

Electoral neural network technologies in the information-analytical system of the security operation center

Andrey Mikryukov

associate Professor, PhD,

*Institute of Digital Economics and Information Technology
of the Plekhanov Russian University of Economics,
117997, Russian Federation, Moscow, Stremyanny Lane, 36
Mikrakov.aa@rea.ru*

Mikhail Mazurov

*professor, doctor of Physical and Mathematical Sciences
Institute of Digital Economics and Information Technology
of the Plekhanov Russian University of Economics,
117997, Russian Federation, Moscow, Stremyanny Lane, 36
Mazurov37@mail.ru*

Abstract: The article describes an approach to improving methods for solving security event classification problems in information and analytical systems of security operation centers based on an ensemble of electoral neural networks, as well as using the adaptive reduction method of a neural network ensemble to improve the classification reliability. Electoral neurons differ from classic neurons in a more efficient way of processing input information. The approach is based on the application of procedures for adaptive reduction of the results of classifiers of the ensemble and the choice of the method of their aggregation. It is shown that the use of the considered approach provides an increase in the efficiency of solving the set task.

Keywords: collective classifier, neural network ensemble, electoral neuron, electoral neural network, information and analytical systems, security operation center.

1 Introduction

Nowadays, neural network approaches, in particular, the use of neural network ensembles, which are an example of collective problem solving, are widely used for solving problems of modeling various systems that are hard to formalize and weakly structured.

The quality of solving a specific task (data mining, forecasting, pattern recognition, classification, etc.) can be significantly improved using neural network ensembles, which assume the formation and training of a finite set of neural networks, the results of which are taken into account in the overall solution. At the same time, individual decisions are coordinated in such a way that the overall final decision is the best.

As you know, one of the fundamental problems of improving the functioning of an ensemble of neural networks in terms of increasing their accuracy and reliability is the generation of the diversity of the ensemble (the difference between individual models) [1]. Aggregation of similar models in an ensemble cannot lead to a significant improvement in the quality of the solution of the problem.

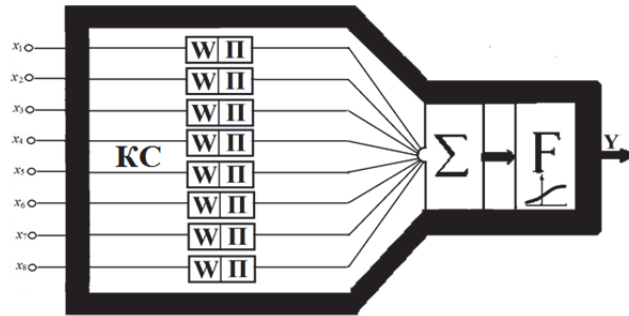
To resolve this contradiction, it was proposed to use approaches based on a collective (ensemble) of neural networks, for the construction of which a new class of artificial neurons, the so-called selective neurons, differing from classical neurons using a more efficient way of processing input information close to a biological neuron are used [2]. The considered

approaches can be used as part of the decision support subsystems of information and analytical systems (IAS) of security operational centers. The next section presents the features of the construction and operation of electoral neural networks.

2 Features of the construction and operation of electoral neural networks

Electoral networks are built on the basis of selective neurons. Teaching neural networks is not carried out by changing the weights of synaptic connections, but by changing the quantity and quality of inhibitory and exciting dendrites (signal transmitters) on the basis of which selective neurons are built [2].

A mathematically artificial neuron usually has a non-linear function of a linear combination of signals with a limited number of inputs. Exit the neuron from the next layer. The communication channels through which signals are received are determined by weights. The task of learning is to minimize the total error value of the output value. Figure 1 shows the McCulloch-Pitts neuron scheme, which is the basic element of the most classic neural networks of practical importance [3].



Legend:

- x_1, x_2, \dots, x_8 are input signals;
- KC is communication channels,
- Σ is adder;
- W is the weight of synaptic connections;
- P is function of the product $x_i * w_i$;
- F is a nonlinear threshold function;
- Y is a set of output signals.

