DE GRUYTER OPEN

Information Technology and Management Science doi: 10.1515/itms-2015-0011

2015 / 18

Hybrid Generalised Additive Type-2 Fuzzy-Wavelet-Neural Network in Dynamic Data Mining

Yevgeniy Bodyanskiy[1], Olena Vynokurova[2], Iryna Pliss[3], Yuliia Tatarinova[4]

[1, 2, 3, 4] *Kharkiv National University of Radio Electronics*

***Abstract* – In the paper, a new hybrid system of computational intelligence is proposed. This system combines the advantages of neuro-fuzzy system of Takagi-Sugeno-Kang, type-2 fuzzy logic, wavelet neural networks and generalised additive models of Hastie-Tibshirani. The proposed system has universal approximation properties and learning capability based on the experimental data sets which pertain to the neural networks and neuro-fuzzy systems; interpretability and transparency of the obtained results due to the soft computing systems and, first of all, due to type-2 fuzzy systems; possibility of effective description of local signal and process features due to the application of systems based on wavelet transform; simplicity and speed of learning process due to generalised additive models. The proposed system can be used for solving a wide class of dynamic data mining tasks, which are connected with non-stationary, nonlinear stochastic and chaotic signals. Such a system is sufficiently simple in numerical implementation and is characterised by a high speed of learning and information processing.**

***Keywords* – Computational intelligence, evolutionary computations, fuzzy neural networks, hybrid intelligent systems.**

I. INTRODUCTION

Nowadays computational intelligence methods and especially hybrid systems of computational intelligence [1]–[3] are wide spread for Data Mining tasks in different areas under uncertainty, non-stationary, nonlinearity, stochastic, chaotic conditions of the investigated objects and, first of all, in control, identification, prediction, classification, emulation etc. These neural networks are flexible because they combine effective approximating properties and learning ability of artificial neural networks [4], [5], transparency and interpretability of the results obtained by using neuro-fuzzy systems [6] and especially type-2 fuzzy-neural networks [7], the possibility of a compact description of the local features of non-stationary signals, providing wavelet neural networks [8] and more advanced wavelet-neuro-fuzzy systems [9], [10].

At present, more and more problems arise where the information should be processed in real time when data arrival frequency may be high enough. Such problems are considered within areas known as Dynamic Data Mining [11] and Data Stream Mining [12], [13].

The most known systems of computational intelligence in this situation are completely ineffective because their learning process is carried out in a multi-epoch batch mode when the training set is defined in the form of observation fixed set that cannot be changed under a learning process. Certainly, there is a wide class of neural networks whose output signal is linearly dependent on tuning synaptic weights, and it follows that these weights can be tuned by the speed recursive procedures. However, these networks suffer from the so-called curse of dimensionality (especially networks using lazy learning), and optimisation of the neuron number is produced using off-line clustering procedures. Neuro-fuzzy systems have undoubted advantages over neural networks and first of all the significantly smaller number of tuning synaptic weights using scatter partitioning of input space. However, in these systems to provide the required approximating properties not only the synaptic weights but also membership functions (centres and widths) must be tuned. Furthermore, the training process of these parameters is performed using backpropagation algorithms, which are very difficult to implement in a real-time mode.

Problems with the centres and widths of membership functions could be overcome by using the concept of type-2 fuzzy systems [14]–[17], but the type of reduction procedure, which is present in these systems, is too tedious from the computational point of view. At the same time, the hybrid wavelet neuro-fuzzy systems [10], [18] are too bulky. Although these systems have increased approximating properties, they require off-line tuning of the membership functions-wavelets.

For processing in the real-time mode, the so-called generalised additive models [19], [20] suit in the best way, but they are not oriented to the processing of non-stationary, nonlinear, chaotic signals.

As a result, the development of hybrid real-time system of computational intelligence is reasonable. Such a system will combine the advantages of traditional neural networks, neuro-fuzzy systems, type-2 fuzzy systems, wavelet neural networks, generalised additive models for solving a wide class of problems, which appear in the Dynamic Data Mining and Data Stream Mining.

II. GENERAL REGULATION ARCHITECTURE OF HYBRID GENERALISED ADDITIVE TYPE-2 FUZZY-WAVELET-NEURAL NETWORK (HGAT2FWNN)

Fig. 1 shows the architecture of the proposed hybrid generalised additive type-2 fuzzy-wavelet-neural network (HGAT2FWNN).

