http://doi.org/10.35784/iapgos.6155 received: 16.04.2024 | revised: 10.09.2024 | accepted: 16.09.2024 | available online: 30.09.2024

MATHEMATICAL MODEL AND STRUCTURE OF A NEURAL NETWORK FOR DETECTION OF CYBER ATTACKS ON INFORMATION AND COMMUNICATION SYSTEMS

Lubov Zahoruiko[1,2], Tetiana Martianova[3], Mohammad Al-Hiari[4], Lyudmyla Polovenko[1], Maiia Kovalchuk[1], Svitlana Merinova[5], Volodymyr Shakhov[6], Bakhyt Yeraliyeva[7]

[1]Vinnytsia National Technical University, Vinnytsia, Ukraine, [2]Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine, [3]Vinnytsia Regional Youth Centre Kvadrat, Vinnytsia, Ukraine, [4]Jadara University, Irbid, Jordan, [5]Vinnytsia Institute of Trade and Economics of State University of Trade and Economics, Vinnytsia, Ukraine, [6]Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Vinnytsia, Ukraine, [7]M. Kh. Dulaty Taraz Regional University, Taraz, Kazakhstan

Abstract. *The paper discusses the principles of creating a mathematical model and system architecture by applying the method of artificial intelligence to detect cyberattacks on information and communication systems, where a neural network capable of learning and detecting cyberattacks is used. The proposed approach, based on the application of the developed mathematical model and architecture of artificial neural networks, as a detector of network attacks on information and communication systems, allows to increase the level of detection of network intrusions into computer systems, Web and Internet resources. An algorithm for processing network traffic parameters in real-time systems by structuring a neural network is proposed, which allows to optimize the redundancy of its multi-level structure at the level of inter-element connections.*

Keywords: neural networks, network traffic, network connection, cyber attack, information and communication system

MODEL MATEMATYCZNY I STRUKTURA SIECI NEURONOWEJ DO WYKRYWANIA CYBERATAKÓW NA SYSTEMY TELEINFORMATYCZNE I KOMUNIKACYJNE

Streszczenie. *W artykule omówiono zasady tworzenia modelu matematycznego i architektury systemu poprzez zastosowanie metody sztucznej inteligencji do wykrywania cyberataków na systemy teleinformatyczne, gdzie wykorzystywana jest sieć neuronowa zdolna do uczenia się i wykrywania cyberataków. Proponowane podejście, oparte na zastosowaniu opracowanego modelu matematycznego i architektury sztucznych sieci neuronowych, jako detektora ataków sieciowych na systemy teleinformatyczne, pozwala na zwiększenie poziomu wykrywania włamań sieciowych do systemów komputerowych, zasobów sieciowych i internetowych. Zaproponowano algorytm przetwarzania parametrów ruchu sieciowego w systemach czasu rzeczywistego poprzez strukturyzację sieci neuronowej, co pozwala na optymalizację redundancji jej wielopoziomowej struktury na poziomie połączeń międzyelementowych.*

Słowa kluczowe: sieci neuronowe, ruch sieciowy, połączenie sieciowe, cyberatak, system teleinformatyczny

Introduction

Ensuring information security is the primary task of every state. In the era of computerization and automation, the problem of protecting Web and Internet resources of information networks and information and communication systems from cyberattacks (CA) is the most problematic among the majority of tasks of ensuring cyberspace security. According to the materials of a number of sources, in recent years the amount of CA and the damage from them have been constantly increasing [3, 9, 11].

Two new aspects were carried out of these statistics:

- Internet resources and small and medium-sized information and communication systems became more often the targets of CAs;
- low power CAs become tools of influence.

Low-power CAs do not require powerful botnets, and conducting them is not followed by noticeable anomalies in communication channel bandwidth. The traffic resulting from such an attack may not differ at all from the normal operation of the victim resource, since botnet clients use a technique to imitate the behavior of legitimate users [16, 20, 21].

1. Classification of the architectures of artificial neural networks for cyberattacks detection

The processing of network traffic parameters for the purpose of CA detection onto information and communication systems, its localization and analysis of the dynamics of changes in automatic mode is an extremely important element of any attack detection system. Neural networks are widely and effectively used to solve the problems of detecting CAs. In this an intensively developing area of research. New tasks related to the development of new network structures are constantly appearing, the purpose of which is to increase the resistance of computer systems (CS) to attacks with full automation of the procedure for their detection. The article proposes and investigates a mathematical model and structure of a neural network capable of detecting cyberattacks.

Over the past few years, one of the most urgent problems in the field of information security is increasing the effectiveness of methods for detecting attacks on information resources of computer systems. At the same time, an important direction of efficiency improvement is the "intellectualization" of attack detection methods through the use of the theory of artificial neural networks (ANNs). The prospects of this approach are confirmed by some successful applications of ANNs in attack detection tools (Cisco products) and a large amount of relevant theoretical and practical studies. At the same time, a large number of false positives, a long term and instability of training, insufficient adaptation to many features of the modern CS, which is primarily related to methodological faults, significantly limit their practical value. Therefore, it became necessary to solve the problem caused by the prospects of use of the existing neural-like means of detecting attacks on the one hand, and by the imperfection of the methods of development and application of such means on the other [7, 10, 14].

