrow HN$.

The security breach characteristics are in fact a detailed description of the security violation of information. A list of the main security violation characteristics is given in Table 1.

Table 1
Characteristics of information security violation and their meanings

Name of characteristic	Characteristic values	Type of characteristic
Object of impact	Workstation Server equipment Network hardware	primary
Type of vulnerability	Information Security vulnerabilities Operating System vulnerabilities Application vulnerabilities	primary
Implementation results	The change Delete Creation Blocking, etc.	primary
Detection time	Time instant for detection of violation by information protection means	primary
Attack ID (computer incident ID)	Name of the attack received from information security tools	primary
User ID	Login ID ID received as a result of analysis	primary
Attack Source Address	Network address and port of attack source	primary
Attack Object Address	Network address, port of the object of attack	primary
The purpose of the violation	Information breach Violation of the integrity of information Information access violation	secondary
Source of violation	External Internal	secondary
Stage of implementation	Intelligence service Penetration Implementation Hiding traces	secondary
Nature of the violation	Intentional Unintentional	secondary
Result	Significant damage Minor damage	secondary
Repeatability	First discovered violation Rediscovered violation	secondary
Attack type	DoS attack Scanning Computer virus Software bookmark, etc.	secondary
Risk level	Critical High Average Low	secondary

By the method of determining the values of the security violation characteristics, we divide the characteristics into primary and secondary. The primary ones are those characteristics whose values are determined by direct measurement or calculation (for example, user identifiers, time, and others). Secondary characteristics are those for the determination of the values of which it is necessary to build functional dependences on the values of diagnostic signs.

To use the results of the diagnosis as new training examples, it is necessary to repeat the training procedure. Accordingly, the problem of minimizing the time T_l devoted to training arises. At the same time, it is necessary to ensure the preservation of previously achieved quality indicators of the artificial neural network Q_f. The objective function for the developed diagnostic model can be represented as

$$T_l \longrightarrow min, Q_n \geq Q_f,$$

where Q_n - the value of the quality index of the artificial neural network after each next training, Q_f - the value of the quality index of an artificial neural network, which obtained during first training.

4. Combined neural network model

To achieve the goals of diagnosing computer incidents, information is collected from numerous elements of the information and communication system and the resulting picture of the distribution of recorded events is compared with previously known distributions for which the values of the security violation characteristics are determined.

Due to the lack of formalized analytical dependencies between the events that have occurred and the values of the security violation characteristics, it is proposed to use feedforward neural networks, for example, a multilayer perceptron, to determine such dependencies. In relation to the task of diagnosing computer incidents, a multi-layer perceptron generalizes diagnostic signs with the subsequent determination of the values of the security violation characteristics.

However, in order for the neural network to be able to generalize the available examples, and not just remember them, it is necessary to strike a balance between the volume of the training sample and the number of neurons in the structure of the neural network. To ensure compliance with the quality requirements of the created neural network, it is proposed to combine a multilayer perceptron with the encoding part of the autoencoder, that is, use a combined artificial neural network. This architecture allows one to build a model for diagnosing computer incidents that overcomes the difficulties with the formation of a large number of training examples and allows one to adequately characterize the violation of information security.

A model of a combined artificial neural network for diagnosing computer incidents is presented in Fig. 1, where in - the size of input layer, gr - the size of the hidden layer of the autoencoder, sk -– the size of the hidden layer of the perceptron, m -– the size of the output layer of the combined neural network. The ratio between the number of neurons has the following form: $in \gg gr > sk > m$.

