UDC 004.032.26

MODIFIED LEARNING ALGORITHM FOR GMDH-WAVELET-NEURO-FUZZY-NETWORK IN INFORMATION TECHNOLOGIES

O. Vynokurova

Control System Research Laboratory, Information technologies security department, Kharkiv National University of Radio Electronics

vinokurova@kture.kharkov.ua

У статті запропоновано модифікований алгоритм навчання МГУА-вейлет-нейро-фаззі-мережі для вирішення задач обробки інформації. Для оптимізації структури вейвлет-нейро-фаззі системи запропоновано використовувати Метод Групового Урахування Аргументів (МГУА). Запропонована система дозволяє вирішувати широке коло задач, включаючи ідентифікацію та прогнозування сигналів, автентифікацію користувачів, класифікацію та кластеризацію інформації, розробку псевдо-випадкових генераторів на основі нейромереж в криптографії та інші. Імітаційне моделювання запропонованої архітектури та модифікованого алгоритму навчання підтверджує ефективність запропонованого підходу.

Ключові слова: гібридна МГУА-вейвлет-нейро-фаззі мережа, вейвлет-нейрон, модифікований алгоритм навчання, МГУА-алгоритми, ідентифікація, прогнозування, автентифікація

In the paper modified learning algorithm for GMDH-wavelet-neuro-fuzzy-network in information technologies is proposed. For Wavelet-Neuro-Fuzzy-Network structure optimization based on Group Method of Data Handling (GMDH) is developed and the method of structure optimization is described. Such hybrid systems can be used for solving many tasks including signal identification and prediction, person authentication, information classification and clustering, developing pseudo-random generator based on neural networks in cryptography and etc. The experimental investigations were carried out and their results accuracy of data processing by optimally constructed Wavelet-Neuro-Fuzzy-Network and network with multilayer feedforward architecture are presented and compared.

Keywords: Hybrid wavelet-neuro-fuzzy-network, wavelet neuron, modified learning algorithm, GMDH algorithms, information technologies, identification, forecasting, authentication

В статье предложен модифицированный алгоритм обучения МГУА-вэйвлет-нейро-фаззи сети для решения задач обработки информации. Для оптимизации структуры вэйвлет-нейро-фаззи системы предложено использовать Метод Группового Учета Аргументов (МГУА). Предложенная система позволяет решать широкий круг задач, включая идентификацию и прогнозирование сигналов, аутентификацию пользователей, классификацию и кластеризацию информации, синтез псевдо-случайных генераторов на основе нейросетей в криптографии и другие. Имитационное моделирование предложенной архитектуры и модифицированного алгоритма обучения подтверждает эффективность развиваемого подхода.

Ключевые слова: гибридная МГУА-вэйвлет-нейро-фаззи-сеть, вэйвлет-нейрон, модифицированные метод обучения, МГУА-алгоритмы, информационные технологии, идентификация, прогнозирование, аутентификация

Introduction

Last years the problem of information processing tasks including identification, authentication, prediction, clustering, developing pseudo-random generator in cryptography and etc is of great importance [1-3]. For its solution various approaches were applied. The most perspective information processing methods are neural networks, especially a fuzzy neural networks and GMDH methods. Earlier it was proved that neural networks are universal approximators and have some remarkable properties, such as parallel processing of information, ability to work with incomplete noisy input data, and learning possibilities to achieve the desired response (output).

The GMDH [4-6], from the other side, uses the principle of self-organization that allows to construct an optimal structure of the forecasting model during the algorithm operation. It's very promising to combine advantages of these both approaches for the solution of the problem – constructing an efficient model for the forecasting in different applications including financial ones.

The goal of the present work is a synthesis of learning algorithm on the turning points of GMDH-Wavelet-Neuro-Fuzzy-Network based on hybrid criterion.

In [7] GMDH-network with neo-fuzzy nodes using triangular and cubic-spline activation membership function was introduced. In this paper as nodes of GMDH-network the wavelet neurons with adaptive membership-activation functions which have an improved approximation and extrapolation properties in comparison with neo-fuzzy neuron is proposed. Also the specialized learning algorithm based on hybrid optimization criterion is introduced.

