IDENTIFICATION NONLINEAR BY FUZZY WAVELET NEURAL NETWORKS

José Medeiros de Araújo Júnior, medeiros_eng@yahoo.com.br
Electrical Engineering Department
Federal University of Piauí, Brazil

Fábio Meneghetti Ugulino de Araújo, meneghet@dca.ufrn.br
Computation and Automation Engineering Department
Federal University of Rio Grande do Norte, Brazil

Abstract. Fuzzy systems and neural networks, when used together, form a great tool for dealing with complex non-linearities and processes poorly defined. However, Neuro-Fuzzy system may require a large number of rules. Willing to deal with this problem, Wavelet Neural-Networks (WNN) has been used in recent years. The combination of these techniques, results in a structure called Fuzzy Wavelet Neural-Network (FWNN). In this work, FWNN was used to identify a real nonlinear system with a considerable complexity grade. This system consists of a multiple sessions tank, a pressure sensor, a water pump, a data acquisition board and a power module. The different sections of the tank, the pressure at the exit of the hose that carries water to the tank, the noise of the sensor, and other imperfections, add more nonlinearities to the system. The objective is to identify this system, which by its complex and nonlinear dynamic making it an ideal system to test the efficiency of FWNN as an identifier. The FWNN was implemented in C++, having as inputs the water level in the tank and the voltage applied to the pump. The training was conducted using a backpropagation algorithm, following by an adequate validation procedure. The results were satisfactory, and demonstrated the ability of the technique FWNN in identifying complex nonlinear systems.

Keywords: Wavelets, Nonlinear Identification, Neuro-Fuzzy, Fuzzy Wavelet Neural Network

1. INTRODUCTION

There has been a growing interest in nonlinear dynamical models (Yilmaz and Oysal, 2010, Banakar and Azeem, 2006, Kara and Eker, 2003). This is due to the fact that real dynamical systems are nonlinear and despite of linear approximations are sufficient, in many applications, the linear models are not satisfactory and nonlinear representations should be used. There are several nonlinear representations that may be used in system identification, such as ARX models, polynomial models, radial basis functions, artificial neural networks, fuzzy systems, neuro-fuzzy hybrid systems, etc.

Fuzzy logic is a proper tool for dealing with complex processes, poorly defined and non-linear (Wang et al, 2006). The rules base of the fuzzy system is usually created using the knowledge of human experts. However, for some complex processes, this knowledge may not be enough and several approaches have been proposed to generate these rules. In the last decade, the use of neural networks has been applied for this purpose (Xie et al, 2009, Uppal et al, 2003, Rodrigues et al, 2009). The union of these two techniques is called neuro-fuzzy system and ANFIS (Adaptive-Network-Based Fuzzy Inference System) proposed by Jang (1997) is one of the best known models.

In order to improve the efficiency of such system, the use of wavelets has gained importance (Banakar and Azeem, 2006, Ahmad et al, 2009). These functions perform signal analysis both in time domain as in frequency domain. Thus, wavelets provide an enhanced ability to track and to analyze signals locally and globally, making it possible to study small and large details. This characteristic leads the neuro-fuzzy system to require fewer neurons to identify systems of greater complexity and also to get a more accurate convergence.

In short, wavelets have properties that allow an optimal analysis of complex nonlinear signals, fuzzy logic reduces the complexity of decision making, and the learning capability of neural networks increases the accuracy of the model. Thus, the combination of the three mentioned techniques results in the structure called Fuzzy Wavelet Neuro-Network (FWNN), which several researchers (Lee et al, 2003, Davani et al, 2010) have used to model nonlinear systems characterized by uncertainty.

This study aims to investigate a process of identifying a nonlinear dynamic system using hybrid neuro-fuzzy structure wavelet. The proposed technique uses a similar architecture to ANFIS. However, the polynomial Takagi-Sugeno is replaced by Wavelets Neural-Networks (network feedforward of three layers, where the activation functions are wavelets). This approach was applied to a plant-level with multiple sections containing several nonlinearities.