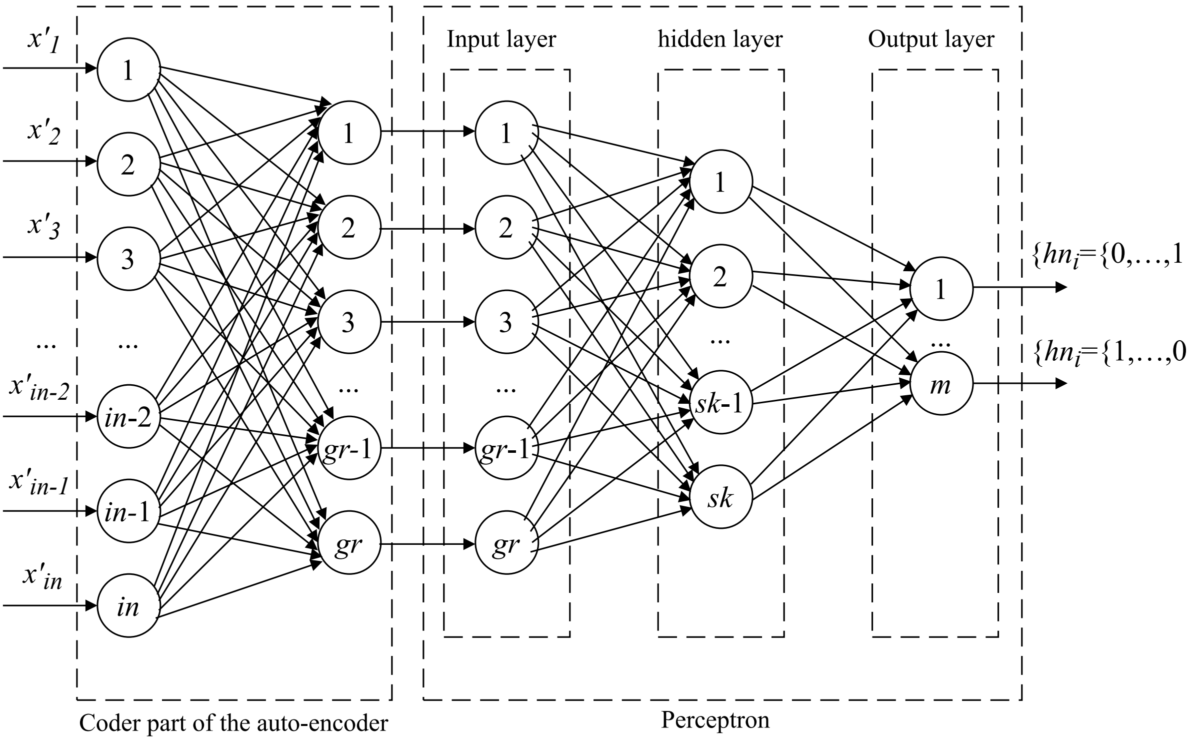

Figure 1: A model of a combined artificial neural network for diagnosing a computer incident

Autoencoder, as it is known, reduces the dimension of the vector of the initial data X arriving at the input of a combined artificial neural network [19]. Thus, the first stage is carried out to generalize the input data (they are compressed), i.e. $F_a : X' \longrightarrow X'_g$. In the hidden layer of the auto-encoder, the group diagnostic features X'_g are formed. Then they go to the input of the multilayer perceptron, in this case a three-layer one, which implements a procedure for processing diagnostic signs and outputs a set of values of the m-ary security violation characteristic, i.e. $F_p : X'_g \longrightarrow HN$.

Thus, the output of the considered model is an assessment of the degree of compliance with the values of the characteristics of the security violation for the set of diagnostic features presented at the input. By training a combined artificial neural network, the weights of neural connections are determined $w_{ij} \in W$, $i = \overline{1, N_{prev}}$, $j = \overline{1, N_{cur}}$, where N_{prev} - – the number of neurons in the previous layer, N_{cur} - the number of neurons in the layer under consideration. These coefficients allow one to record in general form the dependence of the value of the i-th security violation characteristic on the input vector of diagnostic signs X': $hn_i = f_3(f_2(f_1(W_1 \cdot X') \cdot W_2) \cdot W_3)$, where $W_k \in W$, $k = \overline{1,3}$ - the vector of weights of the k-th layer, f_k - the function of activation of neurons of the k-th layer.