1. Wavelet neuron architecture

Let us consider wavelet-neuron architecture [8], shown on the Fig.1. As it can be seen, wavelet neuron is close enough in construction to the neo-fuzzy neuron, however instead of usual tuning synaptic weights it contains wavelet synapses WS_i, $i = 1, 2, \ldots, n$. In this case tuning parameters are not only synaptic weights $w_{ji}(k)$, but center, width and shape parameters of adaptive wavelet membership-activation function $\varphi_{ji}(x_i(k))$.

When a vector signal

$$x(k) = (x_1(k), x_2(k), \ldots, x_n(k))^T = (x(k-1), x(k-2), \ldots, x(k-n))^T$$

is fed to the input of wavelet neuron, the output is determined by both the tunable weights $w_{ji}(k)$ and wavelet membership activation function:

$$\hat{y}(k) = \sum_{i=1}^{n} f_i(x(k)) = \sum_{i=1}^{n} \sum_{j=1}^{h_i} w_{ji}(k-1)\varphi_{ji}(x_i(k)). \tag{1}$$
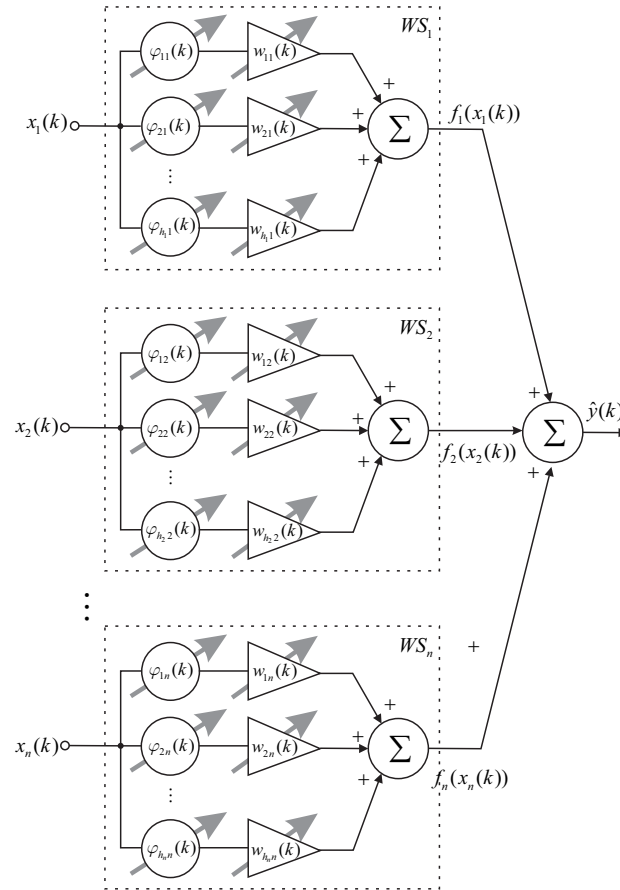
Fig. 1. Architecture of wavelet neuron

In this case we use one-dimensional wavelet membership-activation function, proposed in [9] in the form

$$\varphi_{ji}(x_i(k)) = (1 - \alpha_{ji}(k)t_{ji}^2(k))\exp\left(-\frac{t_{ji}^2(k)}{2}\right), \tag{2}$$

where $t_{ji}(k) = \left(x_i(k) - c_{ji}(k)\right)\sigma_{ji}^{-1}(k)$; $c_{ji}(k)$ is center parameter of wavelet membership-activation function; $\sigma_{ji}(k)$ is width parameter of wavelet membership-activation function; $\alpha_{ji}(k)$ – shape parameter of wavelet membership-activation function.

2. Wavelet neuron learning algorithm

The learning task is to estimate for each iteration k synaptic weights $w_{ji}(k)$, centers $c_{ji}(k)$, widths $\sigma_{ji}^{-1}(k)$ and shape parameter $\alpha_{ji}(k)$ of wavelet membership-activation function, which optimized some prescribed criteria.

2.1. Learning algorithm for synaptic weights, width parameters, shape parameters of wavelet membership-activation function

As the criterion of optimization for wavelet neuron synaptic weights, width and shape parameters of membership-activation function quadratic error function is used in the form

$$E(k) = \frac{1}{2}(y(k) - \hat{y}(k))^2 = \frac{1}{2}e^2(k) = \frac{1}{2}(y(k) - \sum_{i=1}^{n}\sum_{j=1}^{h_i} w_{ji}(k)\varphi_{ji}(x_i(k)))^2 \qquad (3)$$

(here $y(k)$ – external training signal).

