Models of neural networks with fuzzy activation functions

A T Nguyen[1] **and A M Korikov**[1,2]

[1] Tomsk Polytechnic University, 30, Lenina ave., Tomsk, 634050, Russia
[2] Tomsk State University of Control Systems and Radioelectronics, 40, Lenina ave., Tomsk, 634050, Russia

E-mail: nguyenanhtu@tpu.ru

Abstract. This paper investigates the application of a new form of neuron activation functions that are based on the fuzzy membership functions derived from the theory of fuzzy systems. On the basis of the results regarding neuron models with fuzzy activation functions, we created the models of fuzzy-neural networks. These fuzzy-neural network models differ from conventional networks that employ the fuzzy inference systems using the methods of neural networks. While conventional fuzzy-neural networks belong to the first type, fuzzy-neural networks proposed here are defined as the second-type models. The simulation results show that the proposed second-type model can successfully solve the problem of the property prediction for time – dependent signals. Neural networks with fuzzy impulse activation functions can be widely applied in many fields of science, technology and mechanical engineering to solve the problems of classification, prediction, approximation, etc.

1. Introduction
Artificial neural networks (NNs) are a computational approach that was developed based on the way the brain solves problems. To model biological neural systems, the elements of artificial NNs (neuron model) imitate the properties and functioning of biological neurons [1]. In recent years, we have seen increasing interest in NNs, as well as their successful applications in various fields, such as business, medicine, science, mechanical engineering, geology, etc. NNs have also been used to solve the problems of forecasting, classification, management, etc. In this paper, we propose the new model of fuzzy-neural networks (FNNs), in which the fuzzy membership functions (MFs) are used as activation functions.

2. Modeling of membership functions as the activation functions
In previous studies, the activation functions of NNs usually take the forms of a unit step function, a linear threshold or sigmoid function [2]. Moreover, impulse and Gaussian distribution functions [3] can be used for activation functions in an effort to simplify the structures of NNs. The results of the above-mentioned studies have revealed the complexity of activation function selections for the development of NNs. To this end, it is necessary to study the validity for a new model that utilizes MFs from the fuzzy system theory [4, 5] as NN activation functions.

   According to literature [6], the fuzzy number is a convex, normalized fuzzy set which membership function is at least segmentally continuous and has the functional value at precisely one of the x values. This point x is referred to as the mean value of the fuzzy number.

   The support of a fuzzy set is a crisp set and is defined as follows:

$$\mathrm{supp}\left(\tilde{A}\right)=\left[S_L,S_R\right]:\mu_A\left(x\right)>0, \forall x \in \left[S_L,S_R\right].$$

Similarly, the core (kernel) of a fuzzy set is a crisp set:

$$\ker\left(\tilde{A}\right)=\left[K_L,K_R\right]:\mu_A\left(x\right)=1, \forall x \in \left[K_L,\mathrm{K}_R\right].$$

Thus, a fuzzy number is defined by four specific points:

$$\tilde{A}=\left\langle S_L,K_L,K_R,S_R\right\rangle. \tag{1}$$

However, the fuzzy number will be only completely determined by equation (1) if we know the type of MF. When creating a computer application, we use the MF LR-type, which is defined by the following expression [7]:

$$\mu_A\left(x\right)=\begin{cases} f_L\left(x\right), & x \in \left[S_L,K_L\right); \\ 1, & x \in \left[K_L,K_R\right]; \\ f_R\left(x\right), & x \in \left(K_R,S_R\right]; \\ 0, & x \notin \left[S_L,S_R\right]. \end{cases} \tag{2}$$

Equation (2) offers a large variety of MF selections, among which the triangular, trapezoidal and piecewise continuous polynomials are the most commonly-used.

In this paper, we propose a new form of MFs using triangular fuzzy number $N=\langle A, B, C\rangle$:

$$\mu\left(x\right)=f_L\left(x\right)\cdot H\left(x-A\right)\cdot H\left(B-x\right)+f_R\left(x\right)\cdot H\left(x-B\right)\cdot H\left(C-x\right), \tag{3}$$

where $f_L\left(x\right)$, $f_R\left(x\right)$ respectively denote the left and right parts of the MF, which is given by a second-order polynomial, $H\left(x\right)$ is the Heaviside unit function [4].