The system consists basically of a tank, a pressure sensor, a water pump and a data acquisition board. A pump, submerged, is used to feed the tank, pumping water from an external reservoir into the tank. The control signal of this pump is a tension sent by the data acquisition board, iUSBDAQ - U120816 from Hytek Automation, which makes the communication between the computer and the tank. A power module was used to amplify this tension, since the tension provided by the acquisition board is not sufficient to operate the pump.
This system has nonlinearities that must be considered by FWNN. For example, several distinct sections of the system mean that even for constant inputs, the performance level is not linear. Moreover, the pressure outlet of the hose, which lies at the bottom of the tank that causes a change in the pump flow depending on the level of the tank. There are also non-linearities into the hole at the base of the tank whose flow varies depending on the level. There is also the read noise, the delay and inaccuracy of the sensor. In this way, so that the identifier is successful it is necessary to consider these nonlinearities.

For the application of the technique were created three programs. The first program is responsible for collecting data for both the validation procedure of the system as the training of the FWNN structure. Through the acquisition board, voltage values specified in this program are applied to the water pump. The second program implements the FWNN and is responsible for initiating and optimizing the structure used for identification, i.e., do the training. The third program uses the structure optimized by the second program performs validation of the system with a new data set obtained by the first program.

2. NEURO-FUZZY AND WAVELETS NEURAL-NETWORKS

Several techniques have been used to identify nonlinear systems (Palit, 2005), and neural-networks and fuzzy systems are among the most powerful ones. In order to improve the identification of nonlinear models, this technique may be associated with wavelet transform. This function, also called ondeleta, is able to decompose and describe other functions in the frequency domain, so that consider these functions in different frequency ranges. In the next section, it is presented the main concepts of WNN and neuro-fuzzy system ANFIS used in this study.

2.1. Adaptive Neuro-Fuzzy Inference Systems - ANFIS

The parameters adjustment process of Fuzzy systems is complex and can be time consuming, becoming worse with the increasing complexity of the fuzzy structure used. Moreover, these systems do not include in its initial conception, the possibility of automatic adjustment of these parameters, which could reduces the design work, favoring learning and adaptation in real time.

On the other hand, Artificial Neural Networks (ANN) allows the automatic adjust to different operating conditions in real time, due to its learning and generalization ability. Given these facts, the combination of these techniques is a way to create an intelligent system, as shown in (Rodrigues, 2006). This fusion has attracted several researchers in various scientific and engineering areas due to the need solve real world problems (Abraham, 2001). One of the most popular ways of combining fuzzy systems and neural networks is called ANFIS (Jang, 1997).

An ANFIS is considered to be an adaptive network very similar to neural networks, however, has no synaptic weights, but nonadaptive nodes. The ANFIS has the advantage of not needing to create complex mathematical models. ANFIS can learn from the sample data such as the input-output sets from the system and can adapt parameters inside its network (Reddy and Mohanta, 2008). The ANFIS has two types of parameters: the premise parameters specify the shape of the membership function; the consequent parameters are the polynomial coefficients of the linear equations in the fuzzy rule (Phichaisawat et al., 2009).

In the ANFIS, the crisp values are fuzzificated to be fuzzy values (membership grade in fuzzy sets). The membership grade values are multiplied by the weights of each rule outputs. An output is defuzzified to be crisp value by a weighted average of rule outputs. The ANFIS adjusts all the membership function parameters from a given input/output data set using the back-propagation gradient descent and the least squares methods for nonlinear and linear parameters, respectively (Reddy and Mohanta, 2008).

2.2. Wavelets Neural-Networks

Wavelet functions have attracted interest in several areas, both from the view of theoretical and practical. The fast advance of wavelet theory is basically due to its interdisciplinary origins, which has attracted researchers from different fields of knowledge, and by some simple concepts that will be discussed (Candid, 2008).

Wavelets are mathematical functions that analyze data into different frequency components, and study each component with a resolution matching its scale. These functions were developed independently from the application, being studied largely in the fields of mathematics, engineering, quantum physics, and recently have been included in a wide range of applications: computer vision and human, radar and sonar, computer graphics, filter banks, image compression and analysis of medical signs, such as electrocardiograms, mammography, electroencephalograms, etc. (Oliveira, 2003).
These functions will be used as activation function in hidden layer of a feedforward neural network, replacing the sigmoid functions (Abiyev and Kaynak, 2008). This structure, called WNN, has good Generalization Ability, can approximate complex functions and can be easily trained than other networks, such as multilayer perceptrons and radial based.