In general, it can be noted that in the last decade there has been a rapid development of ANN technologies in the field of cyber security. Not only the possibilities of using ANNs in the detection of CAs are diverse, but their architecture is diverse as well. The classification analysis of the structures of such ANNs for detecting cyberattacks is presented in Fig. 1 [14–16].

It should be noted that the structure of any ANN used in many practical fields and, in particular, in cyber security attacks detection, can be considered as a directed graph with strongly connected edges, in which artificial neurons are vertices. According to the architecture of connections, ANNs can be grouped into two classes (Fig. 1): feedforward networks, in which the graphs do not have loops, and recurrent networks, or networks with feedback.

Feedforward networks are static because, for a given input, they produce a single set of output values that are independent of the previous state of the network. Recurrent networks are dynamic, as their neurons inputs are modified due to feedback, which significantly changes the state of the network [7, 14].

Networks using radial basis functions (RBF networks) are a special case of the two-layer feedforward network. Each element of the hidden layer uses a radial basis function, like Gaussian, as its activation function. The radial basis function (kernel function) is centered at a point defined by the weight vector associated with the neuron. Both the position and the width of the kernel function must be trained on selective samples. Of course, there are much fewer kernels than training examples. Each output element computes a linear combination of these radial basis functions [18, 22].

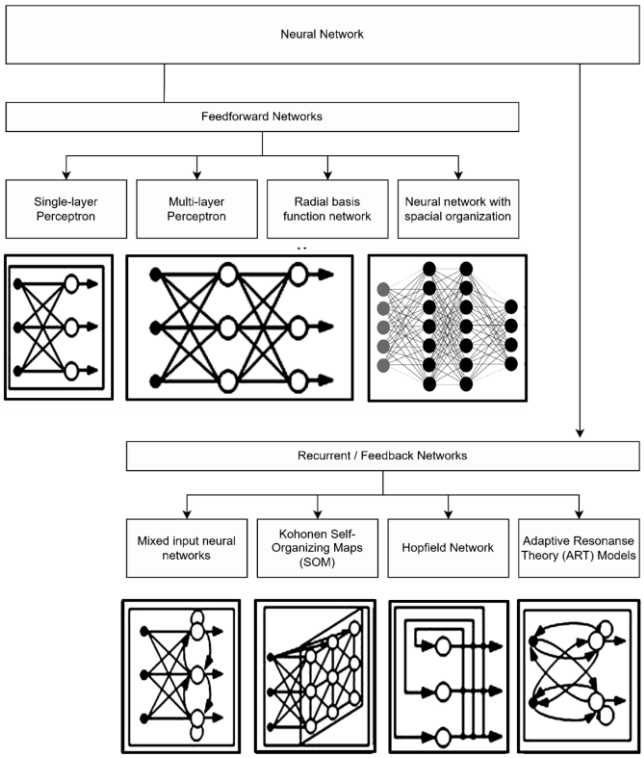


Fig. 1. Architectures classification of ANNs for cyberattack detection

Kohonen's Self-organizing maps (SOMs) have the favorable property of storing topology, which reproduces an important aspect of feature maps in the cerebral cortex of highly organized animals. In topology-preserving mapping, nearby input elements stimulate nearby output elements. In essence, the Kohonen SOM network is a two-dimensional array of elements, and each element is connected to all n input nodes.

This kind of a network is a special case of a competitive learning network in which a spatial environment is determined for each output element.

The stability-plasticity dilemma is an important feature of competitive learning. How to teach new types of CA (plasticity) and at the same time maintain stability so that existing knowledge is not erased or destroyed? Carpenter and Grossberg, who developed the adaptive resonance theory models (ART1, ART2, and ARTMAP), attempted to resolve this dilemma. The network has a sufficient number of output elements, but they are not used until the need arises. An element is considered allocated (not allocated) if it is used (not used). The learning algorithm corrects the existing category prototype only if the input vector is sufficiently similar to it [3, 8, 10].

Hopfield used the energy function as a tool to construct recurrent networks and to understand their dynamics. Hopfield's formalization clarified the principle of preserving information as dynamically stable attractors and popularized the use of recurrent networks for associative memory and for solving combinatorial optimization problems of the CA detection process.

Dynamic change of network states can be done in two ways: simultaneously and asynchronously. In the first case, all elements are modified simultaneously at each time step. In the second case a single element (network connection element) is selected and processed at a time. This item can be chosen randomly. The main feature of the energy function is that in the process of evolution of the network states according to the equation it decreases and reaches a local minimum (attractor) in which the function preserves constant energy [5, 11, 16].