The derivatives of the error function with respect to the tuned parameters can be written in the form

$$\frac{\partial E(k)}{\partial w_{ji}(k)} = -e(k)\varphi_{ji}(x_i(k)) = -e(k)(1 - \alpha_{ij}(k)t_{ji}^2(k))\exp\left(-\frac{t_{ji}^2(k)}{2}\right) = -e(k)J_{ji}^w(k), \quad (4)$$

$$\frac{\partial E(k)}{\partial \sigma_{ji}^{-1}(k)} = - e(k)w_{ji}(k)(x_i(k) - c_{ji}(k))((2a_{ji}(k) + 1)t_{ji}(k) - a_{ji}(k)t_{ji}^3(k))\cdot$$

$$\qquad (5)$$

$$\cdot \exp\left(-\frac{t_{ji}^2(k)}{2}\right) = - e(k)J_{ji}^s(k),$$

$$\frac{\partial E(k)}{\partial \alpha_{ji}(k)} = e(k)w_{ji}(k)t_{ji}^2(k)\exp\left(-\frac{t_{ji}^2(k)}{2}\right) = -e(k)J_{ji}^\alpha(k). \qquad (6)$$

Introducing $(h_i \times 1)$ –vectors of variables $\varphi_i(x_i(k)) = (\varphi_{1i}(x_i(k)),\ldots,\varphi_{h_i i}(x_i(k)))^T$, $w_i(k) = (w_{1i}(k),\ldots,w_{h_i i}(k))^T$, $\sigma_i^{-1}(k) = (\sigma_{1i}^{-1}(k),\ldots,\sigma_{h_i i}^{-1}(k))^T$, $\alpha_i(k) = (\alpha_{1i}(k),\ldots,\alpha_{h_i i}(k))^T$, $t_i(k) = (t_{1i}(k),\ldots,t_{h_i i}(k))^T$, $J_i^w(k) = (J_{1i}^w(k),\ldots,J_{h_i i}^w(k))^T$, $J_i^\sigma(k) = (J_{1i}^\sigma(k),\ldots,J_{h_i i}^\sigma(k))^T$, $J_i^\alpha(k) = (J_{1i}^\alpha(k),\ldots,J_{h_i i}^\alpha(k))^T$, we can rewrite gradient learning algorithm of i-th wavelet-synapse:

$$\begin{cases} w_i(k+1) = w_i(k) + \eta^w(k)e(k)J_i^w(k), \\ \sigma_i^{-1}(k+1) = \sigma_i^{-1}(k) + \eta^\sigma(k)e(k)J_i^\sigma(k), \\ \alpha_i(k+1) = \alpha_i(k) + \eta^\alpha(k)e(k)J_i^\alpha(k). \end{cases} \qquad (7)$$

The convergence rate of learning algorithm can be increased via the use of Gauss-Newton algorithms [10].

Using the sum of two matrix inversion lemma and performing a sequence of obvious transformations [11], we can write on-line learning algorithm in the form

$$
\begin{cases}
w_i(k+1) = w_i(k) + \dfrac{e(k)J_i^w(k)}{r_i^w(k)}, & r_i^w(k+1) = \beta r_i^w(k) + P J_i^w(k) P^2, \\[3mm]
\sigma_i^{-1}(k+1) = \sigma_i^{-1}(k) + \dfrac{e(k)J_i^\sigma(k)}{r_i^\sigma(k)}, & r_i^\sigma(k+1) = \beta r_i^\sigma(k) + P J_i^\sigma(k) P^2, \\[3mm]
\alpha_i(k+1) = \alpha_i(k) + \dfrac{e(k)J_i^\alpha(k)}{r_i^\alpha(k)}, & r_i^\alpha(k+1) = \beta r_i^\alpha(k) + P J_i^\alpha(k) P^2, \ 0 \le \beta \le 1,
\end{cases}
\tag{8}
$$

where β is forgetting factor ($0 \le \beta \le 1$).

2.2. The learning algorithm on the turning points for center parameters of wavelet membership-activation functions

Generally the learning algorithms for neural networks are based on using some learning error function. But in the situation when the processed signal or process is significantly non-stationary using such criteria leads to shift effect, which reduces the accuracy of prediction [12].

For the solving such problem, we need to introduce learning criterion that can be consider the shift effect of prediction signal.