5. Implementation of the model

The implementation of the procedure for diagnosing computer incidents in an information and communication system using a combined artificial neural network involves the following sequence of actions:

- collection and processing of data from information sources;
- preparation of an input data set for an artificial neural network;
- formation of group diagnostic features by the encoder part of the auto-encoder;
- classification of the obtained group diagnostic features on the set of values of the desired characteristic of a security violation.

For each secondary characteristic of a security breach, its own neural network and its own set of training examples are formed. The result of the operation of the set of combined artificial neural networks in parallel mode is a set of values of information security violation characteristics, on the basis of which the development of a response option to a computer incident is further carried out.

The model for diagnosing a computer incident in an information and communication system is shown in Fig. 2. After detecting a computer incident, the information security tools report the primary security violation characteristics. At the same time, the process of determining the secondary characteristics of a security violation is launched. The sources of information transmit data to the feature preparation unit, where they are processed and presented to the required form before being fed with combined artificial neural networks. Having the analysis unit came, the normalized values of diagnostic features are processed in parallel mode by several neural networks at once. At the final stage of diagnosis, a final set of values of the characteristics of the security violation is formed, which includes the values of the primary and secondary characteristics.

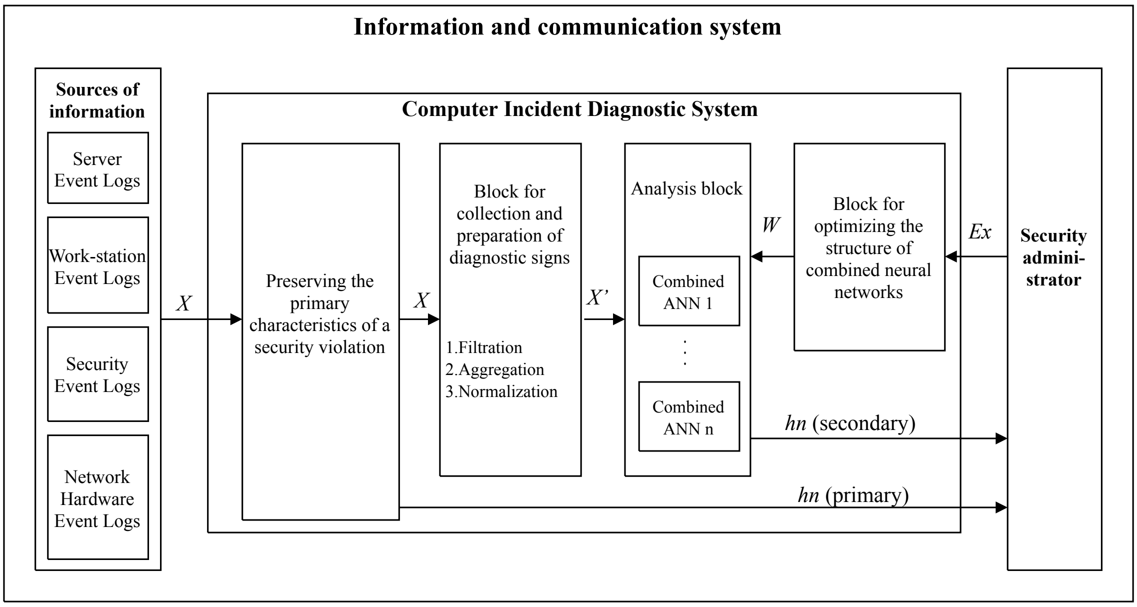
Information and communication system
Sources of information
Server Event Logs
Work-station Event Logs
Security Event Logs
Network Hardware Event Logs
X
Computer Incident Diagnostic System
Preserving the primary characteristics of a security violation
X
Block for collection and preparation of diagnostic signs
1.Filtration
2.Aggregation
3.Normalization
X'
Analysis block
Combined ANN 1
Combined ANN n
W
Block for optimizing the structure of combined neural networks
Ex
Security administrator
hn (secondary)
hn (primary)