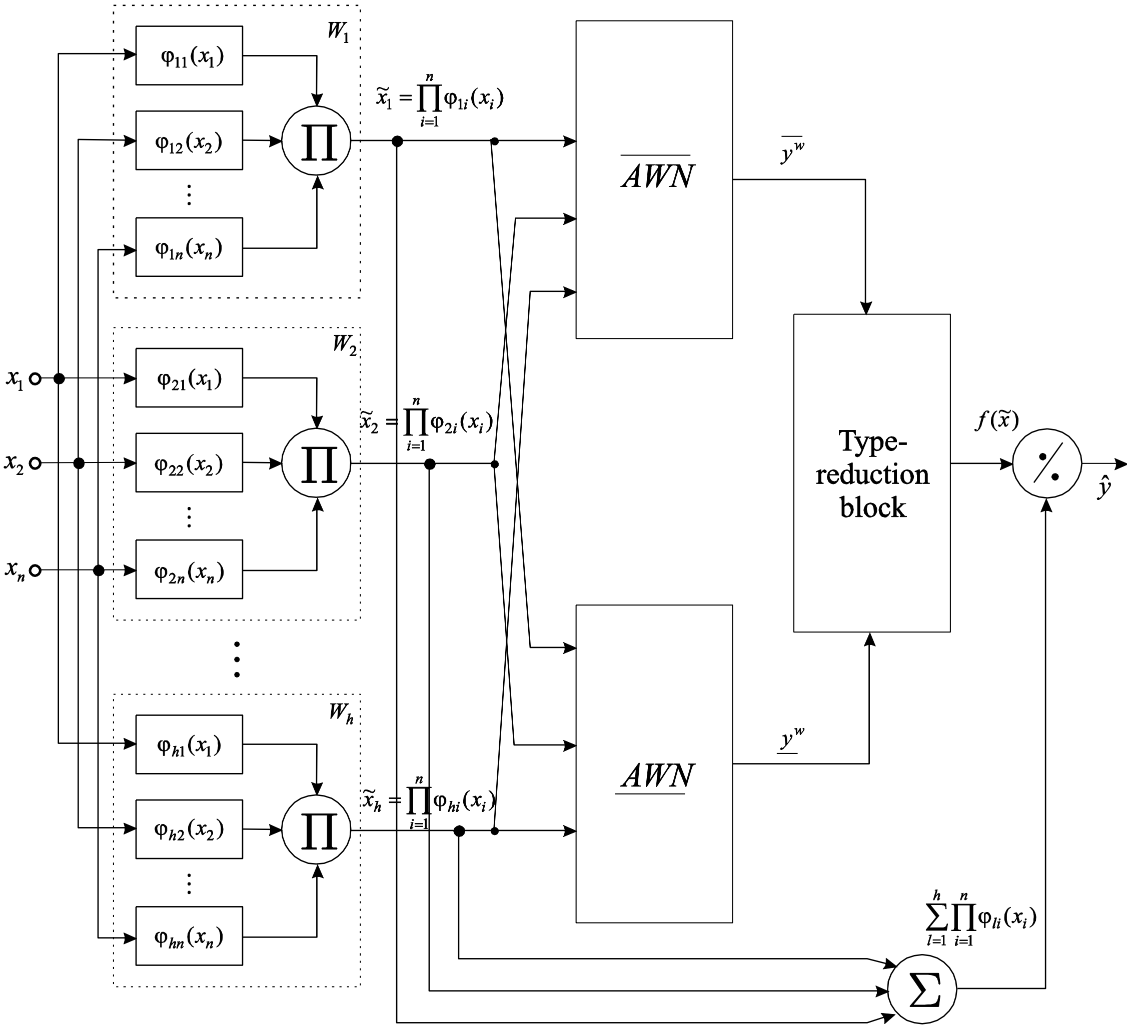


Fig. 1. Architecture of HGAT2FWNN.

This system consists of four layers of information processing; the first and second layers are similar to the layers of TSK-neuro-fuzzy system [21]–[23]. The only difference is that the odd wavelet membership functions "Mexican Hat", which are "close relative" of Gaussians, are used instead of conventional bell-shaped Gaussian membership function in the first hidden layer

$$\varphi_{li}(x_i(k)) = (1 - \tau_{li}^2(k)) \exp(-\tau_{li}^2(k)/2),$$

where $x(k) = (x_1(k), x_2(k), \ldots, x_n(k))^T \in R^n$ is the vector of input signals in the current moment of time $k = 1,2,\ldots$; $\tau_{li}(k) = (x_i(k) - c_{li})\sigma_{li}^{-1}$; c_{li}, σ_{li} are the centre and width of the corresponding membership function implying that $\underline{c} \le c_{li} \le \bar{c}$; $\underline{\sigma} \le \sigma_{li} \le \bar{\sigma}$; $i = 1,2,\ldots,n$; $l = 1,2,\ldots,h$; n is the input number; h is the membership function number.

It is necessary to note that using the wavelet functions instead of common bell-shaped positive membership functions gives the system more flexibility [24], and using odd wavelets for the fuzzy reasoning does not contradict the ideas of fuzzy inference, because the negative values of these functions can be interpreted as non-membership levels [25].

Thus, if the input vector $x(k)$ is fed to the system input, then in the first layer the hn levels of membership functions $\varphi_{li}(x_i(k))$ are computed and in the hidden layer h vector product blocks perform the aggregation of these memberships in the form

$$\tilde{x}_l(k) = \prod_{i=1}^{n} \varphi_{li}(x_i(k)).$$

This means the input layers transform the information similarly to the neurons of the wavelet neural networks [8], [26], which form the multidimensional activation functions

$$\prod_{i=1}^{n} (1 - \tau_{li}^2(k)) \exp(-\tau_{li}^2(k)/2)$$

providing a scatter partitioning of the input space.

As a result, the signals in the output of the second layer can be written in the form

$$\tilde{x}_l(k)=\prod_{i=1}^{n}\left(1-\frac{(x_i(k)-c_{li})^2}{\sigma_{li}^2}\right)\exp\left(-\frac{(x_i(k)-c_{li})^2}{2\sigma_{li}^2}\right).$$

To provide the required approximation properties, the third layer of system is formed based on type-2 fuzzy wavelet neuron (T2FWN) [27], [28]. This neuron consists of two adaptive wavelet neurons (AWN) [26], [29], whose prototype is a wavelet neuron of T. Yamakawa [30]. Wavelet neuron is different from the popular neo-fuzzy neuron [31], [32] that uses the odd wavelet functions instead of the common triangular membership functions. The use of odd wavelet membership functions, which form the wavelet synapses $WS_1,\ldots,WS_l,\ldots,WS_h$, provides higher quality of approximation in comparison with nonlinear synapses of neo-fuzzy neurons. Fig. 2 shows the architecture of wavelet-neuron.

