#90

BICA*AI 2023

Development of a network traffic anomaly detection system based on neural networks.

N.V. Bespalova[a], A.S. Ershov[b], S.A. Sitnikov[b], S.V. Nechaev[c], M. Vanina[d], V. Radygin[e], D. Kupriyanov[e], M.N. Ivanov[a]

[a] Financial University under the Government of the Russian Federation, 49, Leningradsky ave., Moscow, 125993, Russia, [b]Yuri Gagarin State Technical University of Saratov, Polytechnic, 77, Saratov, Russian Federation, [c] ITMO University, Kronverksky pr., 49, Saint Petersburg, Russian Federation, [d]Moscow Technical University of Communications and Informatics, st. People's Militia, 32, Moscow, Russian Federation, [e]National Research Nuclear University "MEPHI", 31 Kashirskoe Shosse, Moscow, 115409, Russian Federation *Ostozhenka Street, Moscow, 119034, Russia*

FINANCIAL UNIVERSITY
UNDER THE GOVERNMENT OF THE RUSSIAN FEDERATION

SUMMARY

The article discusses the issues of ensuring information security using artificial intelligence technologies. As a result of the work, neural networks created on the basis of various variations of input parameters were trained and tested. The resulting solutions handle incoming network traffic on their own, informing security administrators of any unusual network behavior.

INTRODUCTION

Machine learning methods and neural networks are used to solve a wide range of tasks. One of these goals is the task of monitoring and differentiating network anomalies that can negatively affect the security of information systems. The use of a cyber incident response system based on artificial intelligence will provide the necessary support and will allow processinga large number of events simultaneously. At the moment, there are not many solutions based on the use of artificial intelligence in the field of information security. Table 1 provides a comparative analysis of some of the nicknames.

APPROACH

To train the neural network in the presented work, the method of error back propagation is used, which is based on gradient descent. The calculation of weights is considered based on the Widrow-Hoff rule.

METHODS

In the project under consideration, the data set used for the international competition

of tools for discovery and data mining was chosen as a dataset, which was held in conjunction with KDD-99 - an international conference on knowledge discovery and data

mining. For correct processing, the data were normalized using the maximum value normalization method. After preparing the training data, the parameters of the neural network were set: the number of input and output neurons, as well as the number of hidden layers and neurons on each.

RESULTS

After the first run with the initially selected parameters, the percentage of correctly identified situations was 99.96%. However, the running time of the program is too long, and the formation of a normalized dataset with all parameters requires a lot of memory. To optimize the work, an experiment was carried out with a decrease in the number of
input parameters and maintaining the percentage of successful determination. 12, 25 and 33 output parameters were chosen for testing. The results of the experiment are presented in Table 2. With 25 input parameters, the ratio of indicators percentage of correct definition / waiting time is the most optimal. Studies have also been carried out on changing the hidden layers. The best result was
shown by a network with two hidden layers, in which the first is the number of input
neurons multiplied by 4, and the second is multiplied by 2.

ANALYSIS

The implementation of the neural network is carried out using the Java language and the Neuroph library. Neuroph is a lightweight Java neural network framework for developing general neural network architectures. It contains a well-developed open source Java library with a small number of base classes that correspond to the basic concepts of the neural network.

DISCUSSION

As a result of neural network training, 4 different variants were obtained. The differences are determined by the inverse dependence of the number of input parameters on the data processing time. The network with 41 input parameters showed the best detection percentage, while its operating time and memory usage were the highest. Different configurations of neural networks allow you to adapt them to different operating conditions. Adding new attack options does not require major changes in the developed program code.

CONCLUSIONS

In conclusion, a network incident detection program has been successfully created. As a result of the work, neural networks created on the basis of various variations of input parameters were trained and tested. The resulting solutions handle incoming network traffic on their own, informing security administrators of any unusual network behavior. Further improvement will be able to respond to some types of attacks without human assistance, which will significantly speed up the process of processing and resolving incidents, as well as reduce the burden on human resources.