UDC 681.3:519.68

I.A. Tereikovskyi Ass. Prof. (NTUU "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine), V.V. Pogorelov (National Aviation University, Ukraine), O.I. Tereikovskyi (NTUU "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine)

Determination of structural parameters of a multilayer cyber threat detection perceptron

This study proposes a way to increase the effectiveness of cyber threat detection software. The possibility to increase the effectiveness is shown by adapting a multilayer perceptron structure to the specific task conditions. A mathematical tool is proposed to adapt the structure in terms of the relative error minimization during neural network model generalization.

At present, one of the most important areas of modern cyber threat detection software development is introduction of neural network systems intended for security parameters evaluation [2, 3, 8]. In many respects, the prospects of such systems are associated with the use of neural network models (NNM) based on a multilayer perceptron (MLP) [1, 6]. At the same time, real-life experience with neural network detection systems and research results [2, 3, 7] shows the need to improve MLP efficiency by adapting its structure to the specific task conditions. This statement is based on the assumption that the MLP structure has a noticeable effect on its generalization error, which affects MLP efficiency the most.

The following main parameters determine the MLP structure: number of input neurons (N_x), number of hidden neuron layers (H), number of neurons in each hidden layer, and number of output neurons (N_y). The real-life experience and results [5] in this case indicate that the available solutions (1-3) require significant refinement.

Therefore, the purpose of this study is to develop a procedure for determining the structural parameters of an MLP intended for cyber threat detection based on monitored security parameters evaluation.

Based on the results of [1, 4, 7], we can assume in the first approximation that the number of input and output neurons matches the number of registered parameters and the number of recognizable images. These values are given a priori based on the specified task conditions, and are not subject to change. Therefore, the MLP structure can only be adapted by changing the number of hidden neuron layers and the number of hidden neurons in each layer.

According to [3, 6], MLP structural model development is based on determination of the total number of synaptic connections. In accordance with [6, 7], let us take the criterion for minimizing MLP relative error as the basis for such determination:

$$G_{MLP} \to \min \ , \qquad (1)$$

where G_{MLP} is the MLP relative error.

In this case, the MLP relative error term shall mean the ratio between the generalization error (ε) and the number of synaptic connections (L_w). Hence, the formula (1) can be rewritten in the following form:

$$\frac{\varepsilon}{L_w} \to \min . \tag{2}$$

Based on the results of [1, 4], we can assume that for (2) the generalization error and the number of synaptic connections can be taken as analogue quantities. This assumption allows us to determine the optimal number of synaptic connections (L_w^{opt}) by solving the following equation:

$$\frac{\partial L_w}{\partial \varepsilon} = 0 . \tag{3}$$

Below are the details on the equation (3) components. The generalization error is calculated as the sum of the approximation error (ε_a) and model declaration error (ε_o):

$$\varepsilon = \varepsilon_a + \varepsilon_o . \tag{4}$$

In accordance with [2-4], the following should be noted:

- The approximation error correlates with the MLP learning from training data, and the declaration error correlates with the generalization of these data.
- The approximation error is linearly proportional to the ratio of the sum of the number of input and output neurons to the number of synaptic connections.
- Assuming equal computational capabilities are present, the number of MLP synaptic connections is somewhat less than the number of synaptic connections in a two-layer perceptron.
- Each separate class corresponds to its output neuron.
- The number of training cases for each separate class shall exceed the number of input parameters by at least 10 times.
- The total number of all training cases shall be at least 10 times greater than the number of separate classes.
- Typical problems of evaluating the security parameters of information systems have the number of output parameters not exceeding the number of input parameters.

Based on the equations (2-4) and taking into account the above assumptions, we obtain the equations for determining the optimal MLP structure:

$$\begin{aligned} &\left(H^{\min} - 1\right) \times J_w^{\min 2} + \left(N_x + N_y\right) \times J_w^{\min} = \\ &\left(N_x + N_y\right) \times Round\left(0{,}2 \times \sqrt{N_x \times P}\right) \end{aligned} , \tag{4}$$

$$\begin{aligned} &\left(H^{\max} - 1\right) \times J_w^{\max 2} + \left(N_x + N_y\right) \times J_w^{\max} = , \\ &\left(N_x + N_y\right) \times Round\left(10 \times \sqrt{N_x \times P}\right) \end{aligned} \tag{5}$$

where $H^{\min}$ and $H^{\max}$ are the minimum and maximum numbers of hidden neuron layers in MLP; $J_w^{\min}$ and $J_w^{\max}$ are the minimum and maximum numbers of neurons in a single hidden layer, P is the number of training cases, N_x, N_y is the number of input and output neurons.

We recommend further refining of the number of hidden neurons in the range from $H^{\min}$ to $H^{\max}$ by performing simulation modelling using the approximation error minimization criterion. At the same time, further research is required to determine the optimal relationship between the number of hidden layers and the number of neurons in a single hidden layer.

In order to evaluate the informational content of the results obtained, let us consider a change in a search range for finding the optimal number of hidden neurons in a two-layer perceptron. The range change can be estimated by using the expression as follows:

$$\delta = \frac{H^{\max} - H^{\min}}{N^{\max} - N^{\min}} . \quad (6)$$

where $N^{\min}$, $N^{\max}$ are the minimum and maximum numbers of hidden neurons obtained from expressions (7, 8) given in [2].

$$N^{\min} = Round\left(0,2 \times \sqrt{N_x \times P}\right), \quad (7)$$

$$N^{\max} = Round\left(0 \times \sqrt{N_x \times P}\right). \quad (8)$$

Let us consider the task of detecting web-based computer threats. We take into account that for the majority of similar evaluation problems, the number of input parameters is approximately 100 ($N_x \approx 100$), and the number of training cases required for effective MLP training is 20,000 ($P \approx 10{,}000$). We assume that a two-layer perceptron with a single output ($N_Y = 1$) is used. By substituting these parameters into equations (4-8), we obtain $\delta \approx 0.7$. Therefore, the search range for the optimal number of hidden neurons has been narrowed by approximately one and a half.

Conclusions

The study shows that it is possible to increase efficiency of a multilayer perceptron intended for evaluation of technical system parameters by adapting perceptron structural parameters to the specific task conditions. The proposed adaptation ensures minimization of the relative error during neural network model generalization. A mathematical tool has been developed to calculate the most acceptable range of the number of neurons in hidden layers.

References

1. A. Korchenko. Neural network models, methods and tools for evaluation of security parameters of Internet-oriented information systems: monograph/A. Korchenko, I. Tereikovskyi, N. Karpinskyi, S. Tynymbaiev. Kyiv: Nash Format LLC. 2016. — 275 p.
2. O. G. Rudenko. Artificial neuron networks. Study guide./O. G. Rudenko, Ye. V. Bodianskyi. — Kharkiv: Kompania SMIT LLC, 2006. — 404 p.
3. I. Tereikovskyi. Neuron networks for computer data security means / I. Tereykovskiy. Kyiv: Polygraph-Consulting LLC. 2007. — 209 p.
4. I. A. Tereikovskyi. Two-layer perceptron structure optimization intended for detection of anomalous values of computer network operation parameters./I. A. Tereikovskyi//Collected volume of scientific and development papers "Managing the complex systems development". Kyiv National University of Construction and Architecture. — 2011. — Issue 5. — pp. 128-131.
5. Hu, Z., Tereikovskyi, I. A., Tereikovska, L. O., Pogorelov, V. V. I.J. Intelligent Systems and Applications, 2017, 10, 57-62.