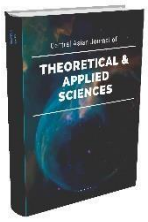
X1



CENTRAL ASIAN JOURNAL OF THEORETICAL AND APPLIED SCIENCES

Volume: 03 Issue: 04 | Apr 2022 ISSN: 2660-5317

Integration of Biometric and Electronic Signatures Using Neural Network Algorithms

Muminova Sunbula, Nomozov Mansurbek

Tashkent University of information technologies named after Muhammad al-Khwarazmi Tashkent, Uzbekistan

Sunbula.axmedova@gmail.com

Received 24th Feb 2022, Accepted 28th Mar 2022, Online 29th Apr 2022

Abstract: *This article provides arguments regarding the possibility of using modern methods of pattern recognition in problems of biometric authentication: fuzzy extractors, artificial multilayer neural networks, deep learning methods, as well as convolutional, evolutionary, small, wide, hybrid neural networks. The results of our own research in this area are presented. Two methods are proposed for integrating biometric and electronic signatures based on dynamic signature parameters, as well as face and keyboard handwriting parameters.*

Keywords: *artificial neural networks, biometrics, electronic signature, pattern recognition.*

I. INTRODUCTION

The trends of the modern information society are associated with the transition of states to the digital economy. This leads to the fact that entire segments of document flow (DO) are transferred to the digital environment: government services, banking services, e-procurement. Although documents are created using software, many are distributed on paper. This is due to the fact that the rate of widespread introduction and assimilation of modern technologies in organizations, as well as the development of legislation, lag behind the potential opportunities that these technologies provide. Therefore, in the foreseeable In the future, workflow will be mixed in many areas of activity, with electronic documents and transactions prevailing.

This article discusses the solution to the problem of integrating biometric and electronic signatures in a hybrid workflow environment [1]. A hybrid document can be in electronic or paper form and contain an autograph image protected with electronic signature (ES), as well as secret or open biometric images, so that you can quickly (within a reasonable time) check its integrity and authenticity, regardless of the form of presentation (paper or electronic). The key attribute of a hybrid document is electronic signature, during the formation of which the biometric image of the subject is used. At the same time, neural network algorithms for converting biometric features into the secret key of the digital signature are used.

To form an EP from biometric data, a hybrid biometrics-code converter is proposed, capable of converting a fuzzy vector, ambiguous biometric parameters of the user into a clear unambiguous key (password) code.

As biometric images, it is proposed to use the so-called biometric handwritten signature (autograph) or the face image and the parameters of the subject's keyboard handwriting when they enter a passphrase. The main focus is on the dynamic biometric images of a person that change over the course of life.

II. DYNAMIC BIOMETRIC IMAGES

There are two options for the implementation of the technology for generating ES secret keys based on biometric images:

1. The first option implies the use of a handwritten signature as a biometric image. In this case, the formation of an electronic signature from a biometric handwritten signature will not significantly affect the existing business processes in the organization, since a signature is a common way of verifying the authenticity of paper documents.

2. The second proposed method for generating an electronic signature is to use data from standard equipment to generate a secret key: a keyboard and a webcam.

In this case, the parameters of the subject's face and keyboard handwriting are used. On the basis of this method, an "extended" security technology for electronic document implementations is also proposed, which is as follows. After the formation of a hybrid document, its owner can restrict the access of other persons to its arbitrary parts, as well as prohibit certain actions (printing, editing, etc.). In this case, the content of each of these parts of the document will be encrypted in the public key of the subject to whom access is granted. If more than one subject has access to one of the parts of the document, several copies of this part are created, each of which is encrypted with the corresponding public key. When a user works with an electronic implementation of a hybrid document, continuous monitoring in real time of the parameters of his face and keyboard handwriting is performed. These parameters are used to generate a private key that is used later to decrypt the corresponding parts of the document (Fig. 1). When fixing changes in the biometric characteristics of the subject, registered in the course of work, the document temporarily "changes" or "hides" its content entirely or completely, blocks some of the functions for editing it. The user will not be able to "bypass" the hybrid document management system, because all confidential information is encrypted on the appropriate cryptographic keys. Information publicly available to all subjects is not encrypted. This active protection technique is called the "living document" technology. To improve the reliability of key generation, you can use a multifactorial method that combines both proposed options.