2.3. Fuzzy Wavelets Neural Networks - FWNN

The use of wavelet function is to improve the computational efficiency of a neuro-fuzzy system, in particular the ANFIS. This is due to the fact that the consequent part of ANFIS in case of modeling complex nonlinear processes may require a high number of rules in order to accurately find a more accurate model to be identified. Increasing the number of rules leads to an increase in the number of neurons in the hidden layer of the network, so that its convergence is slower and often without obtain an accurate model. Thus, the combination of ANFIS and WNN results in a system (FWNN) with good generalization and approximation ability of complex functions that can be implemented and trained in a simple way.

In this work the WNN will be used in place of polynomials Sugeno located in the consequent of the fuzzy rules of the ANFIS. The wavelet used in the WNN is the Mexican Hat, represented by equation 3 and shown en Fig. 1.

\[
\Psi_j(z) = \sum_{i=1}^{m} |a_{i,j}|^{-1/2} (1 - z_{i,j}^2) e^{-\frac{z_{i,j}^2}{2}}
\]  

(3)

where \( i \) is the number of neurons in input layer and \( j \) is the number of neurons in consequent layer, which corresponds to the number of wavelets. The term \( |a_{i,j}|^{-1/2} \) is a normalization factor which ensures that the energy of the wavelet is independent of \( a \) and \( b \), shown in Eq. 4. The variable \( z_{i,j} \) is used to simplify the equation 4 and is defined as follows:

\[
z_{i,j} = \frac{x_i - b_{\psi,i}}{a_{\psi,i}}
\]  

(4)

The parameter \( b \) indicates that the function \( \psi_j(x) \) was shifted to a distance equivalent to \( b \), being the translation parameter. Since varying the parameter \( a \) it is possible to stretch or compress the form of the function, so it is possible to cover different scales large. This parameter is known as the dilation parameter.

![Figure 1. Mexican Hat Wavelet](image)

The rules of the FWNN are described as follows:

If \( x_1 \) is \( A_{11} \) and \( x_2 \) is \( A_{12} \) ... and \( x_m \) is \( A_{1m} \), then:

\[
y_1 = \sum_{i=1}^{m} w_{i1} \left( 1 - z_{i1}^2 \right) e^{-\frac{z_{i1}^2}{2}}
\]  

(5)

If \( x_1 \) is \( A_{21} \) and \( x_2 \) is \( A_{22} \) ... and \( x_m \) is \( A_{2m} \), then:
\[ y_2 = \sum_{i=1}^{m} w_{ij} \left(1 - z_{i2}^2\right) e^{-\frac{z_{i2}^2}{2}} \]  

(6)

If \( x_1 \) is \( A'_{a1} \) and \( x_2 \) is \( A'_{a2} \) ... and \( x_m \) is \( A'_{am} \) then:

\[ y_n = \sum_{i=1}^{m} w_{in} \left(1 - z_{i2}^2\right) e^{-\frac{z_{i2}^2}{2}} \]  

(7)

where \( x_i \) are the inputs, \( y_i \) are the outputs of the rules, \( A'_{ij} \) is a function of relevance to a given rule.

### 3. SYSTEM IMPLEMENTATION

The architecture of the FWNN is based on a basic structure of a Takagi-Sugeno fuzzy system, which can be interpreted as a five layer neural network (ANFIS), with the difference that the resulting functions are WNN. The FWNN implemented has two inputs, with 3 membership functions each. This results in 9 rules, and consequently, 9 wavelets for each input. Figure 2 shows the FWNN, however, for better viewing, it was summarized with the use of 2 membership functions for each entry.