If $f_L\left(x\right)$, $f_R\left(x\right)$ are second-order polynomials and $f_L{}'\left(x\right)=0$, $f_R{}'\left(x\right)=0$ at the characteristic points of the fuzzy values, the MF given by equation (2) could be modified by one of the four following expressions:

$$\begin{cases} f_L{}'\left(A\right)=0; \\ f_R{}'\left(C\right)=0. \end{cases} (4); \quad \begin{cases} f_L{}'\left(B\right)=0; \\ f_R{}'\left(B\right)=0. \end{cases} (5); \quad \begin{cases} f_L{}'\left(A\right)=0; \\ f_R{}'\left(B\right)=0. \end{cases} (6); \quad \begin{cases} f_L{}'\left(B\right)=0; \\ f_R{}'\left(C\right)=0. \end{cases} (7)$$

The shapes of MFs using triangular fuzzy numbers based on the conditions given by equation (4) - equation (7) are shown in figure 1
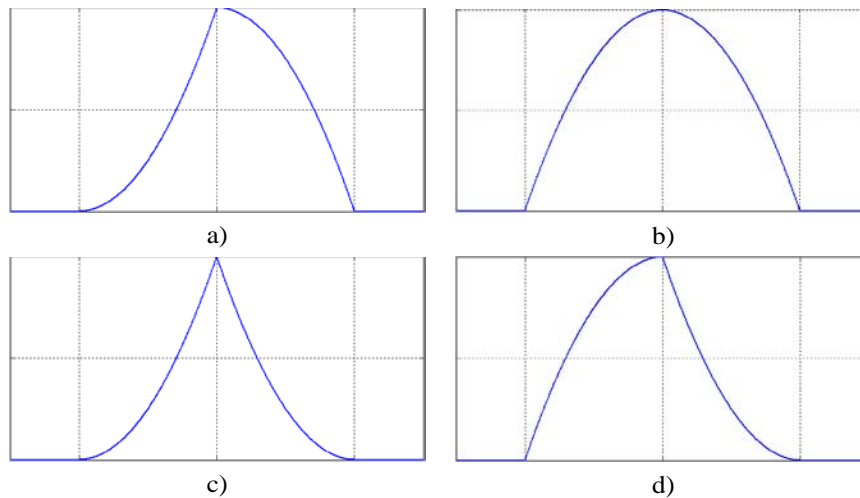

Figure 1. MFs with additional conditions: a) Conditions by equation (4); b) Conditions by equation (5); c) Conditions by equation (6); d) Condition by equation (7).

In next sections, we will investigate the feasibility of fuzzy neuron models and FNNs using these MFs as activation functions.

3. Models of fuzzy neurons

According to [8], a FNN is a clear neural network with a feed-forward signal based on a multi-layer architecture using AND – OR neurons

An AND–neuron is a neuron, in which the multiplication operation of weight w and input x is modeled by conorm $S(w,x)$, and the addition operation of weights is expressed by norm $T(w,x)$.

An OR–neuron is a neuron, in which the multiplication operation of weight w and input x is modelled by norm $T(w,x)$, and the addition operation of weights is given by conorm $S(w,x)$.

The concepts of T–norm and S–conorm were discussed in [8, 9]. T–norm and S–conorm are functions with special properties. They are the actual functions of two variables that are defined in the interval of $[0, 1] \times [0, 1]$, and the values of these functions are in the range of $[0, 1]$. The models of AND–fuzzy neurons and OR–fuzzy neurons are shown in figure 2.
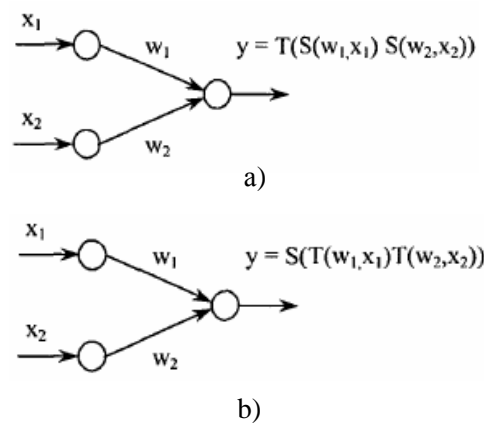

Figure 2. Models of fuzzy neurons:
a) AND–neuron; b) OR–neuron

The FNN, which is mentioned above, is termed the first-type FNN. The present paper proposes the second-type FNN model, whose activation functions have the forms of the MFs given by figure 1.

We can observe that in the first-type FNN model, the fuzzy relationships between neurons are similar to those between the elements of NNs, i.e., the first-type FNNs utilize fuzzy inference systems by NN methods. For the second-type FNN model, fuzziness is an attribute of neurons. While developing the structural representations for FNNs, it is found that it is necessary to study the third type of FNN, which is the combination of the first and second types.

4. Application of second-type FNNs to determine the fundamental frequency of signals with white noise

For many technical applications, it is necessary to identify several important characteristics, such as the frequency and the phase of a fundamental signal that can be interfered by some types of noise. This noise is usually modeled by white noise, which has a constant spectral power density distributed over the whole frequency domain.

In this paper, we investigate the feasibility of the second-type model of FNNs for the problem of determining the fundamental frequency of a time-dependent signal with noise. The methods of so-called 'theory of experiment' [10] are used to process the data of the experiment.

A signal used in this study is given as the combination of a sinusoidal signal and a white noise function:

$$y = a_1 \times \sin(2\pi f t) + a_2 \times r \tag{8}$$

where a_1 is an amplitude of the fundamental sinusoidal signal, a_2 is a constant related to the magnitude of the noise, r denotes a random function, which values are within the interval of [0,1], f is the frequency of the fundamental sinusoidal signal.