A. Parameters of a handwritten signature

In computer representation, a signature can consist of functions of the position of the pen on the tablet x(t), y(t) and the pressure of the pen on the tablet p(t), where t is discrete time. Each handwritten image is subjected to spectral and correlation analysis in order to calculate a fixed number of informative features. This vector consists of both the quantities characterizing the appearance of the image (the distance between certain points of the signature image, the parameters of its slope, width, length), and the dynamics of its reproduction (the amplitudes of the harmonics of the functions x(t), y(t), p(t) corresponding to the oscillation frequency of the signer's hand (about 1-10 Hz), the correlation

coefficients between these and derived functions, the Daubechies wavelet transform coefficients D6). The process of calculating these features is described in more detail in [2].

B. Parameters of the face and keyboard handwriting of the subjects

As parameters of the face, we used some characteristics from [3] and [4], in particular:

- Distances between the eyes, the center of the face, the tip of the nose (in pixels, the values were normalized along the diagonal of the face in the frame).
- The areas of the eyes, nose, mouth (values were normalized by the area of the face).
- Correlation coefficients of brightness and color components of pixels (in accordance with the RGB model) between all pairs of the following areas of the face: eyes, nose, mouth. These signs characterize the asymmetry of the face.
- Parameters characterizing the color of the eyes and skin.

Hold and pause times between keystrokes were used as signs of keyboard handwriting.

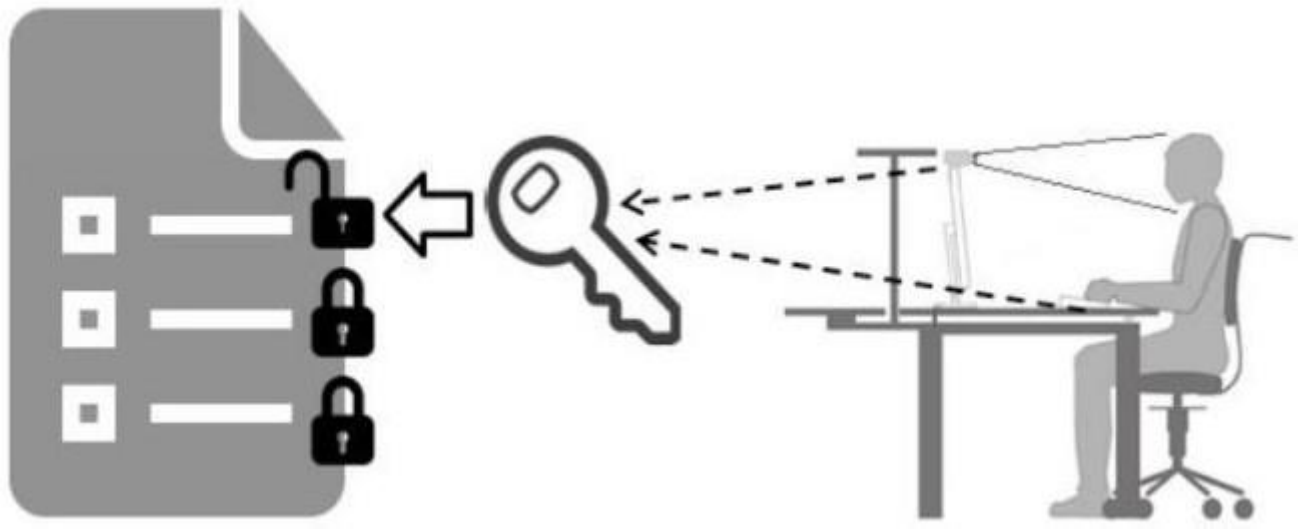


Fig.1. Illustration of the process of gaining access to a fragment of the electronic implementation of a hybrid document

III. METHODS AND TECHNOLOGIES FOR GENERATING A SECRET KEY OF AN ELECTRONIC SIGNATURE FROM BIOMETRIC DATA

The fuzzy extractor is based on the application of error-correcting codes. The cryptographic key is encoded with an error-correcting code, then the encoded sequence of bits is combined with the biometric characteristics of the subject, which are calculated from the data of the training sample. The output is "open string". In the process of authentication, the subject re-submits biometric data, which is added to the "open string" using the XOR operation. As a result, a key is released, the wrong bits of which are corrected. The fundamental disadvantages of fuzzy extractors include [1]:

1. All classical codes introduce redundancy. The greater the correcting ability of the code, the greater the redundancy and the shorter the length of the generated key-password. In a fuzzy extractor, the key length is strictly dependent on the correcting ability of the code.
2. Classical codes cannot correct a large number of errors, therefore, they cannot be used together with low-level formative or correlated biometric parameters (features).
3. Fuzzy extractors quantize “raw” (unprocessed, unenriched) biometric data and do not take into account the parameters of the distribution of feature values; as a result, they should give a higher proportion of errors in comparison with neural network converters biometrics-code , which in turn have these data, encoding them with the weight coefficients of neurons.