Fig. 2 shows the fragment of signal $y(k)$ with shift effect of predictive value $\hat{y}(k)$.
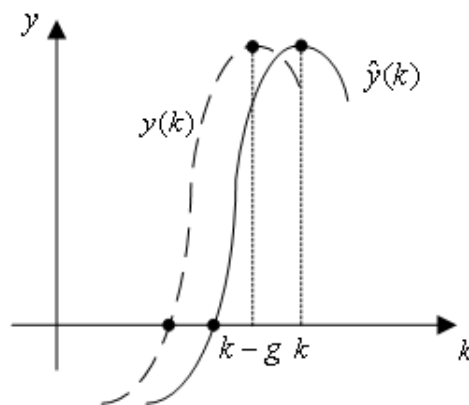

Fig. 2. Signal fragment with prediction shift effect

It can be seen on the fig. 2 that, in order to minimizing of shift effect we need to minimize distance g or at the inflection points of the signal, or at the points of intersection with the axis k. It should be noted, that inflection points of $y(k)$ and

$\hat{y}(k)$ is the intersection axis of $\Delta y(k)$ and $\Delta \hat{y}(k)$ corresponding (here Δ is symbol of first difference).

In the prediction theory of economical time series in German literature such prediction quality criteria as Wegstreke [13, 14] is accepted. This criterion is the estimation of predicting model quality, when its value $+1$ corresponds optimal predicting model, and when its value -1 corresponds incorrect prediction result. Such criterion has form

$$Wegstreke = \frac{\sum_{k=1}^{N} signal(k)(y(k) - y(k-1))}{\sum_{k=1}^{N} |y(k) - y(k-1)|} \qquad (9)$$

where $signal(k)$ is signum function

$$signal(k) = \begin{cases} 1, & if \ \hat{y}(k) - y(k-1) > 0, \\ -1, & if \ \hat{y}(k) - y(k-1) < 0, \\ 0 & otherwise, \end{cases}$$

$y(k)$ is the actual value of process; $\hat{y}(k)$ is prediction value; N is length of learning sample.

Due to the fact, that the sign function is not differentiable, it can be replaced by a hyperbolic tangent function (see fig. 3) with large steepness parameter

$$signal(k) \approx \tanh \gamma(\hat{y}(k) - y(k-1)) = \frac{1 - \exp(-2\gamma(\hat{y}(k) - y(k-1)))}{1 + \exp(-2\gamma(\hat{y}(k) - y(k-1)))},$$
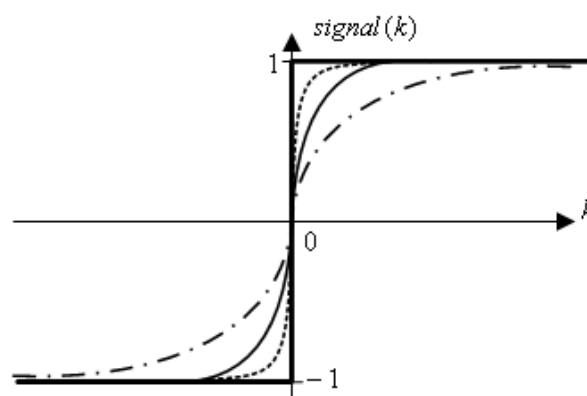
where γ is steepness parameter, $\gamma \gg 1$.

Fig. 3. Signum-function (solid, thick line) and hyperbolic tangent function ($\gamma > 1$ - dotted line, $\gamma = 1$ - solid line, $\gamma < 1$ - dash-dotted line)

Using this hypothesis, we can introduce the learning criterion, which will take into account the shifting effect of prediction in the form

$$E^c(k) = \frac{(\tanh \gamma(\hat{y}(k) - y(k-1)))(y(k) - y(k-1))}{|y(k) - y(k-1)|}, \qquad (10)$$

and in this case the derivative with respect to the parameter $c_{ji}(k)$ has form

$$\frac{\partial E^c(k)}{\partial c_{ji}(k)} = \gamma \, sign(y(k) - y(k-1)) \times$$
$$\times (1 - \tanh^2(\gamma(\hat{y}(k) - y(k-1)))) w_{ji}(k) \frac{\partial \varphi_{ji}(x_i(k))}{\partial c_{ji}(k)}. \qquad (11)$$