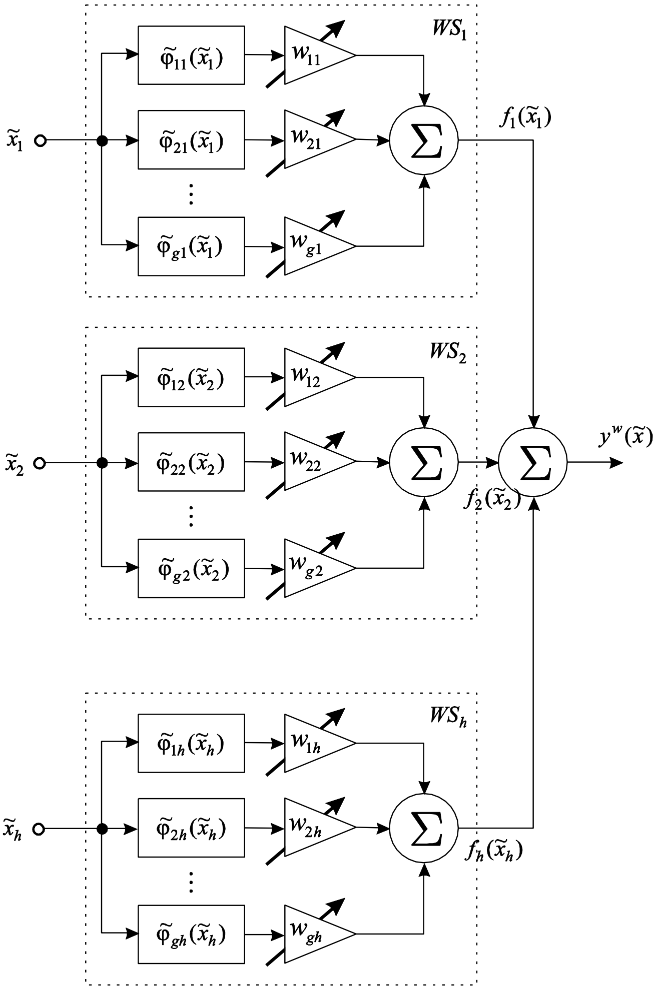


Fig. 2. Wavelet-neuron architecture.

In such a way the wavelet neuron performs the mapping in the form

$$y^w(\tilde{x}(k))=\sum_{l=1}^{h} f_l(\tilde{x}_l(k))$$

where $\tilde{x}(k)=(\tilde{x}_1(k),\ldots,\tilde{x}_l(k),\ldots,\tilde{x}_h(k))^T$, $y^w(\tilde{x}(k))$ is the scalar output of wavelet neuron. Each wavelet synapse WS_l consists of g wavelet membership functions $\tilde{\varphi}_{jl}(\tilde{x}_l)$, $j=1,2,\ldots,g$ (g is a wavelet membership function number in the wavelet neuron) and the same number of the tuning synaptic weights w_{jl}. Thus, the transform is implemented by each wavelet synapse WS_l in the k-th instant of time, which can be written in form

$$f_l(\tilde{x}_l(k))=\sum_{j=1}^{g} w_{jl}(k-1)\tilde{\varphi}_{jl}(\tilde{x}_l(k))$$

(here $w_{jl}(k-1)$ is the value of synaptic weights that are computed based on previous $k-1$ observations), and the general wavelet neuron performs the nonlinear mapping in the form

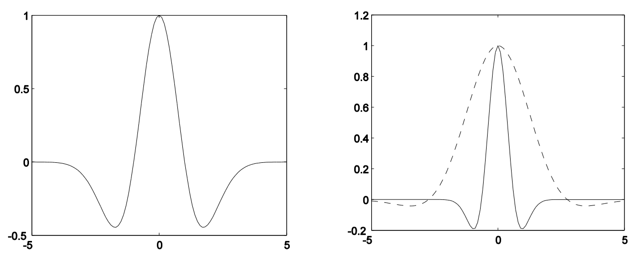
$$y^w(\tilde{x}(k))=\sum_{l=1}^{h}\sum_{j=1}^{g} w_{jl}(k-1)\tilde{\varphi}_{jl}(\tilde{x}_l(k)),$$

i.e., in fact, this is the generalised additive model [18], [19] that is characterised by the simplicity of computations and high approximation properties.

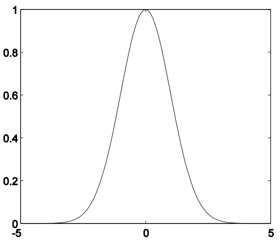
Under uncertain, stochastic or chaotic conditions, it is more effective to use the adaptive wavelet neuron (AWN) [26] instead of common wavelet neuron. The adaptive wavelet neuron is based on the adaptive wavelet function [24] in the form

$$\tilde{\varphi}_{jl}(\tilde{x}_l(k))=(1-\alpha_{jl}\tau_{jl}^2(k))\exp(-\tau_{jl}^2(k)/2)=$$
$$=\left(1-\alpha_{jl}\frac{(\tilde{x}_l(k)-c_{jl})^2}{\sigma_{jl}^2}\right)\exp\left(-\frac{(\tilde{x}_l(k)-c_{jl})^2}{2\sigma_{jl}^2}\right)$$

where $0\le\alpha_{jl}\le 1$ is the shape parameter of adaptive wavelet function, if $\alpha_{jl}=0$ it is conventional Gaussian, if $\alpha_{jl}=1$ it is the wavelet "Mexican Hat", and if $0<\alpha_{jl}<1$ it is some hybrid activation-membership function (see Fig. 3).



a) b)

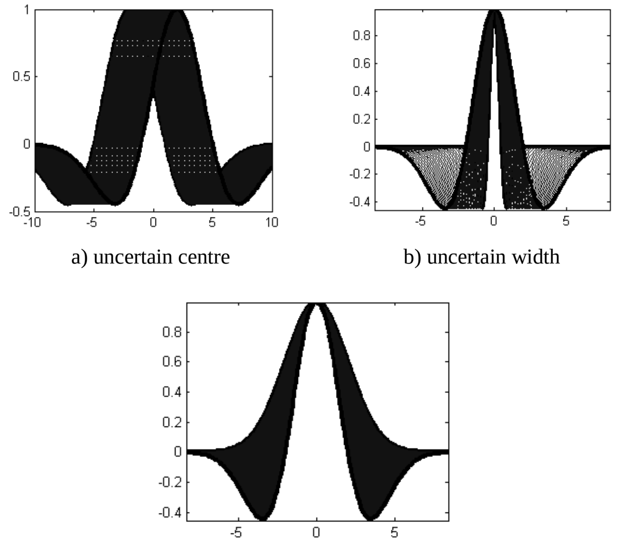


c)

Fig. 3. Adaptive wavelet function.

Here it should be noted that if the learning process of wavelet neuron of T. Yamakawa is the tuning of the synaptic weights w_{jl}, then the learning process of adaptive wavelet neuron consists of the tuning not only synaptic weights but also centres c_{jl}, widths σ_{jl} and shape parameters α_{jl} of wavelet functions. However, if synaptic weights w_{jl} can be tuned using the second-order optimisation algorithms, such as a recurrent least squares method, then the optimisation of the operation speed in the gradient learning algorithms of c_{jl}, σ_{jl}, α_{jl} is significantly difficult.