The duration from 0 to 0.005 s is studied with 500 time steps, and fundamental frequency f is set between 1 and 2 Hz. Therefore, a set of discrete values of function $y(t)$ corresponding to the 500 time steps is used as an input to our FNN. A second-type FNN is applied to predict frequency f. Similar problems may arise in many applications of FNN, especially in the mechanical engineering.

A multilayer unidirectional network is used to create our FNN. The network consists of an input layer, a hidden layer and an output layer. The hidden layer comprises 10 neurons with fuzzy activation functions of types a) or d). A linear activation function is employed for the output layer. The program selects 70% of the input values for the training process; while the rest 30% are used for the validation process. The training process is based on the Levenberg-Marquardt algorithm.
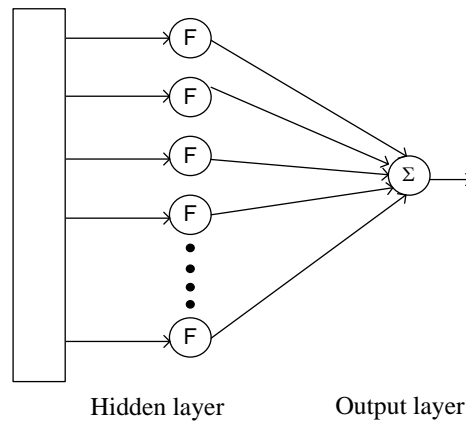

Figure 3. 'Network structure

Table 1. Test results

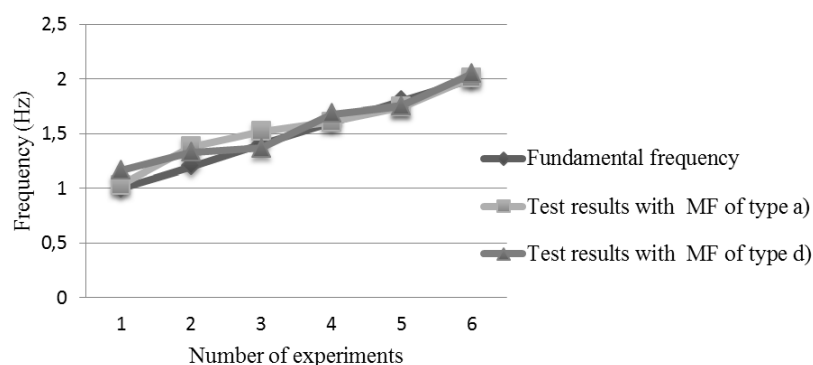| Fundamental frequency f (Hz)	Test results with MF of type a)	Test results with MF of type d)
1	1.030	1.165
1.2	1.379	1.332
1.4	1.516	1.373
1.6	1.606	1.678
1.8	1.739	1.758
2	2.012	2.046

Figure 4. Test results

5. Conclusions

From the test results it follows that the model of second-type fuzzy neurons and fuzzy neural networks can successfully solve the problems of predicting the properties of time–dependent signals. Therefore, neural networks with fuzzy impulse activation functions may be widely used in many fields of science, technology and mechanical engineering for some problems regarding classification, prediction, approximation, etc. However, these models require significant computing resources while dealing with the time series prediction problems. According to Moore's law [11], the number of elements of a single chip doubles every 18 months. Hence, with the fast development of state-of-the-art computers, this problem can be handled soon. For the next study, we will focus on solving the problem regarding computational cost.

References
[1]   Wasserman F 1989 *Neurocomputing technics: the theory and practice* (New York: Van Nostrand Reinhold)
[2]   Galushkin A I 2010 *Neural Networks: basic theory* (Moscow: Hotline – Telecom)
[3]   Aksenov S V and Novoseltsev V B 2006 *The organization and the use of neural networks (methods and technologies)* (Tomsk: Publishing house of the YTL) 128 p
[4]   Yefremov A A 2015 *J. Proc. of Tomsk State Univ. of Control Syst. and Radioelectr.* **2 (36)** 136
[5]   Yefremov A A 2013 *Youth and Modern Information Technology* M **75** pp 129–131
[6]   Beer M and Moller B 2004 *Fuzzy Randomness: Uncertainty in Civil Engineering and Computational Mechanics* (New York: Springer – Verlag)
[7]   Yefremov A A and Korikov A M 2011 *Siberian Journal of Science* **1(1)** 340–343
[8]   Yarushkina N G 2004 *Lectures on Neuroinformatics* (Moscow: Moscow Engineering Physics Institute) part 1 p 151
[9]   Gupta M M and Qi J 1991 *Fuzzy Sets and Systems* **40** 431–450
[10]   Korikov A M 2015 *J. Proc. of Tomsk State Univ. of Control Syst. and Radioelectr.* **2(36)** 148
[11]   Thackray A, Brock D and Jones R 2015 *Moore's Law: The Life of Gordon Moore, Silicon Valley's Quiet Revolutionary* (New York: Basic Books)