Figure 1 – McCulloch-Pitts neuron diagram

A mathematical model of an artificial neuron is presented below. The output value y can be represented by the formula (1).

$$y = F(S) = F\left(\sum_{i=1}^N w_i x_i - \theta\right) = F(\mathbf{w}, \mathbf{x} - \theta), \quad (1)$$

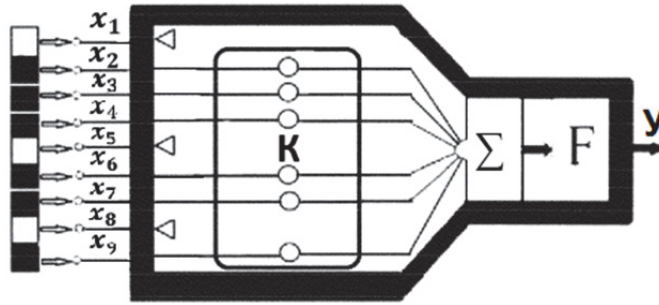
where $F(S)$ is the data transformation function after adder Σ , $\mathbf{x} = (x_1, \dots, x_N)$, $\mathbf{w} = (w_1, \dots, w_N)$, (\mathbf{w}, \mathbf{x}) is the scalar product of vectors \mathbf{w}, \mathbf{x} ; N is the number of inputs of the neuron; θ - threshold of arousal.

However, the classical types of neural networks have a significant drawback - the instability to retraining. If the number of weights in the network is sufficiently large, then the operation of summing their linear combination becomes extremely computationally laborious, and a retraining effect is likely to occur when the network recognizes images on the training set with extremely high accuracy, but at the same time shows a fairly large percentage of error on the test set input data.

Thus, the drawbacks of classic neural networks are a consequence of the use of weights of synaptic connections. In biological neurons there is no weight gradation of synaptic connections. They solve the main tasks of processing input signals by changing the number of dendrites receiving an input signal at the input of a biological neuron and, accordingly, changing the configuration of their clustering [4,5].

The analysis showed that, by analogy with biological mechanisms, it is advisable to use a neural network that includes so-called selective neurons, with controlled clustering of input channels (synapses) in quantity, as well as clustering of input channels in quality (exciting or inhibiting). As an effective classifier of images, it was proposed to use a neural network based on selective neurons. The selective neuron does not have weights of synaptic connections and is close in properties to its biological analog (Figure 2).

Cluster formation includes blocking non-informative communication channels that do not conduct excitatory or inhibitory signals [4-7]. Thus, after learning, each neuron has a cluster of input communication channels with an individual transfer characteristic, where some of the inputs are blocked for signals (the blocked inputs are marked with triangles), which are certainly different from the training code combination.

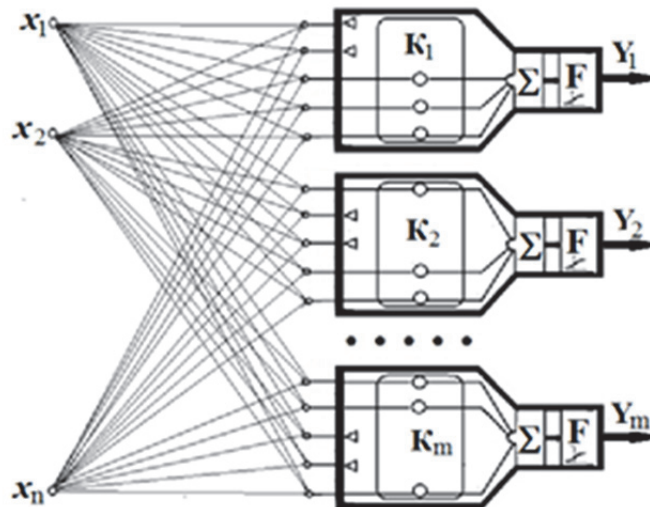


Legend:
 x_1, x_2, \dots, x_9 are input signals in the form of a binary code;
 K is a cluster of input channels of communication, formed in accordance with the binary code at the entrance;
 Σ is adder;
 F is a nonlinear threshold function;
 Y is output signal.

Figure 2 – Diagram of the electoral neuron

A distinctive feature of the electoral neuron is the presence of a cluster formed from a combination of internal communication channels. The cluster is formed in accordance with the specific properties of the input signal; the cluster is connected to an adder, after which a nonlinear threshold signal transformation is performed, which represents the output signal.

Taking into account the drawbacks of the known single-layer perceptron, a selective single-layer perceptron was proposed, based on electoral neurons [6]. An effective method of obtaining selectivity in the system of n neurons is implemented practically. The block diagram of the electoral perceptron is shown in Figure 3.