Figure 2: Model for diagnosing a computer incident in an information and communication system

The set of initial data and diagnostic results can, if necessary, be used as new training examples (datasets). The number of examples in the training set N_{ex} affects the number of neurons in the hidden layer N_{ck} of an artificial neural network, according to the relation established by R. Hecht-Nielsen [20]:

$$N_{ck} = \frac{N_w}{N_x + N_y}, \tag{1}$$

$$\frac{N_y N_{ex}}{1 + \log_2(N_{ex})} \leq N_w \leq N_y(\frac{N_{ex}}{N_x} + 1)(N_x + N_y + 1) + N_y, \tag{2}$$

where N_y - the number of neurons of the output layer, N_{ex} - the number of training examples, N_w - the number of neural connections, N_x - the number of neurons in the input layer.

Accordingly, with the replenishment of the base of training examples, the structure of the neural network should be adjusted. The choice of the number of neurons in the hidden layer of the auto-encoder and perceptron affects the values of the quality indicators of the combined artificial neural network and the duration of the training. In order to minimize the training time and maintain the values of quality indicators at the required level, it is necessary to solve the optimization problem:

$$\begin{cases} T_l(N_{ck}) \longrightarrow min, \\ Q_n(N_{ck}) \geq Q_f \end{cases} \tag{3}$$

To solve this problem, computational experiments were carried out. Based on the obtained experimental values, in the optimization unit of the structure of combined neural networks, the optimal number of hidden layer neurons is selected and the combined neural networks are adjusted.

6. Experimental results

An experimental assessment of the proposed model for diagnosing computer incidents was carried out on the basis of the existing database of information security violations compiled by experts. The database contains a description of two security violation characteristics: “nature of the violation” and “consequences”. It consists of 276 entries, which include sets of diagnostic features from event logs for 20 workstations. To characterize a security violation, the “nature of the violation” of 276 records, 113 corresponds to a deliberate violation, and 163 - to an unintentional violation. To characterize a security violation, the “consequences” of the implementation of the violation of 276 records, 164 correspond to significant damage, and 112 - to minor damage.

Combined neural networks were implemented in the C ++ programming language. The training was carried out on a workstation under the control of the Windows 7 Professional operating system, 16 Gb RAM, AMD Ryzen 5 2600 processor.

The experiments were carried out according to the following scheme. Initially, a database of training examples was formed, which included many pairs of sets of diagnostic signs and the corresponding values of the security violation characteristics. These pairs are designed to train combined artificial neural networks that are part of the computer incident diagnosis model and test their readiness for operation. Then, new training examples for subsequent adjustment of neural networks are stored in the same database. The training is carried out by the method of back propagation of error, during which the weighting coefficients of each communication between artificial neurons are adjusted [21].

Next, the structure of combined neural networks was formed. The size of the input layer for all combined neural networks corresponds to the size of the vector of diagnostic features. To cover the studied diagnostic features for a long period of time during the experiment, the input layer size was 32400. The output layer of each combined neural network is determined by the number of values of the secondary characteristic of the security violation.

The size of the hidden layers of the autoencoder and perceptron was initially selected based on the minimum sufficient according to R. Hecht-Nielsen’s condition (1) and (2) and amounted to 12 and 1, respectively.

For training and quality control of the artificial neural network, the database was divided into two parts. Initially, 180 notes were taken for training and 76 entries for testing. The quality assessment of the combined neural network was carried out according to the generally accepted indicator F-measure (F), the value of which is calculated on the basis of indicators of precision Pr and recall Rc:

$$F = \frac{2 \cdot Pr \cdot Rc}{Pr + Rc},$$

$$Pr = \frac{TP}{TP + FP},$$

$$Rc = \frac{TP}{TP + FN},$$

where TP is the number of records classified as the true value of the characteristic, while it is true, FP is the number of records classified as the true value of the characteristic, while it is actually false, FN is the number of records classified as the false value, while it is true.