To overcome this difficulty, we can set the boundary of possible changes of adaptive wavelet function parameters $\underline{c}_{jl} \le c_{jl} \le \bar{c}_{jl}$, $\underline{\sigma}_{jl} \le \sigma_{jl} \le \bar{\sigma}_{jl}$, $\underline{\alpha}_{jl} \le \alpha_{jl} \le \bar{\alpha}_{jl}$ and introduce the type-2 fuzzy wavelet membership functions. These functions form the type-2 fuzzy wavelet neuron (T2FWN) and are shown in Fig. 4.



a) uncertain centre

b) uncertain width

c) uncertain function shape parameter

Fig. 4. Type-2 fuzzy-wavelet membership function with different type uncertainties.

T2FWN consists of two AWNs, where the signal $\tilde{x}(k)$ is fed to the inputs of it. Thus, one AWN uses low boundary values $\underline{c}_{jl}$, $\underline{\sigma}_{jl}$, $\underline{\alpha}_{jl}$, while the other AWN uses high boundary values $\bar{c}_{jl}$, $\bar{\sigma}_{jl}$, $\bar{\alpha}_{jl}$. It is important to notice that neurons $\underline{AWN}$, $\overline{AWN}$ are trained independently of each other based on common reference signal $y(k)$.

As a result, the transformation that is implemented by $\underline{AWN}$ can be written in the form

$$\underline{y}^{w}(\tilde{x}(k)) = \sum_{l=1}^{h}\sum_{j=1}^{g} \underline{w}_{jl}(k-1)\underline{\tilde{\varphi}}_{jl}(\tilde{x}_l(k)) ,$$

and $\overline{AWN}$ -

$$\bar{y}^{w}(\tilde{x}(k)) = \sum_{l=1}^{h}\sum_{j=1}^{g} \bar{w}_{jl}(k-1)\bar{\tilde{\varphi}}_{jl}(\tilde{x}_l(k))$$

where

$$\underline{\tilde{\varphi}}_{jl}(\tilde{x}_l(k)) = \left(1-\underline{\alpha}_{jl}\frac{(\tilde{x}_l(k)-\underline{c}_{jl})^2}{\underline{\sigma}_{jl}^2}\right)\exp\left(-\frac{(\tilde{x}_l(k)-\underline{c}_{jl})^2}{2\underline{\sigma}_{jl}^2}\right),$$

$$\bar{\tilde{\varphi}}_{jl}(\tilde{x}_l(k)) = \left(1-\bar{\alpha}_{jl}\frac{(\tilde{x}_l(k)-\bar{c}_{jl})^2}{\bar{\sigma}_{jl}^2}\right)\exp\left(-\frac{(\tilde{x}_l(k)-\bar{c}_{jl})^2}{2\bar{\sigma}_{jl}^2}\right).$$

In the type-reduction block, the signals $\underline{y}^{w}(\tilde{x}(k))$ and $\bar{y}^{w}(\tilde{x}(k))$ are united in the simplest way and form the output signal of T2FWN

$$f(\tilde{x}(k)) = c(k)\bar{y}^{w}(\tilde{x}(k)) + (1-c(k))\,\underline{y}^{w}(\tilde{x}(k)) ,$$

where $c(k)$ is the tuning parameter that defines closeness of signals $\underline{y}^{w}(\tilde{x}(k))$ and $\bar{y}^{w}(\tilde{x}(k))$ to reference signal $y(k)$.

Finally, the forth (output) layer of system that consists of elementary sum and division blocks implements the defuzzification in the form

$$\hat{y}(k) = \frac{f(\tilde{x}(k))}{\sum_{l=1}^{h}\tilde{x}_l(k)} = c(k)\frac{\bar{y}^{w}(\tilde{x}(k))}{\sum_{l=1}^{h}\tilde{x}_l(k)} + (1-c(k))\frac{\underline{y}^{w}(\tilde{x}(k))}{\sum_{l=1}^{h}\tilde{x}_l(k)} =$$

$$= c(k)\frac{\sum_{l=1}^{h}\sum_{j=1}^{g}\bar{w}_{jl}(k-1)\bar{\tilde{\varphi}}_{jl}(\tilde{x}_l(k))}{\sum_{l=1}^{h}\tilde{x}_l(k)} +$$

$$+(1-c(k))\frac{\sum_{l=1}^{h}\sum_{j=1}^{g}\underline{w}_{jl}(k-1)\underline{\tilde{\varphi}}_{jl}(\tilde{x}_l(k))}{\sum_{l=1}^{h}\tilde{x}_l(k)} =$$

$$= c(k)\sum_{l=1}^{h}\sum_{j=1}^{g}\overline{w}_{jl}(k-1)\frac{\overline{\tilde{\varphi}}_{jl}(\tilde{x}_l(k))}{\sum\limits_{l=1}^{h}\tilde{x}_l(k)} +$$

$$+ (1-c(k))\sum_{l=1}^{h}\sum_{j=1}^{g}\underline{w}_{jl}(k-1)\frac{\underline{\tilde{\varphi}}_{jl}(\tilde{x}_l(k))}{\sum\limits_{l=1}^{h}\tilde{x}_l(k)} =$$

$$= c(k)\sum_{l=1}^{h}\sum_{j=1}^{g}\overline{w}_{jl}(k-1)\overline{\tilde{\psi}}_{jl}(\tilde{x}_l(k)) +$$

$$+ (1-c(k))\sum_{l=1}^{h}\sum_{j=1}^{g}\underline{w}_{jl}(k-1)\underline{\tilde{\psi}}_{jl}(\tilde{x}_l(k)) =$$

$$= c(k)\overline{w}^T(k-1)\overline{\tilde{\psi}}(\tilde{x}(k)) + (1-c(k))\,\underline{w}^T(k-1)\,\underline{\tilde{\psi}}(\tilde{x}(k))$$