There are three learning paradigms: "supervised", "unsupervised" (self-learning) and mixed. In the first case, the ANN has at its disposal the correct answers (network outputs) for each input element. The weights are adjusted so that the network produces responses that are close to known acceptable responses. The enhanced version of supervised learning assumes that only the critical evaluation of the validity of the ANN output is known, but not the acceptable output values themselves. Unsupervised learning does not require knowing the correct answers to each item of the learning sample. In this case, the internal structure of the data or the correlation between samples in the data system is revealed, which allows the distribution of samples by categories. In mixed learning, some of the weights are determined by a supervised learning, while the other is formed by self-learning. There are four main types of learning rules: error correction, Boltzmann machine, Hebb's rule, and competitive learning [1, 6].

There are many controversial issues in the design of feedforward ANNs. For example, how many layers are needed for this problem, how many nodes should be selected in each layer, how the network will respond to data that is not included in the training sample (what is the network's generalization ability), and what size of the training sample is necessary to achieve a "acceptable" network generalization ability [4, 12, 13].

Although multi-layer feedforward networks are widely used for feature classification and approximation, many parameters still need to be determined by trial and error. Existing theoretical results provide only weak guidelines for choosing these parameters in practical use [17, 19].

The evolution of ANNs caused a lot of enthusiasm and criticism. Some comparative studies, similar to those analyzed above, turned out to be optimistic while others were pessimistic. For many tasks, such as CA detection, no dominant approaches have yet been established. The choice of the best technology should be dictated by the nature of the task. It is necessary to try to understand the possibilities, prerequisites and the field of application of various approaches and to make the most of their additional advantages for the further development of intelligent systems for the CA detection. Such efforts can lead to a synergistic approach that combines ANNs with other technologies for a significant breakthrough in solving current problems. As M. Minsky, one of the founders of the theory of artificial intelligence, noted: "The time has come to build systems beyond individual components. Individual modules are important, but we also need an integration methodology. It is clear that the interaction and joint work of researchers in the field of ANNs and other disciplines will allow not only to avoid repetitions, but also (more importantly) will stimulate and provide new qualities for the development of certain areas".

The article covers a feedforward network with a hierarchical organization of links with the paradigm of a supervised learning method and the methodology of a spacial information integration.

2. Mathematical model and structure of a neural network for cyberattacks detection

In general, the methodology of forming a neural network can be presented in a formalized way, based on the following statements. Let there be a set of input streams (network connections). This brings up a problem. How to organize a real-time parallel processing so as to obtain a computing network that is strictly distributed in time and hierarchy?

The answer to this question can be given by the principle of building the following computing network.

Consider the mathematical model of the multilevel organization of the structure of a neural network [4], which is presented in Fig. 2. For this, it is necessary to obtain a functional description of the basis network. Let there be an information stream of network connections given in the form of sets $M_i(n)$, $i=\overline{1,n}$, where n is a dimension of an *i*-th set.

In general the basis structure of the neural network can be described with help of 6 variations of sets (1)–(6)

$$C_1(i',j')=\bigcup_{i=1}^{n}\bigcup_{j=1}^{n}\left(M\left]\frac{n-(i-1)}{j}\right[\right) \tag{1}$$

$$S_1(i',j')=\bigcup_{i=n+1}^{2n-1}\bigcup_{j=2}^{i}\left(M\left]\frac{i}{j}\right[\right)$$

$$S_1(i',j')=\bigcup_{i=n+1}^{2n-1}M(2n-i),\ \ j=1 \tag{2}$$

$$S_2(i',j')=\bigcup_{i=2n}^{3(n-1)}\bigcup_{j=4}^{i}\left(M\left]\frac{i}{j}\right[\right)$$

$$S_2(i',j')=\bigcup_{i=2n}^{3(n-1)}M(3n-2-i),\ \ j=2 \tag{3}$$

$$S_3(i',j')=M(1),\quad i=3n-2,\quad j=3$$

$$S_3(i',j')=M(1),\quad i=3n-2,\quad j=4 \tag{4}$$

$$S_4(i',j')=\bigcup_{i=3n-1}^{(3n-2)+(n-3)^2}\ \bigcup_{j=i+(7-3n)}^{i}\left(M\left]\frac{i}{j}\right[\right)$$

$$S_4(i',j')=M(1),$$

$$i=\overline{(3n-1),(3n-2)+(n-3)^2},\ j=\overline{5,4+(n-3)^2} \tag{5}$$

$$S_5(i',j')=\bigcup_{i=(3n-1)+(n-3)^2}^{n^2}\ \bigcup_{j=2i-(n-3)^2+9-6n}^{i}\left(M\left]\frac{i}{j}\right[\right)$$

$$S_5(i',j')=M(1),\quad i=\overline{(3n-1)+(n-3)^2,n^2},$$

$$j=\overline{(n-3)^2+6,(n-3)^2+2(3n-5)} \tag{6}$$

$$C_2(i',j')=$$

$$S_1(i',j')\bigcup S_2(i',j')\bigcup S_3(i',j')\bigcup S_4(i',j')\bigcup S_5(i',j') \tag{7}$$

$$C(i',j')=C_1(i',j')\bigcup C_2(i',j')$$

where j - hierarchical level of a set $M_i(n)$ in the basis set.

In expressions (1)÷(6) symbol] [means extraction of an integer part and rounding up to a bigger value.