The dependences of the values of the F-measure indicator obtained during the computational experiment on the number of neurons in the hidden layers of the N_{ae} autoencoder and the N_{perc} perceptron for characterizing the security violation "nature of the violation" are shown in Fig. 3, and to characterize the security violation "consequences" - in Fig. 4. The combined neural network was trained on 180 examples.

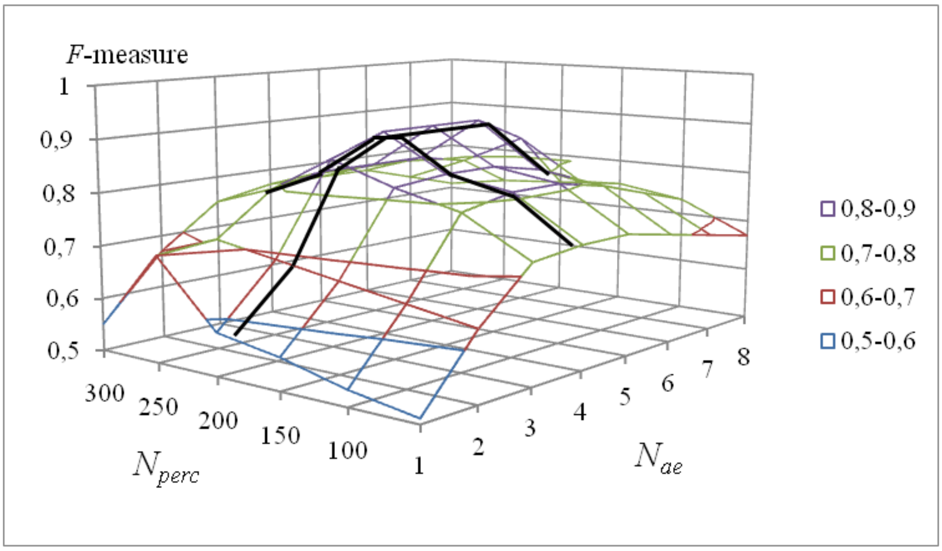

Figure 3: The values of the F-measure indicator for the security violation characteristic "nature of the violation"

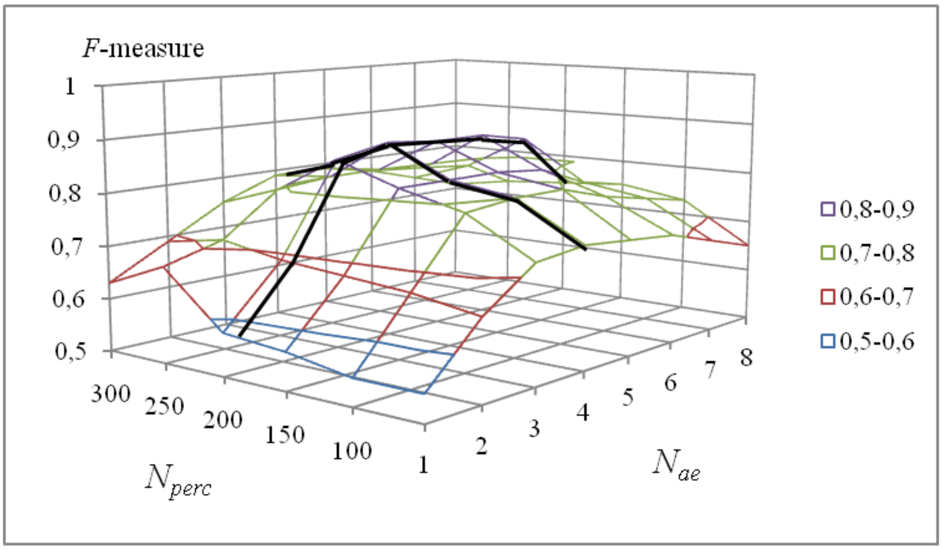

Figure 4: The values of the *F*-measure indicator for the security violation characteristic “consequences”

As can be seen from Fig. 3 and Fig. 4, the combined neural network reaches the maximum value of the *F*-measure index with the number of neurons of the hidden layer of the auto-encoder of about 180 and 4 neurons of the hidden layer of the perceptron. With an increase in the number of hidden layer neurons, an increase in the *F*-measure is not observed and subsequently even decreases due to the occurrence of the negative effect of overfitting.