Using expression (4) we can rewrite learning algorithm on the turning points of wavelet membership-activation function centers based on gradient procedure in the form

$$c_{ji}(k+1) = c_{ji}(k) - \eta_c \gamma \, sign(y(k) - y(k-1)) \times$$
$$\times (1 - \tanh^2(\gamma(y(k) - y(k-1)))) \varphi_{ji}^c(x_i(k)) \qquad (12)$$

where

$$\varphi_{ji}^c(x_i(k)) = w_{ji}(k)\sigma_{ji}^{-1}(k)((2\alpha_{ji}(k)+1)t_{ji}(k) - \alpha_{ji}(k)t_{ji}^3(k))\exp\left(-\frac{t_{ji}^2(k)}{2}\right),$$

$w_{ji}(k)$ is synaptic weight of wavelet neuron; η_c is learning rate parameter, $0 < \eta_c < 1$.

Such learning algorithm has enough high speed and computational simplicity, and main its advantage is possibility of minimizing shift between actual signal and forecast, that it is important for the solving many practical tasks.

3. Wavelet-Neuro-Fuzzy-Network and its architecture optimization using the Group Method of Data Handling

Wavelet-Neuro-Fuzzy-Network is a multilayer feedforward architecture that consists of wavelet neurons. 3-layers Wavelet-Neuro-Fuzzy-Network with *n* inputs and *m* outputs is shown of Fig. 4.

Given architecture is completely coincides with the structure of the 3-layer perceptron, except that the wavelet neurons are used here as an elementary nodes instead of usual neurons.
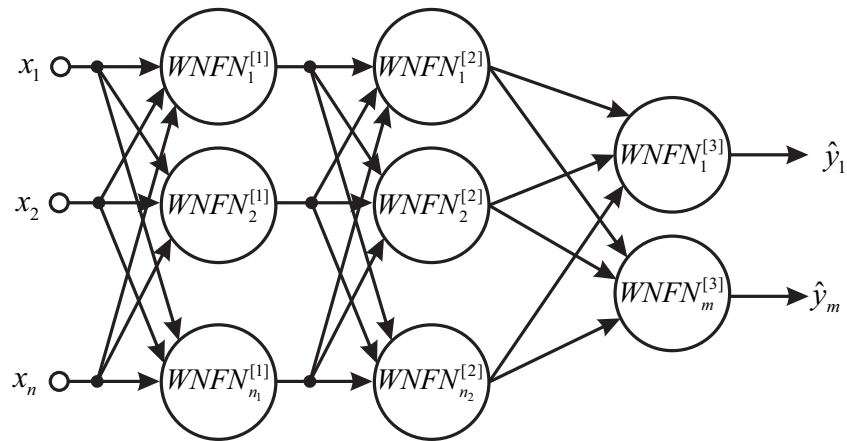
Fig. 4. GMDH-Wavelet-Neuro-Fuzzy-Network

If we use wavelet neurons that have only two inputs, the GMDH can be applied for the synthesis of the Wavelet-Neuro-Fuzzy-Network with optimal architecture.

The main idea of the GMDH algorithm lays in successive synthesis of the neuron layers until the external criterion begins to increase. Algorithm description:

1) Form pairs from the wavelet neuron outputs of the current layer (at the first iteration we use the set of input signals). Each pair is fed to the corresponding wavelet neuron.

2) Using the learning subsample to adjust synaptic weight coefficients of each wavelet neuron.

3) Using the test subsample to calculate the value of the external criterion (regularity) for each wavelet neuron:

$$\varepsilon_p^{[s]} = \frac{1}{N_{nep}} \sum_{k=1}^{N_{nep}} \left(y(k) - \hat{y}_p^{[s]}(k) \right)^2 \tag{13}$$

where N_{nep} is a size of the test subsample, s is the layer number, p is a neuron number in the current layer $p = \overline{1, n_s}$, $\hat{y}_p^{[s]}(k)$ is the *p*-th neuron of the *s*-th layer response signal for the *i*-th input vector.

4) Find the minimal value of the external criteria for all wavelet neurons of the current layer

$$\varepsilon^{[s]} = \min_p \varepsilon_p^{[s]} \tag{14}$$

and check the condition

$$\varepsilon^{[s]} > \varepsilon^{[s-1]}, \tag{15}$$

where $\varepsilon^{[s]}, \varepsilon^{[s-1]}$ are the criterion values for the best neurons of the and *s*-th and (*s*-1)-th layers correspondingly. If the condition (15) is true then return to the previous layer and find the best neuron that provides minimal value of the criterion (13).