The amount of sets of type (1) are defined by the size *n*. The basis set can be described simpler if sets (1)–(6) are supplied in pairs. For instance, a set $S_1(i',j')$ is described by the first expression of the formula (3), when $j=\overline{2,i}$ and the second expression when *j* = 1. So it turns out that the first elements of the set $S_1(i',j')$ reflect the state of the basis network for the first level (*j* = 1). The rest of the elements is formed from the first expression of the formula (6) and reflects the state of the basis network on its further levels. The amount of sets of type (5) is $n-1$. The amount of sets of type (6) is $n-2$. Clearly the amount of sets of type (6) is limited by a single sequence. The amount of sets of type (5) is not constant for different values of *n*. If $n<4$ this set is completely absent. If $n\geq 5$, the amount of these sets is $n-1$. The amount of sets (6) also varies based on *n* but in a more complex manner. For instance, for $n=3$ the amount of such sets is 2, and for $n=4$ the amount will be 6, and for $n=5$ there will be 8 of them.

In general sets (1)–(6) describe the structure of a neural network. Set $C_1(i',j')$ forms the main basis network, and set $C_2(i',j')$ forms the tail basis network. The multiple of sets $C_1(i',j')$ and $C_2(i',j')$ form the complete basis network which is represented by the set (7). Therefore the basis neural network is formed by the main and the tail networks which are integral components to describe the interaction of the parallel data streams.

An example of basis network sets formation (1)–(6) with 4x4 dimensions is displayed on the Fig. 2.

The primary idea of the studied ANN, as follows from the description (1)–(6) of the complete basis network, is as follows. At the first level, the initial transform (specified for a specific task) is applied in a parallel way in independent branches (channels) and thereby information data streams are formed for conversion at subsequent levels. At each subsequent level, the tail element of the network is formed – the last element of the sets (1)–(6) for each value *i* and the new information stream which is the input for the transforms on the next level of the network.

$C_1(i',j')\{$
1
21
321
4221

$S_1(i',j')\{$
33221
232221
14321221

$S_2(i',j')\{$
222221
1322221

$S_3(i',j')\{$
11222221

$S_5(i',j')\{$
1222221
122221
12221
1221
121
11

Fig. 2. Example of the basis network sets formation with dimensions 4x4

Analyzing the sets, we can conclude that the network formed by them implements the mechanism of spacio-temporal convergence of network elements. Moreover, the spectrum of such convergence is well illustrated in Fig. 2, the numerical sequences of which are nothing but a presentation of the convergence degree of numerical information at different levels of the neural network hierarchy.

Discrete transforms, which are traditionally used in practice for the detection of CAs, are mainly focused on the sequential description of the processing of the initial data.

Unlike known transformations [6], which are characterized by the use of a given basis function, for the transformations in this ANN it is proposed to use a computing network as a basis function. Such a network is given in the form of a collection of sets that describe its structure.

This transformation contains the joint development of the ideas of parallelism, hierarchy and mixing in the processing of data streams, which then propagate in the "horizontal" direction of the network, and to the elements of other hierarchical levels – in the "vertical" direction. It is proposed to use the idea of ensuring the maximum correlation of elements within the original data streams and between them by presenting the transformed information environment in the form of a parallel multi-level computing network. Ensuring the maximum

correlation of the elements of such a network is achieved by the organization of the network structure and multiple mixing of data streams at different levels of the hierarchy.

Marked as $M(n)=\bigcup_{i=1}^{n} M_i(n)$ the direct transform in the neural network has the following form

$$F(i', j') = M(n) * C(i', j')$$

Then the reverse transform in the neural network will look like this

$$M(n) = F(i', j') * C'(i', j') \tag{9}$$

where $*$ – is a functional operator of the network transform, $C'(i', j')$ – reverse complete basis network, which is easy to obtain by swapping the top and bottom indices in the expressions (1)–(6).

It is known that one of the natural approaches to expressing algorithms for parallel and pipeline processing is their description in the form of a spacio-temporal recursive program. Formulas for multilevel processing of type (1)–(6) are a recursive definition, which is confirmed by the organization of the multilevel processing routine in the form of a hierarchical network (Fig. 2).

Among the well-known expressions of parallel algorithms, recursive algorithms with spatio-temporal indices are most often used, which allow using the well-known methodology of mapping algorithms to matrix structures and, in particular, to systolic arrays. At the same time, an important condition necessary for the development of matrix processors is the presence features of full regularity and localizability in the algorithms.

Studies conducted for specific multilevel processing have confirmed the regularity and localizability inherent in these algorithms. For example, the algorithm of parallel multi-operand addition, at each level of which a new set of data is formed, in the process of subtracting the minimal element value from all the elements of the current set, that is, the formation of difference sections in the set by the minimal element of the set, can be implemented on a linear systolic array [4]. At the same time, at each level, the common part of all the elements of the set is accumulated, which makes it possible to form the sum of the elements of the set as a result. Thus, for multi-operand addition of type

$$S_n = \sum_{i=1}^{n} a_{i,} \tag{10}$$

for every *j*-th level for the set $M_i(n)$ the functional operator and the formation of a new set can be represented in the following way:

$$\{q_j\} = \left\{\min_i (a_{i,j-1})\right\} \tag{11}$$

$$M_{ij}(n) = \{a_{ij}\} = \{a_{i,j-1} - q_j\} \tag{12}$$

where a_{ij} – are the elements of the set $M_{ij}(n)$.