By increasing the number of examples to 194 and testing on 82 examples, the task of optimizing the training time of a neural network is solved by choosing the minimum sufficient number of hidden layer neurons without losing the quality of the originally created network. The combined neural network has high values of the accuracy indicator for determining the value of the binary characteristic of a security violation with the number of neurons 186 for the hidden layer of the auto-encoder and 4 for the hidden layer of the perceptron. With an increase in the number of neurons, the training time increases, and the values of the accuracy and completeness indicators initially remain the same, but then decrease. With a decrease in the number of neurons of the hidden layers, along with a slight decrease in the training time, the values of the accuracy and completeness indicators decrease. That is unacceptable according to the conditions of the problem.

These results indicate the need to adapt the structure of the combined neural network when replenishing the base of training examples. While for the number of training examples up to 200, the number of neurons in the hidden layer of the auto-encoder is advisable to choose about 185, the number of neurons in the hidden layer of the perceptron is 4. With an increase in the number of training examples by 14, the values of the *F*-measure indicator remained the same with an increase in the number of neurons of the hidden layer of the auto-encoder by 6 and the preservation of 4 neurons of the hidden layer of the perceptron.

In comparison with a multilayer perceptron, a combined neural network allows us to ensure that the requirements for the *F*-measure and training time are met, the values of which are given in Table 2. In this paper, training time the combined neural network is defined as the time period after which the specified requirement for the difference between the received and known output values of the network for all training examples is fulfilled.

Table 2
Comparison of the indicators F-measure and training time for neural network

Number of training examples	Multilayer perceptron		Combined neural network	
	F-measure	Training time, min	F-measure	Training time, min
140	0.70	76.10	0.89	10.32
160	0.72	79.08	0.90	10.76
180	0.72	83.2	0.91	11.01
194	0.72	84.07	0.91	11.17

Thus, the requirement according to expression 4.3 is met. As can be seen from Table 2, the training of a multi-layer perceptron takes much longer on a small number of examples and does not ensure the achievement of the values of the F-measure indicator for a combined artificial neural network.

7. Conclusion

The paper proposed a new model for diagnosing computer incidents based on combined neural networks, and attempted to solve the problem of optimizing the structure of combined neural networks in order to minimize training time when adding new training examples.

The use of an auto-encoder as part of the combined neural network allowed one to switch from an input set of high-dimensional diagnostic features to a compact representation due to compression using the principal component method. The resulting set of group diagnostic features is classified according to the set of values of the characteristics of the security violation. The classification problem is solved by a three-layer perceptron. Reducing the dimension of a set of diagnostic features allows the training of a combined neural network on a fairly small base of training examples.

An experimental evaluation of the proposed model on a test bench made it possible to identify the optimal number of neurons in the hidden layers of the combined neural network in the interest of minimizing the training time.

Given the simplicity of the proposed model, it seems possible its practical application as part of the information protection system of a typical information and communication system.

The directions of further research are associated with the study of the use of combined neural networks to determine the values of the secondary characteristics of information security violations, depending on other previously defined characteristics, including primary ones, by adding recurrent neural networks to the structure. It is also interesting to study the effect of dropout for neurons of hidden layers of the combined neural network on the value of the F-measure indicator.

Acknowledgments

This research is being supported by the grant of RSF #18-11-00302 in SPIIRAS.

References

[1] A. Malikov, V. Avramenko, I. Saenko, Model and method for diagnosing computer incidents in information and communication systems based on deep machine learning, Information and Control Systems 103 (2019).