The application of this approach in relation to dynamic biometric images did not give good results [5, 6].

Artificial neural networks (ANNs) consist of interconnected computational elements (neurons) capable of learning, leading to the exposure of the quality of resolving the problem. Classical ANNs encode data on the features of signs by weighting coefficients of neuron synapses.

Special attention is currently being paid to deep learning technologies. The popularization of this direction is supported by large organizations (NVIDIA Corporation, Intel Corporation,

Google, Inc., Microsoft Corporation). Today, “deep” learning usually means iterative tuning of multilayer feedforward neural networks, in which the “error backpropagation” algorithm is used in one form or another.

It has two types of implementation: batch or stochastic gradient descent (in the second case, ANN optimization methods are used) [7]. Attempts are being made to apply deep learning methods for authenticating subjects based on dynamic biometric characteristics [8]. However, it is still difficult to use these techniques in real practice, since in order to achieve acceptable indicators, a significant amount of training sample is required (hundreds of examples of a biometric image and more).

In addition to large multilayer ANNs, active research is being carried out in the field of so called shallow networks.

These ANNs are capable of universal approximation, but this requires a potentially unlimited number of hidden neurons, which plays the role of the complexity of the ANN model and is a critical factor for practical implementation. A number of works are known in which the constraints of small ANNs were estimated and a number of theorems were formulated [9].

Lower bounds for the complexity of non-deep networks are obtained depending on the ratio between the range of values of the function being approximated and the dimension of the input

Attempts to create procedures for automatic estimation of the minimum required number of neurons based on the data of the training sample were undertaken earlier. In [11], an empirical relationship is given between the number of examples of the training sample and the size of the network in order to determine the upper threshold for the number of neurons in the hidden layer. These results are important for the development of biometric authentication methods, since they are related to the substantiation of the complexity of the ANN depending on theconstraints on the volume of the training sample. Small networks can learn from much fewer examples.

Neural networks that implement the change in weights and topology using evolutionary algorithms belong to the TWEANNs (Topology & Weight Evolving Artificial Neural Networks) group of networks.

This strategy for constructing and teaching ANN belongs to the category of reinforcement learning methods and has found application in conditions when it is almost impossible to perform supervised learning. Evolutionary augmentation of a neural network topology (NeuroEvolution of Augmenting Topologies, NEAT) uses genetic algorithms to adapt both topology and ANN weights.

The method uses a variation of parametric mutation, which is based on evolutionary strategies and evolutionary programming. Evolution begins with an ANN without hidden neurons and moves towards a more complex structure. This approach finds application in long-functioning and constantly learning ANNs based on large data volumes (car autopilot systems, obstacle detection, etc.) [12]. To date, with relatively small training samples, the evolutionary approach is successfully used to select topologies and weights in an ANN with one hidden layer. Only since 2014 availability hardware resources made it possible to apply neuroevolution to deep and convolutional neural networks, but already on very large samples [13]. An evolutionary approach can be used as the basis for mechanisms that allow changing ANN parameters in proportion to changes in the user's dynamic biometric image with time and depending on his state. In a sense, an analogue of "small" ANNs is the so-called "wide" neural networks. These ANNs are perceptrons, which consist of a large number of neurons, but a small number of layers (one or two) and are fundamentally different from shallow networks. The main difference is that for training "broad"

ANN does not use the principle of "error backpropagation". A non-iterative and absolutely stable learning algorithm for "wide" ANNs was first proposed in Uzbekistan several years ago for solving problems of biometric authentication [14]. Training is performed layer by layer, each neuron is trained independently of the other neurons in the network, based on the parameters of the feature distribution law calculated from the training sample data. To configure the automaton to recognize a certain subject, 20 examples of his image are enough. The high speed of work allows these algorithms to be implemented on a low-performance computing device.