Legend:
 K_i is formed clusters of communication channels;
 Σ is adder;
 F is a non-linear function;
 Y_i is output signal.

Figure 3 – Structure of the selective single-layer perceptron based on selective neurons

In a single-layer electoral perceptron, clusters of specialized communication channels, dendrites, are configured in each electoral neuron and tuned to the corresponding characteristic vectors of the input signals. Triangles denote blocked communication channels from among the input, which are not essential characteristic vectors for the object at the input of the perceptron.

Mathematically, a selective neuron can be described as follows. The output value of y can be represented by the formula

$$y = F\left(\sum_{i \in K} x_i - \theta\right), \quad (2)$$

where K is a cluster of permissible values that are part of a cluster of an electoral neuron. Unlike formula (1) in formula (2) there are no weight coefficients w_i .

Possible characteristic code combinations of objects at the input of a neuron in the form of vectors are represented by expressions (3).

$$\mathbf{x}_1 = (x_{11}, \dots, x_{1n}); \dots; \mathbf{x}_m = (x_{m1}, \dots, x_{mn}), \quad (3)$$

where n is the number of elements of the code combination; m is the number of objects.

All possible code combinations of input objects form a matrix A .

$$\mathbf{A} = \begin{pmatrix} x_{11}, x_{12}, \dots, x_{1n} \\ \dots & \dots & \dots \\ x_{m1}, x_{m2}, \dots, x_{mn} \end{pmatrix}$$

Let the i -th neuron contain a cluster of connections characterized by a code combination $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$. Upon receipt at the input of the i -th neuron of the code combination of the input object we get

$$S_{ij} = \sum_{k=1}^n x_{ik} x_{kj} = (\mathbf{x}_i, \mathbf{x}_j)$$

The values of the sums S_{ij} are equal to the elements of the matrix B (4).

$$B = A * A^T, \quad (4)$$

where A^T is the transposed matrix A . In total we get $m \times m$ sums S_{ij} . The largest will be the sum (5).

$$S_{ii} = \sum_{j=1}^n x_{ij} x_{ji} = (\mathbf{x}_i, \mathbf{x}_i) = N_i = N, \quad (5)$$

where N_i is the number of units in the code combination $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$. The ratio $S_{ij} < N$ used to recognize input features.

Select the threshold value $U_n = \alpha N_i$, where $0 < \alpha < 1$. Then the output is

$$y = \begin{cases} 1 & S_{ik} \geq U_n \\ 0 & S_{ik} < U_n \end{cases}$$

This property characterizes the selective properties of the considered perceptron in the recognition of various input objects.

The graphical interpretation of the selective properties of a single-layer perceptron is represented by the matrix values $B = A \cdot A^T$ as a graph in three-dimensional space. From a physical point of view, the values $S_{ij} = f(i, j)$ of the expression are the matrix values $B = A \cdot A^T$ equal to the sum of the products of the input signal x_j after it passes through the cluster of communication prechannels, combining the working (unlocked) inputs of the i -th neuron, to the corresponding values of the threshold nonlinear function F at the perceptron output.

To build the graph in three-dimensional space, the tools of the Matlab-7 application software package are used. The axes O_x and O_y are the values of i and j . As an example, selective recognition was implemented for 20 works of art in the field of 150×100 . A graphical interpretation of the selective properties of a single-layer perceptron based on selective neurons is illustrated in Figure 4, which presents the results of processing (filtering) input signals (objects).

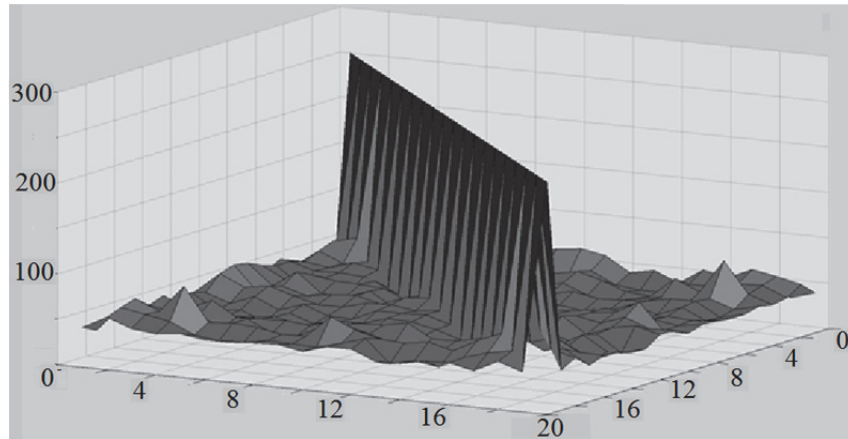


Figure 4 – Results of processing (filtering) input signals (objects) selective single layer perceptron