![Figure 2. Structure of FWNN](image)

Note that the weights of the synapses were defined as 1 (one) and the adjustable variables (antecedents and consequences) are distributed in the functions of neurons, represented by square nodes. The data stream can be analyzed layer by layer, as shown below:

Layer 1: This layer consists of adaptive nodes that generate degrees of membership based on input signals, using some parameterized membership function such as the bell shaped function. (This function was chosen because it is continuously differentiable, which allows the application of backpropagation to update its parameters). The outputs of the neurons in this layer are called fuzzy entries, defined according to Equation 8.

\[ \mu_{j,i}(x) = \frac{1}{1 + \frac{|x - c_j|}{a_j}}^{2\beta_j} \]  

(8)

where \( j \) refers to the neuron of layer 01 and \( i \) indicates the input of the system.
Layer 2: Nodes in this layer are fixed nodes, which represent the activation weight of each rule. The output of each node is the result of an AND operation (or MIN) of all layer input signals, as can be seen in Equations 9 to 12.

\[
f(\mu_A, \mu_B) = w_i = \mu_A(x) \cdot \mu_B(y)
\]

\[
f(\mu_A, \mu_B) = w_2 = \mu_A(x) \cdot \mu_B(y)
\]

\[
f(\mu_A, \mu_B) = w_3 = \mu_A(x) \cdot \mu_B(y)
\]

\[
f(\mu_A, \mu_B) = w_4 = \mu_A(x) \cdot \mu_B(y)
\]

Layer 3: The outputs of layer 3 represent the normalization of the weights of activation of rules, i.e., normalization of output neurons of the previous layer, as shown in the following equation. By being more didactic to visualize the FWNN structure and be simpler to implement the backpropagation algorithm, the normalization of fuzzy rules was realized in this layer.

\[
g(\cdot) = w_i = \frac{w_j}{w_1 + w_2 + w_3 + w_4}, i = 1, 2, 3, 4
\]

Layer 4: The nodes of this layer are the activation functions of the WNN. The function used was the Mexican Hat, was chosen because it is continuously differentiable. Thus, from the entries, normalized rules and adjustable parameters of these functions, each node will have its output calculated according to Equation 14.

\[
h(\cdot) = w^m \sum_{i=1}^{m} w_{\psi_{i,j}} |a_{i,j}|^{-1/2} (1 - z_{i,j}^2) e^{-z_{i,j}^2 / 2}
\]

Layer 5: In this layer, all outputs of neurons in layer 4 are summed, resulting in the output of FWNN.

\[
s(\cdot) = \sum h(\cdot)
\]

From network presented, can clearly identify the neurons that require learning are present in layers 1 and 4, because in the first layer are located in the membership functions and layer 4, the weights w and the parameters of the wavelets, Which sets the Implications of the rules.

Tuning parameters of a system in FWNN model can be obtained through techniques such as adaptive backpropagation algorithm, which use Eq. 16. In this equation, the variable x is the learning rate and its initialization value in this work is \(0.0001\). A detailed explanation of this equation together with the equations used to find the value of the variable \(\delta_{k,i,j}\) can be found at (Rodrigues, 2006). The equations to find the value of the \(\frac{\partial f_{k,i}}{\partial a}\) are shown below. Equations 17, 18 and 19 are utilized to adjust the wavelets parameters.

\[
\Delta a_{k,i,j} = -\eta \sum_{i=1}^{L} \delta_{k,i,j} \frac{\partial f_{k,i}}{\partial a}
\]

\[
\frac{\partial h(\cdot)}{\partial w_{\psi}} = w \cdot \psi
\]
\[ \frac{\partial h(\cdot)}{\partial a_\psi} = \frac{-w_\psi \cdot (3.5z^2 - z^4 - 0.5) \cdot e^{-\frac{z^2}{2}}}{\sqrt{a_\psi^3}} \]  

(18)

\[ \frac{\partial h(\cdot)}{\partial b_\psi} = \frac{-w_\psi \cdot (3z - z^3) \cdot e^{-\frac{z^2}{2}}}{\sqrt{a_\psi^3}} \]  

(19)

Equations 20, 21 and 22 are utilized to adjust the membership functions parameters.

\[ \frac{\partial \mu}{\partial a} = \frac{2b}{a} \mu(1 - \mu) \]  

(20)

\[ \frac{\partial \mu}{\partial b} = \left\{ \begin{array}{ll} -2\ln\left[ \frac{x-c}{a} \right] \mu(1 - \mu) & \text{se } x \neq c \\ 0 & \text{se } x = c \end{array} \right. \]  