Within the framework of the theory of "wide" ANNs, procedures for assessing the information content of features (through the areas of intersection of the probability density functions) began to be applied [15]. For the first time, it was proposed to create synapses taking into account the informativeness of signs, and to establish the number of neuron inputs, proceeding from the general informativeness of signs

This requirement is logical and allows one to justify the choice of many ANN parameters. The theory of "wide" ANNs has much in common with the methods of mathematical statistics and probability theory. The integration of different mathematical apparatuses made it possible to get off the ground in the issues of biometric authentication.

"Wide" neural networks can be configured to generate a fixed bit sequence upon arrival.

To the input of an image belonging to a certain class, and a random uniformly distributed sequence of bits ("white" noise) when an unknown image arrives at the input. Thus, the network data can be used to integrate biometric and electronic signatures.

IV. A HYBRID APPROACH TO THE CONSTRUCTION OF BIOMETRICS-CODE CONVERTERS THAT GENERATE A SECRET KEY OF AN ELECTRONIC SIGNATURE

In recent years, the active development of "wide" ANNs has taken place mainly along the path of hybrid neural network algorithms. After the rejection of the "back propagation of the error," it became possible to change not only the activation function of the neuron, but also its functional (the weighted summation functional was always used in the perceptron). In particular, for more efficient processing of weakly correlated biometric parameters, quadratic forms (1). Recent studies have shown that many functionals process data much more efficiently than perceptron adders and are able to work with strongly correlated features [17] These studies also show that highly correlated traits allow the creation of special neurons, the efficiency of which is much higher than that of neurons focused on processing independent traits [18]. For example, such neurons can be constructed on the basis of difference (2) or hyperbolic (3) multidimensional Bayesian functional. Thus, the correlation dependence between features can be perceived by the neural network as a special type of information about the image.

This circumstance radically changes the approach: from there is no need to get rid of uninformative and correlated features, they must be processed by separate types of neurons.

Hybrid "wide" ANNs have common features with networks of radial-basis functions, but they are more flexible. They can consist of several layers formed from different types of neurons and have cross-connections.

$$Ï = \sqrt{\sum_{j=1}^{N} \frac{(m_i - a_i)^2}{{\sigma_i}^2}} \quad (1)$$

where ai - is the value of the i-th biometric parameter (neuron input), mi and σi are the mathematical expectation and the root-mean-square deviation of the i-th feature (for the "Own" image), respectively, n is the dimension of the functional (the number of features , neuron inputs).

$$d_t = \sum_{j=1}^{N} \left| \frac{|m_t - a_t|}{\sigma_t} - \frac{|m_j - a_j|}{\sigma_j} \right|, j \neq t \quad (2)$$

$$d_t = \sum_{j=1}^{N} \left(\frac{(m_t - a_t)^2}{\sigma_t} - \frac{(m_j - a_j)^2}{\sigma_j} \right), j \neq t \quad (3)$$

where aj is the value of the i-th parameter (neuron input) with a high value of the correlation module |ri, t |in relation to the t-th biometric trait. That is, the tables of input connections of the Bayesian functional should be formed in such a way that the parameters enriched by it were as strong as possible correlated with each other.

After the enrichment of the input data of the neuron, the calculated value of the functional enters the activation function. In the simplest case, the activation function is a threshold howl. It is such a case that is considered within the framework of this article. The considered hybrid ANNs are focused on pattern recognition with a high dimension of the feature space and the presence of restrictions on the volume of

the training sample. Building and training these networks deterministic, while the correlation of biometric parameters is determined only on the basis of a small training sample, and the parameters themselves are determined in advance. Learning a "wide" neural network is layer-by-layer, i.e. each subsequent layer is trained on the output values of the neurons of the previous layer, perceiving them as feature values. It can be said that each layer of the "wide" network consists of several subnets, each of which has its own specificity in the tuning of neurons.

When implementing a biometrics-to-code converter in practice, it is important to take care that the user's biometric image and key are not compromised. The network segment, which is a perceptron, can be considered sufficiently protected from this threat [19].

It is impossible to extract the data of the training sample from the weights of neurons in a reasonable time and to recreate the standard of the biometric image, as well as to extract the user's private key (this is a computationally difficult and poorly formalized task). However, quadratic forms and Bayesian functionals operate directly with the parameters of the distribution laws of features, which leads to the need to store these parameters.