The results of processing the input signals shown in Figure 4 confirm the selective properties of a single-layer perceptron based on selective neurons, which are represented by a rather pronounced diagonal part.

Due to the peculiarities of its construction and operation, the selective neural network provides:

- selective recognition of input signals without using the weighting of their synaptic connections;
- the ability to encode an input signal of a certain type by the channel number or the number of the recording neuron;
- compression of input information, due to the preservation of only that information about objects that falls into a given channel or recording neuron;
- increase the speed of operation;
- improving the reliability of recognition of objects with a large number of them;
- a significantly greater degree of adequacy of the selective neuron to a biological single-layer perceptron.

Taking into account the considered features of electoral neurons, proposals were developed for improving methods for classifying security events in information and analytical systems of situational information security operational control centers based on an ensemble of electoral neural networks.

3 Features of the composition of artificial neural networks in ensembles

A promising direction for the improvement of artificial neural networks (ANN) is the union (composition) of many individual ANNs in an ensemble [8]. In this case, the errors of individual classification algorithms are mutually compensated. Ensemble organization is considered in a number of works [9-14]. In [12], the efficiency of using an ensemble organization for image recognition was experimentally proved.

When constructing an ensemble of neural networks, a finite set of previously trained neural networks is simultaneously used, the output signals of which are combined into a joint estimate, superior in quality to the results obtained using the local networks included in the ensembles.

The ensemble $H(\bar{x})$ of models $h_i(\bar{x})$ ($i=1,2,\dots,N$) is a composition of algorithmic operators $h_i: R^d \rightarrow R$ and the corrective operation $F: R^N \rightarrow R$, in which the set of estimates $h_1(\bar{x}), h_2(\bar{x}), \dots, h_N(\bar{x})$ corresponds to the final grade $H(\bar{x})$ [11]:

$$H(\bar{x}) = F(h_1(\bar{x}), h_2(\bar{x}), \dots, h_N(\bar{x})) \quad (6)$$

As you know, the fundamental task in the construction of ensembles is the generation of the diversity of the ensemble (or differences in individual models) [15].

Obviously, the aggregation of similar models in an ensemble cannot lead to a significant improvement in the quality of the solution of the problem. The ensemble of models can be better than the individual models included in the ensemble for the following reasons [16]:

- The ensemble reduces the rms error. The use of an ensemble of models averages the error of each individual model and reduces the influence of instabilities and randomness in the formation of hypotheses. Solving the problems of classification and regression is a search for hypotheses about the properties of the system or the next state of the system. If you use a sufficiently large number of models, trained in approximately the same set of examples, then you can reduce the instability and randomness of the result obtained by combining the results. Averaging over the set of models built on the basis of independent training sets, always reduces the expected value of the root-mean-square error.

- Ensembles of models trained on different subsets of the source data have a greater chance of finding the global optimum, since they are looking for it from different starting points.

- The combined hypothesis may not be in the set of possible hypotheses for basic classifiers, that is, when building a combined hypothesis, the set of possible hypotheses is expanded.

Several approaches are applied to the construction of model ensembles. Most often, an ensemble consists of basic models of the same type, which are trained in different sets of training samples (Figure 5).