where

$$\underline{w}(k-1) = (\underline{w}_{11}(k-1), \underline{w}_{21}(k-1), \ldots, \underline{w}_{g1}(k-1), \underline{w}_{12}(k-1), \ldots,$$

$$\underline{w}_{jl}(k-1), \ldots, \underline{w}_{gh}(k-1))^T, \;\; \underline{\tilde{\psi}}(\tilde{x}(k)) = (\underline{\tilde{\psi}}_{11}(\tilde{x}(k)), \underline{\tilde{\psi}}_{21}(\tilde{x}(k)),$$

$$\ldots, \underline{\tilde{\psi}}_{g1}(\tilde{x}(k)), \underline{\tilde{\psi}}_{12}(\tilde{x}(k)), \ldots, \underline{\tilde{\psi}}_{jl}(\tilde{x}(k)), \ldots, \underline{\tilde{\psi}}_{gh}(\tilde{x}(k))^T,$$

$$\overline{w}(k-1) = \left(\overline{w}_{11}(k-1), \overline{w}_{21}(k-1), \ldots, \overline{w}_{jl}(k-1), \ldots, \overline{w}_{gh}(k-1)\right)^T,$$

$$\overline{\tilde{\psi}}(\tilde{x}(k)) = \left(\overline{\tilde{\psi}}_{11}(\tilde{x}(k)), \overline{\tilde{\psi}}_{21}(\tilde{x}(k)), \ldots, \overline{\tilde{\psi}}_{jl}(\tilde{x}(k)), \ldots, \overline{\tilde{\psi}}_{gh}(\tilde{x}(k))\right)^T.$$

III. ADAPTIVE LEARNING OF HGAT2FWNN

The learning process of the proposed system of computational intelligence is the tuning of synaptic weight vectors $\underline{w}(k)$ and $\overline{w}(k)$ of the neurons $\underline{AWN}$ and $\overline{AWN}$ and the scalar parameter $c(k)$ in the type-reduction block.

Since the output signals of wavelet neurons depend linearly on the synaptic weights, for their settings it is possible to use the exponentially weighted recursive least squares method, which is, in fact, the second-order optimisation procedure

$$\begin{cases} \underline{w}(k) = \underline{w}(k-1) + \dfrac{\underline{P}(k-1)\left(y(k) - \underline{w}^T(k)\underline{\tilde{\psi}}(\tilde{x}(k))\right)}{\beta + \underline{\tilde{\psi}}^T(\tilde{x}(k))\,\underline{P}(k-1)\underline{\tilde{\psi}}(\tilde{x}(k))}\underline{\tilde{\psi}}(\tilde{x}(k)), \\ \underline{P}(k) = \dfrac{1}{\beta}\left(\underline{P}(k-1) - \dfrac{\underline{P}(k-1)\underline{\tilde{\psi}}(\tilde{x}(k))\underline{\tilde{\psi}}^T(\tilde{x}(k))\,\underline{P}(k-1)}{\beta + \underline{\tilde{\psi}}^T(\tilde{x}(k))\,\underline{P}(k-1)\underline{\tilde{\psi}}(\tilde{x}(k))}\right), \end{cases}$$

$$\begin{cases} \overline{w}(k) = \overline{w}(k-1) + \dfrac{\overline{P}(k-1)\left(y(k) - \overline{w}^T(k)\overline{\tilde{\psi}}(\tilde{x}(k))\right)}{\beta + \overline{\tilde{\psi}}^T(\tilde{x}(k))\overline{P}(k-1)\overline{\tilde{\psi}}(\tilde{x}(k))}\overline{\tilde{\psi}}(\tilde{x}(k))), \\ \overline{P}(k) = \dfrac{1}{\beta}\left(\overline{P}(k-1) - \dfrac{\overline{P}(k-1)\overline{\tilde{\psi}}(\tilde{x}(k))\overline{\tilde{\psi}}^T(\tilde{x}(k))\overline{P}(k-1))}{\beta + \overline{\tilde{\psi}}^T(\tilde{x}(k))\overline{P}(k-1)\overline{\tilde{\psi}}(\tilde{x}(k))}\right) \end{cases}$$

where $0 < \beta \le 1$ – the forgetting factor.

To calculate the parameter $c(k)$ we can introduce an optimal adaptive procedure.

For off-line learning, we can write the learning error in the form

$$e(k) = y(k) - \hat{y}(k) = y(k) - c\overline{y}^w(k) - (1-c)\,\underline{y}^w(k) =$$

$$= y(k) - c\overline{y}^w(k) - \underline{y}^w(k) + c\,\underline{y}^w(k)$$

and the global learning criterion for the type-reduction block in the form

$$\begin{aligned} E(k) &= \sum_k \left(y(k) - \hat{y}(k)\right)^2 = \\ &= \sum_k \left(y(k) - c\overline{y}^w - (1-c)\,\underline{y}^w(k) \right)^2 = \\ &= \sum_k \left(y(k) - \underline{y}^w(k) \right)^2 + \\ &+ 2\sum_k \left(c(y(k) - \underline{y}^w(k))(\underline{y}^w(k) - \overline{y}^w(k)) \right) + \\ &+ \sum_k c^2 \left(\underline{y}^w(k) - \overline{y}^w(k) \right)^2. \end{aligned} \tag{1}$$

Using criterion (1), we can provide optimisation in the form

$$\frac{\partial E(k)}{\partial c} = 2\sum_k \left((y(k) - \underline{y}^w(k))(\underline{y}^w(k) - \overline{y}^w(k)) \right) +$$

$$+ 2\sum_k c\left(\underline{y}^w(k) - \overline{y}^w(k) \right)^2 = 0$$

and the expression for calculating parameter c can be written as follows

$$c = -\frac{\sum\limits_k \left((y(k) - \underline{y}^w(k))(\underline{y}^w(k) - \overline{y}^w(k)) \right)}{\sum\limits_k \left(\underline{y}^w(k) - \overline{y}^w(k) \right)^2}. \tag{2}$$

Using the global learning criterion (1), we can obtain the optimal adaptive procedure for tuning parameter c.