At every level the sum is calculated using the following:

$$S_j = S_{j-1} + q_j \sum_{i=1}^{n} f_{ij} \tag{13}$$

where f_{ij} is formed like this:

$$f_{ij} = \begin{cases} 1, & if \quad a_{ij} \geq 0 \\ 0, & if \quad a_{ij} < 0 \end{cases} \tag{14}$$

whereas $\sum_{i=1}^{n} f_{ij}$ in formula (13) defines the amount of non-negative elements in the $M_{ij}(n)$ set. As a result of projecting the multi-operand addition process for a single set $M_i(n)$ of operands on a matrix structure with usage of the graph theory [4] a corresponding resulting signal flow graph (SFG) was obtained in a form of a horizontal linear systolic array [4]. The Fig. 3a demonstrates the resulting SFG and the Fig. 3b shows the detailing of the systolic array nodes. The analysis of SFG processing can be taken out in parallel along the entire output bus from all the n nodes and the output data in a form of a number of the terminal processing level. It's worth mentioning that two entry points are required for the variable $\alpha_0^j = \infty$ which are used to determine the minimal element of the set $M_{ij}(n)$.

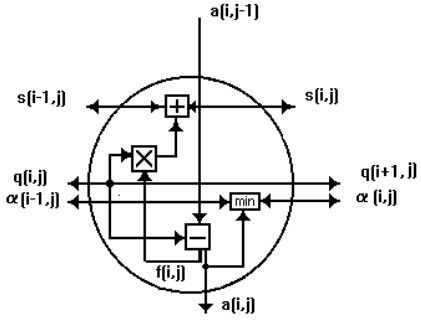
Thus, the fact that the proposed algorithm of multilevel processing belongs to the class of regular iterative algorithms (RIA) indicates the possibility of its implementation on systolic matrices and, in particular, on a linear systolic structure [16].

For the physical modeling of the processing environment, a Texas Instruments signal processor was used, which performs parallel calculations and is capable of multiplying and adding fractional numbers in one operation cycle. The use of a multi-processor computer for detecting CA in ANN, for example, a four-processor computer, allows to reduce the processing time at a level by an average of 1.2 times.

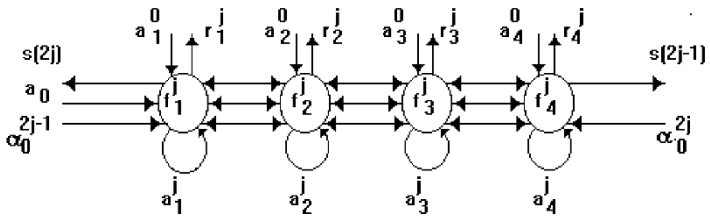
The training result is formed from the time-decorrelated activities of the tail PEs, each of which has the form of an even bump, along the slopes of which, at the given moment in time, the activity of the analyzed network connection standardly decreases. In this case, the coding costs are formed as the cost of describing the center of the tail resulting PE activity bump plus the cost of describing how its actual tail PE activity differs from the desired even activity bump.

Then the activity of the entire processing environment is formed as a spatio-temporal structure of the sequence activity of even bumps of individual tail PEs.

These smooth bumps of tail PE activity are decorrelated in time and form a straight ridge of the processor environment activity. The cost of describing the center of such an activity ridge and the cost of describing how it differs from the desired activity ridge is an objective measure of coding cost.



a) Resulting SFG



b) Detailed display of the systolic array nodes

Fig. 3. Systolic array SFG analysis

Using the coding of the center of activity ridges, we can come to the conclusion that the organization of a multi-stage processor environment in the form of an ANN at the algorithmic level is a convenient way of selecting parallel hierarchies of more effective codings.

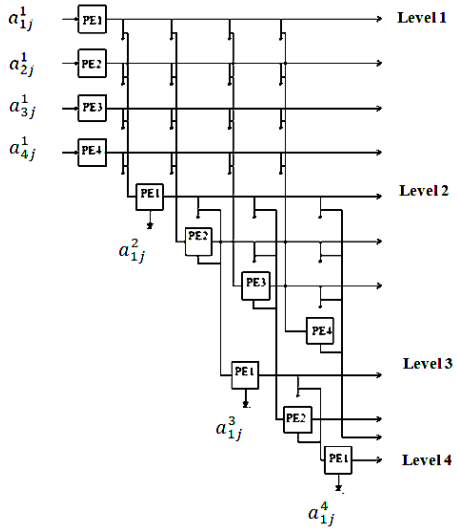


Fig. 4. Structural scheme of a multi-stage processor environment

A number of works have been published on this topic, primarily works by D. Ballard from the University of Rochester, E. Sonda from MIT, as well as R. Zemel and J.E. Hinton from the University of Toronto, which explore various options for solving this problem. The authors of this article, continuing the aforementioned research and summarizing the known results in this field [5] offer their version of solving this problem based on the multi-stage presentation in the form of a ridge of activities of the tail PE of individual elements.