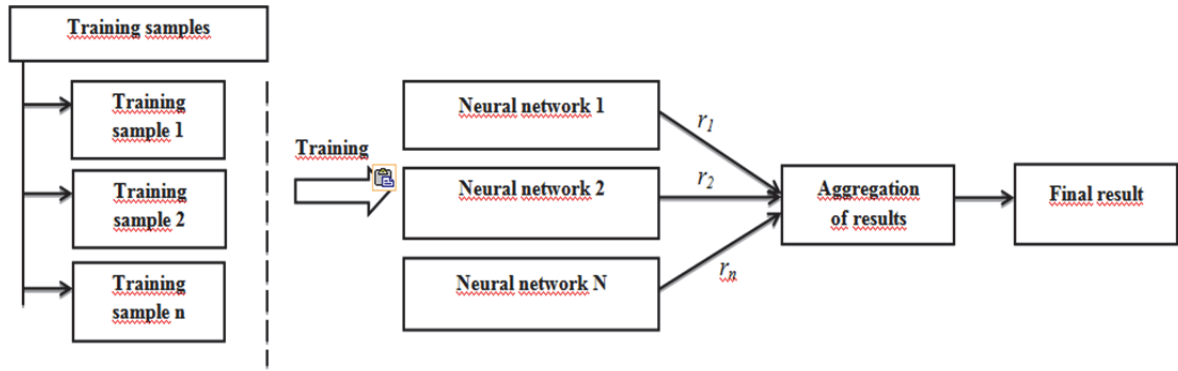


Figure 5 – The architecture of the ensemble of neural networks

For the formation of the output value of the ensemble under certain states of the outputs of the models, the most common are the following algorithms [17]:

1. *Voting*. Used in classification tasks. The class chosen by a simple majority of ensemble models is chosen.
2. *Weighted voting*. It differs from simple voting by assigning weights (points) for the results of different models. The points take into account the accuracy of the work of different classifiers.
3. *Averaging (weighted or unweighted)*. It is used in solving with the help of the ensemble for regression, when the outputs of the models will be numeric. The output of the entire ensemble can be defined as a simple average value of the outputs of all models. If weighted averaging is performed, the outputs of the models are multiplied by the corresponding weights.

The simplest example of a vote is simple voting:

$$Y(x) = F(y_1(x), y_2(x), \dots, y_k(x)) \frac{1}{m} \sum_{i=1}^m y_i(x) \quad (7)$$

where $x = (x_1, x_2, \dots, x_k)$ is the input vector, k is the number of inputs of the neural network, $Y = (Y_1, Y_2, \dots, Y_k)$ is the vector of values of the output signal of the neural network ensemble, F is the function to get the resulting solutions.

Simple voting represents a special case of weighted voting:

$$Y(x) = F(y_1(x), y_2(x), \dots, y_k(x)) = \frac{1}{m} \sum_{i=1}^m a_i y_i(x), \quad \sum_{i=1}^m a_i = 1, \quad a_i \geq 0, \quad (8)$$

where a_i is the weighting factor of the i -th model.

Weighted voting is a special case of a mixture of experts:

$$Y(x) = F(y_1(x), y_2(x), \dots, y_k(x)) = \frac{1}{m} \sum_{i=1}^m a_i(x) y_i(x), \quad \sum_{i=1}^m a_i(x) = 1, \quad \forall x \in X, \quad (9)$$

In the general case, the solution of the classification problem on the basis of the ANN team consists, firstly, in the formation and training of a finite set of ANN involved in the solution, and secondly, in determining the method of coordinating individual decisions so that the final solution is the best.

At the next stage, those ANNs are selected from the common pool, with the help of which the final decision will be formed. The final solution is a certain function, the input parameters of which are the particular solutions of the ANN included in the ensemble:

$$R = f(r_1, r_2, \dots, r_n), \quad (10)$$

where R is the total solution, r_i is the individual solution of the i -th ANN, n is the total number of ANN in the ensemble. The function f determines the method of generalization of individual decisions.

The solution to the problem of classification consists in choosing the number of one of the classes $A_j, j = 1, 2, \dots, J$, or choosing the empty set in the case of non-classification. Each partial solution r_i may take the value of the image or be an empty set if the pattern does not belong to the field of competence of the private classifier. The area of competence is understood as a subset of objects of a feature space within which the scope of a private classifier with a given subset of recognizable images is defined. Synthesis of the function f is the central task of using neural network ensembles.

Each of the solutions r_i can be assigned a weight, as well as a competence area. The collective decision is determined by the set of individual decisions r_i , which belong to the field of competence $K(r_i)$.

Thus, the decision of the team is determined by a set of individual decisions corresponding to the areas of their competence. This approach contains the ANN, whose solutions correspond to the field of competence. At the next stage, the ensemble output value is calculated by one of the above mentioned algorithms.