Based on (2), we can write the value of parameter c for k and $k+1$ instant time

$$c(k) = -\frac{\sum\limits_{i=1}^{k} \left((y(i) - \underline{y}^w(i))(\underline{y}^w(i) - \overline{y}^w(i)) \right)}{\sum\limits_{i=1}^{k} \left(\underline{y}^w(i) - \overline{y}^w(i) \right)^2} \tag{3}$$

and

$$c(k+1)=-\frac{\sum_{i=1}^{k+1}\left((y(i)-\underline{y}^{w}(i))(\underline{y}^{w}(i)-\bar{y}^{w}(i))\right)}{\sum_{i=1}^{k+1}\left(\underline{y}^{w}(i)-\bar{y}^{w}(i)\right)^{2}}. \quad (4)$$

After a series of simple transformations, we can obtain the optimal adaptive procedure of type reduction model

$$c(k+1)=-\frac{\sum_{i=1}^{k}\left((y(i)-\underline{y}^{w}(i))(\underline{y}^{w}(i)-\bar{y}^{w}(i))\right)}{\sum_{i=1}^{k}\left(\underline{y}^{w}(i)-\bar{y}^{w}(i)\right)^{2}+\left(\underline{y}^{w}(k+1)-\bar{y}^{w}(k+1)\right)^{2}}-$$
$$-\frac{\left((y(k+1)-\underline{y}^{w}(k+1))(\underline{y}^{w}(k+1)-\bar{y}^{w}(k+1))\right)}{\sum_{i=1}^{k}\left(\underline{y}^{w}(i)-\bar{y}^{w}(i)\right)^{2}+\left(\underline{y}^{w}(k+1)-\bar{y}^{w}(k+1)\right)^{2}}. \quad (5)$$

Considering that $\sum_{i=1}^{k}\left((y(i)-\underline{y}^{w}(i))(\underline{y}^{w}(i)-\bar{y}^{w}(i))\right)=$ $=-c(k)\sum_{i=1}^{k}\left(\underline{y}^{w}(i)-\bar{y}^{w}(i)\right)^{2}$ we can rewrite (5) in the form

$$c(k+1)=\frac{c(k)\sum_{i=1}^{k}\left(\underline{y}^{w}(i)-\bar{y}^{w}(i)\right)^{2}}{\sum_{i=1}^{k}\left(\underline{y}^{w}(i)-\bar{y}^{w}(i)\right)^{2}+\left(\underline{y}^{w}(k+1)-\bar{y}^{w}(k+1)\right)^{2}}-$$
$$-\frac{\left((y(k+1)-\underline{y}^{w}(k+1))(\underline{y}^{w}(k+1)-\bar{y}^{w}(k+1))\right)}{\sum_{i=1}^{k}\left(\underline{y}^{w}(i)-\bar{y}^{w}(i)\right)^{2}+\left(\underline{y}^{w}(k+1)-\bar{y}^{w}(k+1)\right)^{2}}. \quad (6)$$

Moreover, after introducing some change to parameter notations in (6), we can obtain the final form of adaptive learning procedure [28]:

$$\begin{cases}\gamma(k)=\gamma(k-1)+\left(\underline{y}^{w}(k)-\bar{y}^{w}(k)\right)^{2},\\ c(k)=c(k-1)\dfrac{\gamma(k-1)}{\gamma(k)}+\dfrac{\left(y(k)-\bar{y}^{w}(k)\right)\bar{y}^{w}(k)\left(\underline{y}^{w}(k)-\bar{y}^{w}(k)\right)}{\gamma(k)}.\end{cases} \quad (7)$$

It is clear that the condition of optimal learning is the closeness of signals $\underline{y}^{w}(k)$ and $\bar{y}^{w}(k)$, "subtending" of type-2 fuzzy wavelet membership function to common "Mexican Hat" wavelet and approximating of parameter $c(k)$ to the value 0.5.

IV. Experimental Results

To demonstrate the efficiency of the proposed hybrid generalised additive type-2 fuzzy-wavelet-neural network and its learning algorithm, some experiments were performed: prediction of hourly energy consumption time series in one of German federal lands [33].

The input number of the proposed hybrid generalised additive type-2 fuzzy-wavelet-neural network was $n=7$. The initial values of the synaptic weights were assumed to be zero. The mean squared error (MSE) and mean absolute percentage error (MAPE) were used as criteria for the quality of emulation.

Fig. 5 shows the results of energy time series prediction. The two curves, representing the actual (dot line) and forecasting (solid line) values, are almost indistinguishable.

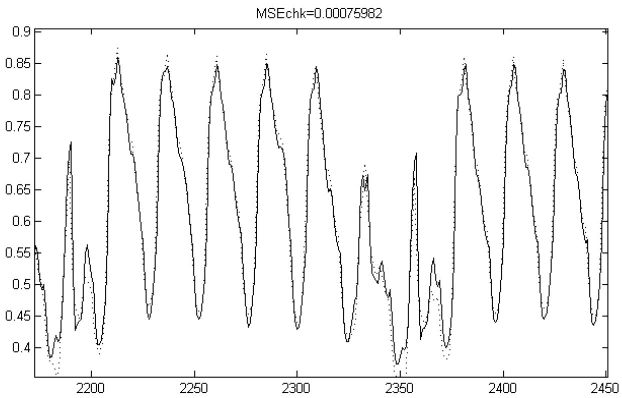


Fig. 5. Results of prediction of energy consumption time series.

Table I shows the comparative analysis of prediction process based on hybrid generalised additive type-2 fuzzy-wavelet-neural network, hybrid generalised additive wavelet-neural network, wavelet neuron and ANFIS.

TABLE I

The Comparison Analysis of Energy Consumption Time Series Prediction

Neural Network	MSE	MAPE
Hybrid generalised additive type-2 fuzzy-wavelet-neural network	0.00075	3.7 %
Hybrid generalised additive wavelet-neural network	0.0017	5.06 %
Wavelet neuron	0.0039	8.7 %
ANFIS	0.0072	12.3 %

Thus, as it can be seen from experimental results, the proposed type-2 wavelet-neuro-fuzzy systems with the learning algorithm have the best quality of prediction.