It is obvious that natural neural networks have some redundancy of their assembly scheme and perform their own reassembly during the learning process. At each moment of time, the neural network strives to find an outline of active neurons, the experience of which allows more accurate presentation of objects at this stage. By computer standards, neurons work very slowly. But this slowliness is compensated both by the parallel work of a large number of neurons and by the rough hierarchical organization of the neural network.

This is very well consistent with the theory of multi-stage integration of S. Zeki. According to his ideas, integration is not a one-time event that happens by the convergence of signals in some higher zone. On the contrary, it is a parallel-hierarchical process, each of the levels of which makes a direct contribution to the evaluation of incoming visual information.

Mechanisms of pre-postsynaptic and pre-modeling convergence are characteristic of memory processes in natural neurons. They implement the mechanism of spatio-temporal correlation of neuron activity. Such a mechanism is one of the main features of the structural organization of the proposed network.

The inside phenomenon described by Fischbach is associated with a complex hierarchical process of spacio-temporal interaction of many coalitions of neurons, which is also an indirect confirmation of the rationality of building the spatial structure of a neural network.

According to the works of Rock and Palmer, one of the most informative properties of objects in the surrounding world is uniform connectedness, which explains the strong tendency of the eye-sight system to perceive any homogeneous connected area as a separate unit. The law of connectedness formulated by them is a candidate for the fundamental principles of grouping.

The formalization of the concept of connectedness in the spatio-temporal domain of the considered neural network, whose weight relationships are described in quantitative ratios of the network connections spatial connectivity, helps to approach the difficult problem of detecting attacks.

Although researchers have developed a number of neural network structures with powerful learning algorithms that have significant practical value, it is still unknown what representations and learning procedures the brain actually uses.

It is likely that the study of neural networks which have the natural computational manner and optimal learning for this network will help to study the methods that have been discovered by nature.

Thus, the analyzed approach [4] to the organization of ANNs could be more convenient and model neural networks in a natural manner. The structural organization of the ANN allows you to implement it in the form of an ordinary three-layer structure, in which input elements (parameters of the network connection) can be used as an input layer, and all other network levels with their deterministic and not random connections as a hidden layer. And for the output layer – the output elements of the ordinary three-layer network connected to the hidden layer by random connections [4]. With such a scheme of ANN implementation it becomes possible to include equally important temporal component in addition to its spacial components into its graphical representation. This allows to introduce the temporal correlation of network elements. Practically, this approach belongs to the group of methods that use correlations between the activity of the hidden element and the activity of the input element. At the same time, new possibilities for the organization of unsupervised learning appear, which seek a compromise solution between the two extreme approaches with purely distributed and purely localized representations.

3. Experimental research of the effectiveness of the proposed ANN for cyberattacks detection

To verify the effectiveness of the proposed mathematical model and ANN architecture, experimental studies were conducted in which the developed network was tested for the detection of cyber attacks.

The question of choosing a database for experimental research on cyberattacks does not have a simple solution because widely used databases contain many outdated types of attacks, and more modern databases have a specific structure that requires complex preprocessing and are used only by individual researchers, which prevents comparison of quality indicators of work results. To date, we can single out the two most common training databases with known attacks - DARPA and KDD Cup1999 Data.

The DARPA (Defense Advanced Research Project Agency) training database was formed as part of a study of the capabilities of various intrusion detection systems. This study used network traffic data and file system information to identify simulated intrusions performed by experts during recording network dumps. Information about DARPA attacks is stored in the form of a text descriptions, which indicates the time of the attack, its duration, the address of the victim, the name of the attack, the category of the attack, and other parameters.

Unlike the DARPA training data, the KDD Cup1999 Data database has about 5,000,000 records and contains processed information in the form of arrays of 42 key values.

The database contains 22 types of cyberattacks, which are divided into 4 main classes - denial of service (DoS), acquiring of unauthorized access rights by an unregistered user (R2L), unauthorized elevation of privileges (U2R) by a registered user, and port scanning (Probe). The number of database entries for each type of attack is shown in table 1. The number of records corresponding to the absence of an attack is 972,781.

Table 1. Characteristic of records of KDD Cup1999 Data database

Class of attack	Type of attack	No. of records
DoS	neptune	1072017
	smurf	2807886
	Pod	264
	teardrop	979
	land	21
	back	2203
U2R	buffer_overflow	33
	loadmodule	9
	perl	3
	rootkit	10
R2L	guess_passwd	53
	ftp_write	8
	imap	12
	phf	4
	multihop	7
	warezmaster	20
	warezclient	1020
	spy	2
Probe	portsweep	10413
	ipsweep	12481
	satan	15892
	nmap	2316

Since the main prerequisite to use the expert knowledge in ANNs is insufficient completeness of training data, the main attention was focused on recognizing U2R attacks, for which the number of records in KDD Cup1999 Data is the smallest.