Currently, the most developed methods for building ensembles of neural networks are: equal or unequal voting for classification problems and simple or weighted averaging for regression problems [15].

Analysis of existing approaches has shown that they do not always provide the necessary quality of making the final decision (accuracy and validity).

The most significant of their disadvantages include:

- the dependence of the final result on the reliability of determining the coefficients of competence, which can lead to an incorrect result;
 - often the training sample contains noise emissions, leading to an increase in the probability of erroneous decisions by private classifiers, since an attempt to tune the noise to a training algorithm impairs the approximating capabilities of the network;
 - the need to use a large number of examples of training samples for the successful implementation of the algorithm;
 - The architecture of a neural network ensemble is often redundant, which does not contribute to an increase in the accuracy of solving classification problems and leads to a significant increase in the required computing resource.
- The effect of these disadvantages can be significantly reduced by using a neural network ensemble based on selective neural networks.

4 An integrated approach to solving the problem of classifying security events based on neural network ensembles

It is known that neural network models are used to solve classification problems in the decision support subsystems of information analysis systems (IAS). They have a number of features and advantages: they are adaptive self-learning systems, their use allows solving problems that are difficult or impossible to solve by traditional methods due to the absence of formalized mathematical descriptions of the processes of the object of study. Neural network models have an associative memory and in the process of work they accumulate and summarize information, from which their effectiveness increases over time. Their use is based on training the neural network to extract information from experimental data, which ensures the objectivity of the results and increases their reliability and reliability [18,19].

One of the drawbacks of the existing classification models used in the information-analytical system of the security operational center is the need to improve accuracy and reliability when solving problems of classification of security events due to their significant increase and diversity.

In order to improve the quality of solving the problem of classification of security events, the use of an ensemble based on electoral neural networks has been proposed. This approach allows, using the advantages of selective neurosites, to improve the accuracy and reliability of recognition of the input signal, to reduce the computational costs of implementing the functioning of the neural network ensemble. Improving the accuracy and reliability of solving classification problems is provided by applying the method of adaptive reduction of a neural network ensemble, as well as choosing and justifying the method of aggregation of results (composition of outputs of private classifiers).

The approach includes a set of steps:

1. Formation of the initial set (pool) of neural networks included in the ensemble (determination of the number of hidden neuron layers, activation functions, size of the training sample).
2. Reduction of the neural network ensemble (selection of the best classifiers based on the calculation of the classification reliability coefficients, comparison of the obtained values of the coefficients with the specified threshold values and evaluating the convergence of the results obtained by private classifiers).
3. Selection and justification of the method of aggregation of results (voting method).

The task of classification is to determine by its quantitative characteristics of an unknown object that it belongs to a particular image.

At the first stage, neural networks that meet the specified requirements are selected (generated). The architecture and parameters of the artificial neural network - the classifier, the size of the training sample are determined. In this case, neural networks are formed in the form of simpler structures, in contrast to the traditional approach to the synthesis of a classifier using a single neural network. Simple networks are fairly easy to learn and less prone to retraining.

At the second stage, the selection of the best classifiers is performed. To assess the competence of the classifier, a special algorithm (referee) is used. Under the competence of the classifier in this area of the space of representation of objects of classification is understood its accuracy probability of correct classification of objects whose description belongs to this area.

To formalize the method of aggregation of results (voting scheme), the coefficient $\mu_{ij} \leq 1$ of classification reliability by a private classifier is used. The coefficient μ_{ij} is the fraction of objects with a given image value j that fall within the competence area of the i -th classifier

$$\mu_{ij} = \frac{F_i(j)}{F(j)}, \quad (11)$$

where $F(j)$ is the cumulative frequency of solutions in the source database, $F_i(j)$ is the cumulative frequency of solutions of the image j for the i -th private classifier in its own area of competence. The voting function q_j of the j -th class is represented by the expression:

$$q_j = \sum_i \mu_{ij}, \quad j = 1, 2, \dots, J. \quad (12)$$

Если паттерн X не принадлежит области компетенции частного i -го классификатора, то значение $\mu_{ij}=0$. В этом случае, результаты работы частного классификатора в дальнейшем не учитываются, т.е. выполняется процедура редукции нейросетевого ансамбля.