V. Conclusion

In this paper, the hybrid generalised additive type-2 fuzzy-wavelet-neural network – HGAT2FWNN – has been proposed. This system combines advantages of neuro-fuzzy system Takagi-Sugeno-Kang, wavelet neural networks, type-2 fuzzy logic and generalised additive models of Hastie-Tibshirani. The proposed hybrid system is characterised by computation simplicity (especially the type-reduction block), improving approximation and extrapolation properties as well as high speed of learning process. HGAT2FWNN can be used to solve a wide class of tasks in Dynamic Data Mining, which are related to on-line processing (prediction, emulation, segmentation, on-line fault detection etc.) of non-stationary stochastic and chaotic signals corrupted by disturbances.

References

[1] L. Rutkowski, *Computational Intelligence: Methods and Techniques.* Berlin: Springer-Verlag, 2008. http://dx.doi.org/10.1007/978-3-540-76288-1

[2] K.-L. Du and M.N.S. Swamy, *Neural Networks and Statistical Learning.* London: Springer-Verlag, 2014. http://dx.doi.org/10.1007/978-1-4471-5571-3

[3] C.L. Mumfold and L.C. Jain, *Computational Intelligence Collaboration, Fusion and Emergence*. Berlin: Springer-Verlag, 2009.

[4] S. Haykin, Neural Networks. *A Comprehensive Foundation Upper Saddle River.* N.J.: Prentice Hall, 1999.

[5] O. Nelles, *Nonlinear System Identification*. Berlin: Springer, 2001.

[6] J.-S. R. Jang, C. T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence.* N.J.: Prentice Hall, 1997.

[7] R.A. Aliev and B.G. Guirimov, *Type-2 Fuzzy Neural Networks and Their Applications*. Switzerland: Springer; 2014.

[8] A.K. Alexandridis and A.D. Zapranis, *Wavelet Neural Networks: With Applications in Financial Engineering, Chaos, and Classification*. John Wiley & Sons, 2014. ttp://dx.doi.org/10.1002/9781118596272

[9] R. Abiyev and O. Kaynak, "Fuzzy wavelet neural networks for identification and control of dynamic plants – a novel structure and a comparative study," *IEEE Trans. on Industrial Electronics*, vol. 2, no. 55, pp. 3133–3140, Aug. 2008. http://dx.doi.org/10.1109/TIE.2008.924018

[10] Ye. Bodyanskiy and O. Vynokurova, "Hybrid adaptive wavelet-neuro-fuzzy system for chaotic time series identification," *Information Science*, no. 220, pp. 170–179, Jan. 2013. http://dx.doi.org/10.1016/j.ins.2012.07.044

[11] E. Lughofer, *Evolving Fuzzy Systems - Methodologies, Advanced Concepts and Applications*. Springer, 2011. http://dx.doi.org/10.1007/978-3-642-18087-3

[12] C. Aggarwal, *Data Streams: Models and Algorithms (Advances in Database Systems)*. Springer, 2007. http://dx.doi.org/10.1007/978-0-387-47534-9

[13] A. Bifet *Adaptive Stream Mining: Pattern Learning and Mining from Evolving Data Streams*. Amsterdam: IOS Press, 2010.

[14] N.N. Karnik and J.M. Mendel, "Application of type-2 fuzzy logic systems to forecasting of time-series," *Information Sciences*, vol. 120, pp. 89–111, Nov. 1999. http://dx.doi.org/10.1016/S0020-0255(99)00067-5

[15] J. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Upper-Saddle River, NJ: Prentice Hall, 2001.

[16] R. John and S. Coupland, "Type-2 fuzzy logic: A historical view," *IEEE Computational Intelligence Magazine*, vol. 2, no. 1, pp. 57–62, Feb. 2007. http://dx.doi.org/10.1109/MCI.2007.357194

[17] M. Singh, S. Srivastava, M. Hanmandlu and J.R.P. Gupta, "Type-2 fuzzy wavelet networks (T2FWN) for system identification using fuzzy differential and Lyapunov stability algorithm," *Applied Soft Computing*, vol. 9, no. 3, pp. 977–989, Jun. 2009. http://dx.doi.org/10.1016/j.asoc.2008.03.017

[18] R.H. Abiyev, O. Kaynak, T. Alshanableh and F Mamedov, "A type-2 neuro-fuzzy system based on clustering and gradient techniques applied to system identification and channel equalization," *Applied Soft Computing*, vol. 11, no. 1, pp. 1396–1406, Jan. 2011. http://dx.doi.org/10.1016/j.asoc.2010.04.011

[19] T. Hastie and R. Tibshirani, *Generalized Additive Models.* Chapman and Hall/CRC, 1990.

[20] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2013.

[21] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. on Systems, Man and Cybernetics*, vol 15, no. 1, pp. 116–132, Jan-Feb. 1985. http://dx.doi.org/10.1109/TSMC.1985.6313399

[22] M. Sugeno and G.T. Kang, "Structure identification of fuzzy model," *Fuzzy Sets and Systems*, vol. 28, pp. 15–33, 1988. http://dx.doi.org/10.1016/0165-0114(88)90113-3

[23] H. Takagi and I. Hayashi, "NN-driven fuzzy reasoning," *Int. J. of Approximate Reasoning*, vol. 5, no. 3, pp. 191–212, May 1991. http://dx.doi.org/10.1016/0888-613X(91)90008-A

[24] S. Mitaim and B. Kosko, "What is the best shape for a fuzzy set in function approximation?" in *Proc. 5th IEEE Int. Conf. on Fuzzy Systems "Fuzz-96"*, vol. 2, pp. 1237–1213, 1996. http://dx.doi.org/10.1109/fuzzy.1996.552354

[25] S. Mitaim and B. Kosko, "Adaptive joint fuzzy sets for function approximation," in *Proc. Int. Conf. on Neural Networks "ICNN-97"*, pp. 537–542, 1997. http://dx.doi.org/10.1109/icnn.1997.611726

[26] Ye. Bodyanskiy, N. Lamonova, I. Pliss, and O. Vynokurova, "An adaptive learning algorithm for a wavelet neural network," *Expert Systems*, vol. 22, no. 5, pp. 235–240, 2005. http://dx.doi.org/10.1111/j.1468-0394.2005.00314.x

[27] Ye. Bodyanskiy and O. Vynokurova "Hybrid type-2 wavelet-neuro-fuzzy network for businesses process prediction", *Business Informatics*, no. 21, pp. 9–21, Feb. 2011.