[2] R. Vaarandi, A data clustering algorithm for mining patterns from event logs, in: Proceedings of the 3rd IEEE Workshop on IP Operations & Management (IPOM 2003)(IEEE Cat. No. 03EX764), IEEE, 2003, pp. 119–126.

[3] Z. Kurd, Artificial neural networks in safety-critical applications, Ph.D. thesis, University of York, 2005.

[4] H.-J. Cheng, A. Kumar, Process mining on noisy logs—can log sanitization help to improve performance?, Decision Support Systems 79 (2015) 138–149.

[5] R. J. C. Bose, R. S. Mans, W. M. van der Aalst, Wanna improve process mining results?, in: 2013 IEEE symposium on computational intelligence and data mining (CIDM), IEEE, 2013, pp. 127–134.

[6] V. I. Gorodetski, O. Karsayev, A. Khabalov, I. Kotenko, L. J. Popyack, V. Skormin, Agent-based model of computer network security system: A case study, in: International Workshop on Mathematical Methods, Models, and Architectures for Network Security, Springer, 2001, pp. 39–50.

[7] I. Kotenko, O. Polubelova, I. Saenko, The ontological approach for siem data repository implementation, in: 2012 IEEE International Conference on Green Computing and Communications, IEEE, 2012, pp. 761–766.

[8] I. Kotenko, I. Saenko, Creating new-generation cybersecurity monitoring and management systems, Herald of the Russian Academy of Sciences 84 (2014) 424–431.

[9] I. Bogachev, A. Levents, E. U. Chye, Application of an artificial neural network for classification of telemetric data in compression systems, Information and Control Systems 82 (2016).

[10] M. V. Vyucheyskaya, I. N. Kraynova, A. V. Gribanov, Neural network technologies in medical diagnosis (review). p. 284–294, Medical biology research 3 (2018) 284–294.

[11] F. Lv, C. Wen, Z. Bao, M. Liu, Fault diagnosis based on deep learning, in: 2016 American Control Conference (ACC), IEEE, 2016, pp. 6851–6856.

[12] D.-Q. Zou, H. Qin, H. Jin, Uilog: Improving log-based fault diagnosis by log analysis, Journal of computer science and technology 31 (2016) 1038–1052.

[13] Q. Fu, J.-G. Lou, Y. Wang, J. Li, Execution anomaly detection in distributed systems through unstructured log analysis, in: 2009 ninth IEEE international conference on data mining, IEEE, 2009, pp. 149–158.

[14] T. Nolle, A. Seeliger, M. Mühlhäuser, Unsupervised anomaly detection in noisy business process event logs using denoising autoencoders, in: International conference on discovery science, Springer, 2016, pp. 442–456.

[15] M. Sakurada, T. Yairi, Anomaly detection using autoencoders with nonlinear dimensionality reduction, in: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, 2014, pp. 4–11.

[16] M. Alkasassbeh, An empirical evaluation for the intrusion detection features based on machine learning and feature selection methods, arXiv preprint arXiv:1712.09623 (2017).

[17] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, Y. Bengio, An empirical evaluation of deep architectures on problems with many factors of variation, in: Proceedings of the 24th international conference on Machine learning, 2007, pp. 473–480.

[18] V. Avramenko, A. Malikov, Diagnosing computer security incidents based on a combined artificial neural network, Information Security. Inside (2019) 72–76.

[19] P. Baldi, Autoencoders, unsupervised learning, and deep architectures, in: Proceedings of ICML workshop on unsupervised and transfer learning, 2012, pp. 37–49.

[20] R. Hecht-Nielsen, Kolmogorov's mapping neural network existence theorem, in: Proceedings of the international conference on Neural Networks, volume 3, IEEE Press New York, 1987, pp. 11–14.

[21] S. Khaikin, Neural networks: full course, M.: Williams 1104 (2006).