Суммирование производится по всем оставшимся классификаторам. Решение о принадлежности паттерна X к одному из классов A_j принимается в соответствии с правилом:

If the pattern X does not belong to the area of competence of the private i -th classifier, then the value $\mu_{ij}=0$. In this case, the results of the work of the private classifier are not subsequently taken into account, i.e. The procedure for reducing the neural network ensemble is performed.

Summation is performed for all remaining classifiers. The decision on the belonging of the pattern X to one of the classes A_j is made in accordance with the rule:

$$\text{If } q_{j^*} = \max_j q_j, \text{ then } X \in A_{j^*}. \quad (13)$$

Such an approach is not always justified in conditions of noisy initial data, in connection with which the refusal of classification can be recorded. The choice of a strategy for combining the decisions of private classifiers usually does not require large computational resources, but at the same time provides a higher quality collective decision. In this case, one of the strategies can be used: selection and merging [20]. In the first case, each subspace of solutions corresponds to a separate classifier, in the second case, private classifiers are used throughout the solution space.

Of the technologies that ensure the effective projection of solutions of private classifiers to the target space, the most acceptable are [9,20]:

- the solution template method (the simplest method);
- weighted averaging;
- the method of multi-tiered generalization (uses a two-step procedure for forming decisions of classifiers with a non-linear combination of individual solutions), which has various modifications.

The choice of the preferred aggregation method is based on the rule of minimum classification error.

On the basis of the developed model of the selective neural network ensemble, representing a collection of neural networks consisting of selective single-layer perceptrons, the modeling of the classification of network attacks on a corporation's information system was carried out. Of the technologies that ensure the effective projection of the solutions of private classifiers onto the target space, the method of multi-tiered generalization, having various modifications, was applied. A computational experiment to test the proposed approach was carried out using test data from the repository [21].

The following network traffic parameters were used as source data:

1. Network recipient service (http, ssh, rdp, https);
2. The duration of the session in seconds;
3. The number of packets in the session;
4. The number of requests per session;
5. The number of responses in the session;
6. The number of requests with suspicious sequences of characters;
7. The number of authorization attempts per minute;
8. The number of erroneous packets in a session;
9. Connection status (active, established, closed);
10. The identity of the ports of the sender and receiver (identical / different).

The result of the work of a trained neural network ensemble is the classification of network traffic into 2 classes (legitimate traffic or suspicious traffic).

A computational experiment showed an increase in the reliability of solving the classification problem on the basis of neural network ensembles by an average of 8-12%, and the quality of diversity (by diversity is understood as the degree of uncorrelated errors of members of the neural network ensemble, the significant effect of which is confirmed, including experimentally) of models of selective neural networks calculated as the difference between the quadratic error of the ensemble and the average error of the individual models of electoral neural networks improved by 10-14%. At the same time, the volume (complexity) of calculations is significantly reduced when training the classifier electoral neural networks.

A promising direction is the development of collective methods for classifying security events, taking into account their considerable diversity and correlation. The most interesting results can be obtained when applying in the architecture of neural network ensembles of neural networks based on new neural elements using memristor technology.

5 Conclusion

Studies have shown that the use of a new class of neural networks based on selective neurons to build neural network ensembles in the decision support subsystems of information and analytical systems can significantly improve the quality (accuracy and reliability) of decision making by an ensemble of neural networks. The key difference between a selective neuron and a classical McCulloch-Pitts neuron is that it does not have weights of synaptic connections and is close in its properties to its biological analogue.

The functioning of electoral neural networks is based on controlled clustering of input channels, which is individual for each neuron of the network, which, in turn, reduces the computational resources for its training and functioning. Taking into account the peculiarities of the electoral neurons considered, proposals were developed for improving the classification methods based on an ensemble of electoral neural networks.

An approach is proposed to solve the problems of classification of security events in the information-analytical system of the situational information security operational center based on an ensemble of electoral neural networks, which allows to obtain more accurate and reliable results.

An improved integrated approach has been developed that provides more efficient classification results. The approach is based on the application of the procedure for adaptive reduction of the results of private classifiers and the procedure for choosing the method of aggregation of the results of private classifiers.

The results of the computational experiment showed an increase in the reliability of the solution to the classification problem (reduction of the classification error) based on neural network ensembles by an average of 8-12%, the quality of the diversity of models of electoral neural networks improved by 10-14%. At the same time, the volume (complexity) of calculations is significantly reduced when training the classifier electoral neural networks.