If the server on which the table of neural network functionals is stored is not trusted, then there is a threat of restoring fragments of the key and the standard of the user's biometric image from the data of the table of neural network functionals. In this case, to protect the biometric standard at the storage stage, the mechanism of a secure neural network container can be used [19]. This mechanism consists in the fact that the parameters of the neurons of the hybrid network are encrypted at the outputs of the neurons of the perceptron. If the key fragment is correctly issued by the neurons of the perceptron, the parameters of other neurons will be decoded. Otherwise, the network will generate noise because the decrypted values of the weight coefficients will not correspond to the standard of the subject.

Correcting codes from [20] make it possible to safely store error syndromes and do not make it possible to recover the key without presenting a biometric image that is sufficiently close to the authentic one. In combination with the mechanism of a secure neural network container [19], these noise-immune codes [20] solve the problem of secure storage of a table of neural network functionals. However, the question of the influence of the mechanism of a protected neural network container on the probability of erroneous decisions made by a "wide" neural network remains open.

In [21], the following hybrid network model is proposed, shown in Figure 2.

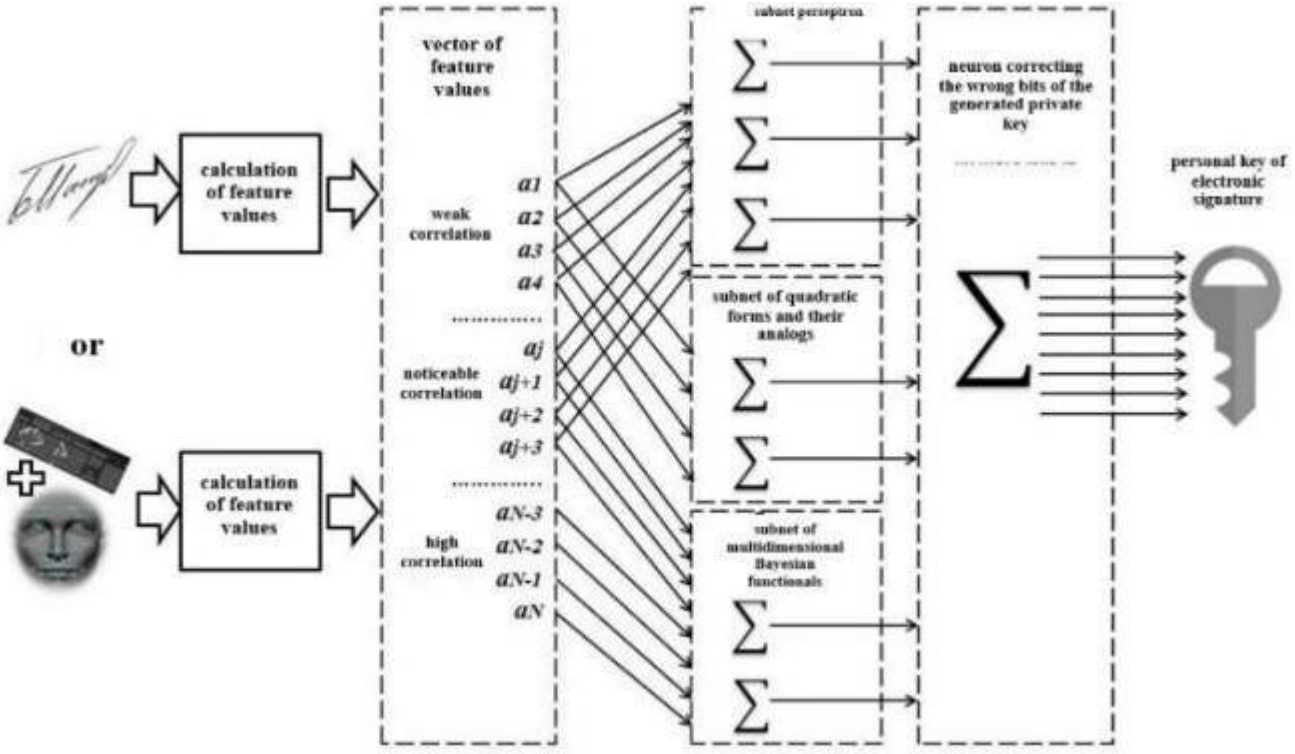


Fig.2. The scheme of work of the converter biometrics-code based on hybrid neural networks