Otherwise, select *F* best neurons according to the criterion (13) value and go to the step 1 to construct the next layer of neurons.

5) Determine the final structure of the network. Moving backward from the best neuron of the (m-1)-th layer along the input connections and passing successively all the layers of neurons, preserve in the final structure only such neurons that are used in the next layer.

After the GMDH finishes its functioning it can be said that the final optimal structure of the Wavelet-Neuro-Fuzzy-Network is synthesized. As it can be readily seen we obtain not only optimal structure, but also trained neural network that is ready to process new data. One of the most important advantages of GMDH usage for the Wavelet-Neuro-Fuzzy-Network architecture synthesis is a capability to use simple but quick learning procedures for the wavelet neuron weights adjustment because network is trained layer-by-layer.

Conclusion

In the paper the modified learning algorithm for GMDH-wavelet-neuro-fuzzy-network in information technologies is proposed. Using Group Method of Data Handling algorithms we can synthesize an optimal architecture of the Wavelet-Neuro-Fuzzy-Network. Theoretical justification and experimental results prove the efficiency of the developed approach to the Wavelet-Neuro-Fuzzy-Network architecture self-organization.

References

1. Desai V.V. Pseudo random number generator using time delay neural network / Desai V.V., Patil R., Deshmukh V.B., Rao D.H. // World Journal of Science and Technology. – 2012. - 2(10). – P. 165-169.

2. Du K.-L. Neural Networks and Statistical Learning / Du K.-L., Swamy M.N.S. - Springer-Verlag, London, 2014. – 824 p.

3. Lian S. Secure Hash Function based on neural network / Lian S., Sun J., Wang Z. // Neurocomputing. – 2006. – 69(16-18). – P. 2346-2350.

4. Ivakhnenko A.G. Inductive sorting-out GMDH algorithms with polynomial complexity for active neurons of neural network / Ivakhnenko A.G. // Neural Networks. – 1999. – 2. – P. 1169-1173.

5. Ivakhnenko A.G. Heuristic self-organization in problems of engineering cybernetics / Ivakhnenko A.G. // Automatica. – 1970. - 6(2). – P. 207-219.

6. Ivakhnenko A.G. Self-organization of neuro net with active neurons for effects of nuclear test explosions forecasting // Ivakhnenko A.G. // System Analysis Modeling Simulation. – 1995. – 20. – P. 107-116.

7. Bodyanskiy Ye. The neo-fuzzy neural network structure optimization using the GMDH for the solving forecasting and classification problems / Bodyanskiy Ye., Zaychenko Yu., Pavlikovskaya E., Samarina M., Viktorov Ye. // Proc. Int.Workshop on Inductive Modeling, Krynica, Poland. – 2009. – P. 77-89.

8. Bodyanskiy Ye. An adaptive learning algorithm for a wavelet neural network / Bodyanskiy Ye., Lamonova N., Pliss I., Vynokurova O. // Expert Systems. – 2005. - 22(5). – P. 235-240.

9. Bodyanskiy Ye. Compartmental adaptive wavelon and its learning algorithm // Bodyanskiy Ye., Vynokurova O. // Control systems and computers. – 2009. - 1 (219). – P. 47-53 (In Russian).

10. Shepherd A.J. Second-Order Methods for Neural Networks // Shepherd A.J. - London: Springer-Verlag, 1997. – 145 p.

11. Bodyanskiy Ye. An adaptive learning algorithm for a neuro-fuzzy network. / Bodyanskiy Ye., Kolodyazhniy V., Stephan A. // Computational Intelligence and Applications [Ed. by B. Reusch]. Berlin-Heidelberg-New-York: Springer. - 2001. – P. 68-75.

12. Lekutai G. Self-tuning control of nonlinear systems using neural network adaptive frame wavelets / Lekutai G., van Landingham H. F. // Proc. IEEE Int. Conf. on Systems, Man and Cybernetics. 1997. – 2. – P. 1017-1022.

13. Baumann M. Nutzung neuronaler Netze zur Prognose von Aktienkursen / Baumann M. // Report Nr. 2/96, TU Ilmenau. - 1996. – 113 p.

14. Fueser K. Neuronale Netze in der Finanzwirtshaft / Fueser K. - Wiesbaden: Gabler. - 1995. – 437 p.