The main directions of development of approaches to improving the efficiency of neural network classification systems include the following: implementation of neural network classifier architectures based on hybrid models of classical and selective neural networks using multi-agent and cognitive technologies.

References

- [1] Kuncheva L.I. Combining Pattern Classifiers: Methods and algorithms. John Wiley & Sons, Hoboken, NJ, 2004.
- [2] Mazurov M. Ye. Pulsed neuron, close to real. The patent for the invention number 2598298. 08/30/2016.
- [3] Aleksandrov Yu.I., Anokhin K.V., Sokolov E.N., Grechenko T.N. and others. Neuron. Signal processing. Plastic. Modeling. Fundamental guide. Publishing house of Tyumen State University. 2008. 548 s.
- [4] Mazurov, M.E. Neuron, modeling the properties of a real neuron. The patent for the invention № 2597495. 08/22/2016.
- [5] Mazurov M.E. Election neural networks for the recognition of complex objects. Q: Mathematical biology and bioinformatics: proceedings of the VI International Conference. M.: MAX Press. 2016
- [6] Mazurov, M.E. One-layer perceptron based on selective neurons. The patent for the invention № 2597497. 08/22/2016.
- [7] Mazurov, M.E. One-layer perceptron, which simulates the properties of a real perceptron. The patent for the invention № 2597496. 08/22/2016.
- [8] Bishop C.M. Neural Networks for Pattern Recognition. - Oxford University Press, 1995. - 496 p.
- [9] Zhou Z.-H. Ensemble Methods: Foundations and algorithms. Chapman & Hall / Crc Machine Learning & Pattern Recognition. 2012. 236 p.
- [10] Kuncheva L.I. Combining Pattern Classifiers: Methods and algorithms. John Wiley & Sons, Hoboken, NJ, 2004.
- [11] Terekhov S. A. The genial committees of smart machines // Scientific session MEPhI-2007. IX All-Russian Scientific and Technical Conference "Neuroinformatics-2007": Lectures on neuroinformatics. Part 2. - M.: MEPI, 2007. - P. 11-42.
- [12] Vorontsov K. V. Lectures on algorithmic compositions. [Electronic resource]. URL: <http://www.ccas.ru/voron/download/Composition.pdf> (appeal date: 1.07.2019).
- [13] Goncharov M. Ensembles models. [Electronic resource]. URL: http://www.business_data_analytics.ru/download/ModelEnsembles.pdf (appeal date: 1.07.2019).
- [14] Borovikov V.P. Neural networks. Statistica Neural Networks. Methodology and technology of modern data analysis. 2nd ed., Pererab. and add. - M.: Hotline - Telecom, 2008. - 392 p.
- [15] Yu.V. Voevodin, L.G. Komartsova The use of a genetic algorithm to optimize the parameters of the neural network in the tasks of classification // Informatics: problems, methodology, technology. - M.: Publishing House of Moscow State Technical University named after Bauman, 2005. - From 42-46.
- [16] Plumton C.O., Kuncheva L.I. Choosing parameters for Random Subspace Assignments for fMRI classification, Proceedings of Multiple Classifier Systems (MCS 10), Cairo, Edgip, LNCS 5997, 2010, pp.54-63.
- [17] Paklin N. B., Oreshkov V. I. Business analyst: from data to knowledge. - SPb.: Peter, 2013. - 704 p.
- [18] Bova V.V., Dukkart A.N. The use of artificial neural networks for the collective solution of intellectual problems. Problems of knowledge representation in integrated support systems for management decisions. Izvestia SFU. Technical science. - 2010. - № 7 (108). - P.131-138.
- [19] A.A. Morozov, V.P. Klimenko, A.L. Lyakhov, S.P. Aleshin The State and Prospects of Neural Network Modeling of DSS in Complex Socio-Technical Systems // Mathematical Machines and Systems. - 2010. - №1.
- [20] Wolpert D.H. Stacked generalization. Neural Networks, 1992, no.5, pp.241-259.
- [21] Frank A., Asuncion A. UCL Machine Learning Repository. University of California, School of Information and Computer Science, Irvine, CA, 2010. [Electronic resource] URL: <http://archive.ics.uci.edu/ml> (circulation date 1.07.2019).