[28] Ye. Bodyanskiy, I. Pliss and O. Vynokurova, "Type-2 Fuzzy-Wavelet-Neuron for Solving Data Mining Problems," in *Proc. East West Fuzzy Colloquium 2012, 19-th Zittau Fuzzy Colloquium.* Zittau / Goerlitz: HS, pp. 96–103, 2012.

[29] Ye. Bodyanskiy, O. Kharchenko and O. Vynokurova, "Least squares support vector machine based on wavelet-neuron," *Information Technology and Management Science*, vol. 17, pp. 19–24, 2014. http://dx.doi.org/10.1515/itms-2014-0002

[30] T. Yamakawa, "A novel nonlinear synapse neuron model guaranteeing a global minimum – Wavelet neuron," in *Proc. 28-th IEEE Int. Symp. on Multiple-Valued Logic*. Fukuoka, Japan: IEEE Comp. Soc., pp. 335–336, 1998.

[31] T. Yamakawa, E. Uchino, T. Miki and H. Kusanagi, "A neo-fuzzy neuron and its applications to system identification and prediction of the system behaviour," in *Proc. 2-nd Int. Conf. on Fuzzy Logic and Neural Networks "IIZUKA-92"*, Iizuka, Japan, pp. 477–483, 1992.

[32] E. Uchino and T. Yamakawa, "Soft computing based signal prediction, restoration and filtering," in *Intelligent Hybrid Systems: Fuzzy Logic, Neural Networks and Genetic Algorithms*, Da Ruan Eds., Boston: Kluwer Academic Publisher, 1997, pp. 331–349. http://dx.doi.org/10.1007/978-1-4615-6191-0_14

[33] P. Otto, Ye. Bodyanskiy, V. Kolodyazhniy, "A new learning algorithm for a forecasting neuro-fuzzy network," *Integrated Computer-Aided Engineering*, 10, N4, Amsterdam: IOS Press, pp. 399–409, Apr. 2003.

Yevgeniy Bodyanskiy. In 1971 he graduated with honour from Kharkiv National University of Radio Electronics. In 1980 he defended the Ph.D. thesis. In 1984 he received an academic title of Senior Researcher. In 1990 he defended the Doctor Thesis (*Dr. habil. sc. ing.*). In 1994 he received an academic title of Professor. His major field of research is evolving hybrid systems of computational intelligence.

Since 1974 he has been working at Kharkiv National University of Radio Electronics. In 1974–1976 he was a Researcher; in 1977–1983 he was a Senior Researcher; in 1986–1991 – the Head of the Control Systems Research Laboratory; in 1991–1992 – a Fellow Researcher; since 1992 he has been a Professor of Artificial Intelligence Department at KhNURE, Head of the Control Systems Research Laboratory KhNURE. He has more than 600 scientific publications, including 40 inventions and 12 monographs. Research interests include hybrid systems of computational intelligence: adaptive, neuro-, wavelet-, neo-fuzzy-, real-time systems, including problems connected with control, identification, forecasting, clustering, diagnostics, fault detection in technical, economical, medical and ecological objects.

He is the IEEE senior member, member of 4 scientific and 7 editorial boards.

Contact data: Office 511, Lenin av., 14, Kharkiv, 61166, Ukraine.
E-mail: yevgeniy.bodyanskiy@nure.ua

Olena Vynokurova. In 2002 she graduated with honour from Kharkiv National University of Radio Electronics. In 2002–2005 she was a Doctoral Student at the Artificial Intelligence Department. In 2005 she defended the Ph.D. thesis. In 2007 she received an academic title of Senior Researcher. In 2012 she defended the Doctor Thesis (*Dr. habil. sc. ing.*). Her major field of research is evolving hybrid neuro-fuzzy systems for dynamic data mining.

Since 2002 she has been working at Kharkiv National University of Radio Electronics. In 2002–2005 she was a Researcher; in 2005–2010 – a Senior Researcher; in 2010–2014 – a Leading Researcher and since 2014 she has been a Leading Researcher of Control Systems Research Laboratory; since 2013 she has been a Professor of the Department of Information Technology Security at Kharkiv National University of Radio Electronics. She has more than 130 scientific publications, including 4 monographs. Research interests include evolving hybrid systems of computational intelligence: wavelet neural networks, hybrid wavelet neuro-fuzzy systems, identification, forecasting, clustering, diagnostics, fault detection in technical, economical, medical and ecological objects.

Contact data: Office 517, Lenin av., 14, Kharkiv, 61166, Ukraine.
E-mail: olena.vynokurova@nure.ua

Pliss Iryna. In 1970 she graduated with honour from Kharkiv National University of Radio Electronics. In 1979 she defended the Ph.D. thesis. In 1983 she received an academic title of Senior Researcher. Her major field of research is evolving hybrid systems of computational intelligence.

She is a Leading Researcher of the Control Systems Research Laboratory at Kharkiv National University of Radio Electronics. She has more than 155 scientific publications. Research interests include hybrid systems of computational intelligence: adaptive, neuro-, fuzzy- real-time systems, including problems connected with control, identification, forecasting, clustering, diagnostics, fault detection in technical, economical, medical and ecological objects.

Contact data: Office 511, Lenin av., 14, Kharkiv, 61166, Ukraine.
E-mail: iryna.pliss@nure.ua

Tatarinova Yuliia. In 2014 graduated with honour from Kharkiv National University of Radio Electronics. She is a Doctoral Student at Kharkiv National University of Radio Electronics. Her major field of research is hybrid neuro-fuzzy systems for data mining problems, cryptography.

She has more than 10 scientific publications. Research interests include evolving hybrid systems of computational intelligence: wavelet neural networks, hybrid wavelet neuro-fuzzy systems, forecasting, clustering in cryptography.

Contact data: Office 517, Lenin av., 14, Kharkiv, 61166, Ukraine.
E-mail: yullia.tatarinova@gmail.com