Approbation of the developed ANN on the training database KDD Cup1999 Data showed a sufficiently high accuracy of recognition of all types of attacks.

For a comparative analysis of the detection of network attacks by the proposed ANN and other methods of intelligent analysis of network traffic for the purpose of detecting cyber attacks, the data presented in works [8, 20] were used.

The conducted experiments showed sufficiently high accuracy of attack detection and proved the correct choice of the developed mathematical model and ANN architecture for intelligent data analysis and their high efficiency. The use of various methods, the possibility of setting internal parameters and threshold values allow you to achieve the optimal ratio of system performance and the accuracy of recognition of attacks on Web and Internet resources.

Further research can be directed to the combinatorics of the application of methods of intelligent data distribution with other known methods and systems of intrusion detection. In order to solve this problem, it is proposed to refine the learning algorithm and the structure of the ANN.

Table 2 contains the results of detection of various types of attacks by various methods. As can be seen from the presented results, the proposed approach shows better results in cyber attacks detection. It should be noted that the percentage of false positives, when a legitimate connection is mistaken for an attack, is less than 10%. It is also worth mentioning that the table presents the average results for four classes of network attacks.

If we take separate types of attacks, then here it is possible to outline the attacks, the detection of which occurs with 98% probability (for example neptune, teardrop), and attacks, the detection of which does not occur (for example guess_passwd, spy).

Table 2. Results of network attacks detection

Attack detection methods	DoS, %	Probe, %	R2L, %	U2R, %
Neural network	98.0	92.8	36.7	30.8
Gaussian Classifier	82.4	90.2	9.6	22.8
K – NN	97.3	87.6	6.4	29.8
Nearest Neighbour Method	97.1	88.8	3.4	2.2
Leader algorithm	97.2	83.3	1.0	6.6
Hypersphere algorithm	97.2	84.8	1.0	8.3
Fuzzy Art Map	97.0	77.2	3.7	6.1
Decision Tree	97.0	80.8	4.6	1.8

4. Conclusions

The proposed approach, based on the application of the developed mathematical model and ANN architecture as a detector of network attacks on information and communication systems, allows to increase the level of detection of network intrusions on computer systems, Web and Internet resources. Detection of some types of attacks occurs with a 98% probability with a negligible level of false detections. In addition, this approach is not demanding on system resources and is able to detect some unknown types of attacks on information and communication systems (an ANN that is trained on one type of attack often shows good results in detecting other types of attacks, that is, on the data on which training was not conducted).

References

[1] Andrushchenko M. et al.: Hand Movement Disorders Tracking By Smartphone Based On Computer Vision Methods. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska – IAPGOS 14(2), 2024, 5–10 [https://doi.org/10.35784/iapgos.6126].

[2] Avrunin O. et al.: Improving the methods for visualization of middle ear pathologies based on telemedicine services in remote treatment. IEEE KhPI Week on Advanced Technology – KhPI Week 2020, 347–350 [https://doi:10.1109/KhPIWeek51551.2020.9250090].

[3] Bezobrazov S. et al.: Artificial intelligence for sport activitity recognition. 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications – IDAACS 2019, V. 2, 628–632.

[4] Bisikalo O. et al.: Parameterization of the Stochastic Model for Evaluating Variable Small Data in the Shannon Entropy Basis. Entropy 25, 2023, 184 [https://doi.org/10.3390/e25020184].

[5] Dhangar K., Kulhare D., Khan A. A.: Proposed Intrusion Detection System. International Journal of Computer Applications 65(23), 2013, 46–50.

[6] Emelyanova Yu. G. et al.: Neural network technology for detecting network attacks on information resources. Software systems: theory and applications 3(7), 2011, 3–15.

[7] Haykin S.: Neural Networks and Learning Machines. Pearson Education, 2009.

[8] Kolodchak O. M.: Modern methods of detecting anomalies in intrusion detection systems. Bulletin of the Lviv Polytechnic National University. Series "Computer Systems and Networks" 745, 2012, 98–104.

[9] Korobiichuk I. et al.: Cyberattack classificator verification. Advanced Solutions in Diagnostics and Fault Tolerant Control, Springer International Publishing, 2018, 402–411.

[10] Lee J. et al.: Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles. IEEE Access 7, 2019, 165607–165626 [https://doi.org/10.1109/ACCESS.2019.2953095].

[11] Likhouzova T. A., Nosenko K. M., Pivtorak O. I.: Review of attack detection systems in network traffic. Adaptive automatic control systems 1(24), 2014, 67–75.

[12] Meleshko Ye.: Method of collaborative filtration based on associative networks of users similarity. Advanced information systems 2(4), 2018, 55–59.

[13] Naseer S., Saleem Y., Khalid S.: Enhanced network anomaly detection based on deep neural networks. IEEE Access 6, 2018, 48231–48246.