For each subject, a separate neural network is formed and trained according to the data of his training sample. Each network is capable of generating an electronic signature key up to 2048 bits long. The reliability of the biometric authentication (identification) system is determined by the following indicators. The probability of a type 2 error (False Acceptance Rate, FAR) should be as low as possible. At the same time, the probability of a mistake of the 1st kind ("false rejection in admitting oneself", False Rejection Rate, FRR) should be acceptable, since frequent refusals create inconvenience. FAR and FRR also occur when an invalid key is generated (when FAR = FRR, one speaks of Equal Error Rate, EER).

In this work, a neural network biometrics-code converter with the indicated architecture was implemented as a software package with interfaces for entering signatures (handwritten passwords), a face, and keyboard handwriting (using a graphics tablet, a web camera, and keyboard). Using biometric data collected in the study [21], hybrid networks (Fig. 2) were reconfigured individually for each subject. Then, the same 90 subjects were recruited as in [21], who repeatedly went through the biometric identification procedure (the data were presented to all neural network transducers at once). Next, we checked the correctness of the keys generated using 90 hybrid networks (the key should be correct only in one of 90 cases). As a result, the following probabilities of erroneous decisions were achieved:

- for subject recognition / key generation from handwritten images: FRR = 18% with FAR <0.01% (EER = 3.5%);
- for subject recognition / key generation based on keyboard handwriting and face: FRR = 6.8% with FAR <0.01% (EER = 1.7%);

- for subject recognition / key generation based on handwritten images, keyboard handwriting, and face: FRR = 2.5% with FAR <0.01% (EER = 0.9%).

During the experiment, the mechanism of a secure neural network container was not used, however, changes in the subject's dynamic biometric image over time were taken into account (training of the biometrics-code converter and its testing were carried out on different days with a break from one to several weeks). In total, about 4500 attempts were made to go through the identification procedure and subsequent authentication.

CONCLUTION

As part of this work, an analytical study of methods for constructing biometrics-code converters used for the integration of biometric and electronic signatures as the basis of biometric authentication systems and electronic signature systems with biometric activation: fuzzy extractors, artificial multilayer neural networks, methods.

"Deep learning", as well as convolutional, evolutionary, small, "wide", hybrid neural networks. A model of a hybrid neural network based on perceptron, networks of quadratic forms, and multidimensional difference and hyperbolic Bayes functionals was tested. The high efficiency of using the model for solving the problem of integrating biometric and electronic signatures has been experimentally confirmed.

REFERENCES

1. Lozhnikov P.S. Biometric protection of hybrid document flow: monograph / Novo sibirsk: Publishing house of the SB RAS, 2017 .-- 130 p.
2. Lozhnikov P.S., Sulavko A.E., Eremenko A.V., Volkov D.A. Methods of Generating Key Sequences based on Parameters of Handwritten Passwords and Signatures // Information. - 2016. - No. 7 (4). - P. 59; DOI: 3390 / info7040059.
3. Vasiliev V.I., Lozhnikov P.S., Sulavko A.E., Zhumazhanova S.S. Assessment of the identification capabilities of biometric features from standard peripheral equipment // Information security issues. - 2016. - No. - S. 12-20.
4. Lozhnikov PS, Sulavko AE, Buraya EV, Pisarenko V.Yu. Authentication of computer users based on keyboard handwriting and facial features // Cybersecurity Issues. - 2017. - No. 3. - S. 24-34.
5. Lozhnikov P.S., Sulavko, A.E., Volkov D.A. Application of Noise Tolerant Code to Biometric Data to Verify the Authenticity of Transmitting Information / Control and Communications (SIBCON), 21-May 2015, Omsk, Russia. - P.1-3; DOI: 10.1109 / SIBCON.2015.7147126.
6. Lozhnikov P.S., Sulavko A.E., Eremenko A.V., Buraya E.V. Methods of Generating Key Sequences Based on Keystroke Dynamics // X
7. International IEEE Scientific and Technical Conference “Dynamics of Systems, Mechanisms and Machines” (Dynamics), November 15-17, 2016, Omsk, Russia. - P. 1-5; DOI: 10.1109 / Dynamics.2016.7819038.
8. Yasuoka Y., Shinomiya Y., Hoshino Y. Evaluation of Optimization Methods for Neural Network // Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems, 25-28 August 2016, Sapporo, Japan; DOI: 10.1109 / SCIS-ISIS.2016.0032.