[14] Pakhomova V. M., Konnov M. S.: Research of two approaches to detect network attacks using neural network technologies. Science and Transport Progress 3(87), 2020, 81–93.

[15] Shestak Ya. et al.: Minimization of Information Losses in Data Centers as one of the Priority Areas of Information Security Technologies. IEEE 9th International Conference on Problems of Infocommunications, Science and Technology – PIC S&T, 2022, 227–230.

[16] Timchenko L. I. et al.: Approach to parallel-hierarchical network learning for real-time image sequences recognition, Proc. Machine Vision Systems for Inspection and Metrology VII, Boston (Massachusetts USA), 1999.

[17] Timchenko L. et al.: Q-processors for real-time image processing. Proc. SPIE 11581, 2020, 115810F.

[18] Turlykozhayeva D. et al.: Routing Algorithm for Software Defined Network Based on Boxcovering Algorithm. 10th International Conference on Wireless Networks and Mobile Communications – WINCOM, 2023.

[19] Turlykozhayeva D. et al.: Routing metric and protocol for wireless mesh network based on information entropy theory. Eurasian Physical Technical Journal 46, 2008, 90–98.

[20] Ulichev O. S. et al.: Computer modeling of dissemination of informational influences in social networks with different strategies of information distributors. Proc. SPIE 11176, 2019, 111761T.

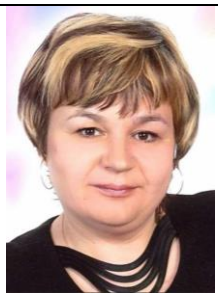
[21] Wu Y., Wei D., Feng J.: Network Attacks Detection Methods Based on Deep Learning Techniques: A Survey, Wiley, Open Access, 2020 [https://doi.org/10.1155/2020/8872923].

[22] Zh Z. Z. et al.: Cluster router based on eccentricity, Eurasian Physical Technical Journal 19(3(41)), 2022, 84–90.

Ph.D. Lubov Zahoruiko
e-mail: l.zahoruiko@donnu.edu.ua

Candidate of Technical Sciences, associate profesor of the department of applied mathematics and cyber security. Vinnytsia National Technical University, Vinnytsia, Ukraine, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine
Published more than 30 scientific works. Field of scientific interests: neural networks, image processing and recognition, cyber security and information protection.

https://orcid.org/0000-0002-6958-8696



Ph.D. Tetiana Martianova
e-mail: zagorujko2016@gmail.com

Candidate of Technical Sciences, associate profesor of the department of applied mathematics and cyber security. Vinnytsia Regional Youth Centre Kvadrat, Vinnytsia, Ukraine
Published more than 17 scientific works. Field of scientific interests: neural networks, image processing and recognition, cyber security and information protection.

https://orcid.org/0000-0002-8700-8050

Ph.D. Mohammad Al-Hiari
e-mail: hiari5355@yahoo.com

Ph.D. Jadara University/Irbid-Jordan, Faculty of Science and Information Technology, Department of Computer Networks and Cyber Security,
Field of scientific interests: neural networks, image processing and recognition, cyber security and information protection.

https://orcid.org/00900002-2770-5417

Ph.D. Lyudmyla Polovenko
e-mail: l.polovenko@donnu.edu.ua

Doctor of Philosophy, associate professor, Vinnytsia National Technical University, Vinnytsia, Ukraine.
Published 10 scientific works.
Field of scientific interests: neural networks, image processing and recognition, cyber security and information protection.

https://orcid.org/0000-0002-9909-825X

D.Sc. Maiia Kovalchuk
e-mail: maya.kovalchuk@gmail.com

Doctor of Pedagogical Sciences, associate professor, Professor of the Department of Higher Mathematics, Vinnytsia National Technical University, Vinnytsia, Ukraine.
Field of scientific interests: mathematical modeling, pedagogic, cyber security and information protection.

https://orcid.org/0000-0002-1895-1715

Ph.D. Svitlana Merinova
e-mail: s.merinova@vtei.edu.ua

Doctor of Philosophy, associate professor, Vinnitsa Institute of Trade and Economics of Kyiv National University of Trade and Economics. Graduated Vinnitsa State Polytechnic Institute in 1991, specialization "System engineer".
Field of scientific interests: neural networks, image processing and recognition, cyber security and information protection.

https://orcid.org/0000-0001-6563-5320

Ph.D. Volodymyr Shakhov
e-mail: shahovu2016@gmail.com

Doctor of Pedagogical Sciences, profesor at Department of Psychology at Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Ukraine.
Field of scientific interests: mathematical modeling, pedagogic, cyber security and information protection.

https://orcid.org/0000-0003-15 35-2802

Ph.D. Bakhyt Yeraliyeva
e-mail: yeraliyevabakhyt81@gmail.com

Senior lecturer of the Information Systems Department, Faculty of Information Technology, M. Kh. Dulaty Taraz Regional University, Taraz, Kazakhstan
Research interests: Fiber Optic Technologies, Information Systems, Internet of Things and Blockchain technologies. H-index in scientometric databases

https://orcid.org/0000-0002-8680-7694