9. Hafemann L. G. et al. Writer-independent Feature Learning for Offline Signature Verification Using Deep Convolutional Neural Networks //2016 International Joint Conference on Neural Networks (IJCNN), 24-29 July 2016. - P. 2576-2583; DOI: 10.1109 / IJCNN.2016.7727521
10. Kůrková V., Sanguineti M. Probabilistic Lower Bounds for Approximation by Shallow Perceptron Networks // Neural Networks. – 2017. – Vol. 91.– P. 34-41.
11. Kůrková V., Sanguineti M. Model Complexities of Shallow Networks Representing Highly Varying Functions // Neurocomputing. – 2016. – Vol. 171. – P. 598-604.
12. Rogers L.L., Dowla F.U.: Optimization of Groundwater Remediation Using Artificial Neural Networks with Parallel Solute Transport Modeling //. Water Resources Research. – 1994. – Vol. 30(2). – P. 457– 481.
13. Stanley K.O. Efficient Evolution of Neural Networks Through Complexification. PhD Thesis. Department of Computer Sciences, The University of Texas at Austin, 2004.
14. Koutn´ık J., Schmidhuber J., Gomez F. Evolving Deep Unsupervised Convolutional Networks for Vision-Based Reinforcement Learning // 2014 Annual Conference on Genetic and Evolutionary Computation. – 2014. – P. 541–548.
15. Volchikhin V.I., Ivanov A.I., Funtikov V.A. Bystryye algoritmy obucheniya neyrosetevykh mekhanizmov biometriko-kriptograficheskoy zashchity informatsii. Monografiya. Penza: Izd-vo Penzenskogo gosudarstvennogo universiteta. – 2005. – S. 273.
16. Sulavko A.E., Fedotov A.A., Eremenko A.V. Users' Identification through Keystroke Dynamics Based on Vibration Parameters and Keyboard Pressure // XI International IEEE Scientific and Technical Conference "Dynamics of Systems, Mechanisms and Machines" (Dynamics), 14-16 November, 2017, Omsk, Russia. – P. 1-7.
17. Ivanov A.I. Mnogomernaya neyrosetevaya obrabotka biometricheskikh dannykh s programmnym vosproizvedeniyem effektov kvantovoy superpozitsii. Penza: Izd-vo PNIEI. – 2016. – S. 133.
18. Ivanov A.I., Lozhnikov P.S., Sulavko A.Ye. Otsenka nadezhnosti verifikatsii avtografa na osnove iskusstvennykh neyronnykh setey, setey mnogomernykh funktsionalov Bayyesa i setey kvadratichnykh form // Komp'yuternaya optika. – 2017. – T. 41. – №5. – S.765-774; DOI: 10.18287/2412-6179-2017-41-5-765-774.
19. Ivanov A.I., Lozhnikov P.S., Vyatchanin S.E. Comparable Estimation of Network Power for Chisquared Pearson Functional Networks and Bayes Hyperbolic Functional Networks while Processing Biometric Data / Control and Communications (SIBCON), 29-30 June 2017, Astana, Kazakhstan. – P.1-3; DOI: 10.1109/ SIBCON.2017.7998435.
20. Akhmetov B.S., Ivanov A.I., Funtikov V.A., Bezyayev A.V., Malygina Ye.A. Tekhnologiya ispol'zovaniya bol'shikh neyronnykh setey dlya preobrazovaniya nechetkikh biometricheskikh dannykh v kod klyucha dostupa: Monografiya. / Almaty: TOO «Izdatel'stvo LEM». – 2014. – S. 144.
21. Bezyayev A. V., Ivanov A. I., Funtikova YU. V. Optimizatsiya struktury samokorrektiruyushchegosya biokoda, khranyashchego sindromy oshibok v vide fragmentov khesh-funktsiy // Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. – 2014. – № 3(13). – S. 4 -13.
22. Lozhnikov P.S., Sulavko A.Ye. Generation of a Biometrically Activated Digital Signature Based on Hybrid Neural Network Algorithms // Journal of Physics: Conf. Series. –2018. – № 1050; DOI: 10.1088/1742-6596/1050/1/012047