(21)

\[ \frac{\partial \mu}{\partial c} = \left\{ \begin{array}{ll} -\frac{2b}{x-c} \mu(1 - \mu) & \text{se } x \neq c \\ 0 & \text{se } x = c \end{array} \right. \]  

(22)

Despite the use of a training algorithm that performs the adjustment of parameters, a good initialization of the system enables it to obtain a better convergence of training and a better model identified. Several methods are proposed by the literature to initialize the wavelet parameters, such as orthogonal least square procedure and the clustering method (Abiyev and Kaynak, 2008). In this paper, the parameters were initialized by a heuristic technique, which uses the domain of input space (Awad et al., 2004). This technique says that the parameters should be calculated according to the equations:

\[ b = 0.5 \cdot (x_{\text{min}} + x_{\text{max}}) \]  

(23)

\[ a = 0.2 \cdot (x_{\text{max}} - x_{\text{min}}) \]  

(24)

where \( x_{\text{min}} \) and \( x_{\text{max}} \) represent the minimum and maximum values of the signal being identified.

All programs developed to implement the technique were written in C++ Builder compiler using C++, version 5.0. This language is best suited to implement techniques that will be implemented in real systems. In general, for the application of the technique were created three programs with different characteristics. The first program is responsible for the communication with the real system and collects data for validation and training procedures to find the linear models and rules for the use of backpropagation algorithm with the complete system. The sampling period used has a value of \( T = 0.1 \) seconds and was obtained based on practical experiences with the plant. This program has performed, to facilitate interpretation of the user, converting the data read from the pressure sensor in volts to centimeters.

The second program is responsible for initiating and optimizing the structure used for identification. The program takes the number of inputs, the membership functions number of each input and the number of wavelets (ie, the number of neurons of the hidden layer of WNN). Then the program initializes the membership functions and the wavelets, and
reads a file containing the training set. Using this file, the program performs the backpropagation algorithm for optimization of membership functions and wavelets. Throughout the learning procedure the best setting for the membership functions and the wavelets is stored in a data file. The third program utilizes the enhanced data and performs validation of the system with a new data set obtained by the first program.

The aim of applied the FWNN technique in the multiple sessions tank will be to identify the behavior of the tank when the water pump is subjected to several levels of tension. To achieve this goal, two entries were used in the structure of the identifier, the current level of the tank (current state) and the value of tension applied in the pump.

4. EXPERIMENT

The system consists basically of a 53x30x10cm tank made of acrylic and manufactured by hand, a pressure sensor, a water pump, a data acquisition board and a power module. Figure 3 illustrates the idealized multiple session tank. Because it was built by hand, this plan had some flaws in relation to the idealized theoretical model. For example, all sections should present 10 cm high, but showed a 13 cm already had another 8 inches tall. Some vertical edges that should be presented a few degrees of tilt visibly identifiable. Besides the acrylic tank, other components that form the system as the pressure sensor, water pump and plate data acquisition, stand out.

Installed in a hole on the tank, the pressure sensor MPX5010DP Motorola is used to inform the water level in the multiple sessions tank. Another component is the pump used to feed the tank, pumping water from an external reservoir into the tank, has a flow rate of approximately 6.5 liters per minute, when an applied voltage 12Volts DC with 2 amp of current, and a reflux of 0.9 liters per minute when the applied voltage is 0 volts.

To establish the communication between the computer and the power module that drives the pump and the sensor was used the data acquisition board iUSBDAQ - U120816 from Hytek Automation. This board has analog inputs and outputs, allowing the output voltage (0 to 2.4 volts) to activate the pump, and input voltage (proportional to the height of water level) provided by the sensor reading of the tank.

The motor drive signal from the data acquisition board has a very small current, about 20 milliamps, and needed to go through an amplifier. The power module model UPM2405 of Quanser Consulting-240 was used to amplify the current and, purely by the characteristics of the amplifier, the voltage was multiplied by 5. As the submersible pump used to work with 12 volts, then the voltage applied to the data acquisition board will be up to 2.4 volts, which multiplied by a gain of 5 to maximum tension generated by the pump. The power module was also used to power the pressure sensor with 5 volts.

The tank system has developed some nonlinearities that can be noted immediately. The various different sections of the system require, even for constant inputs, a nonlinear performance level. The pressure at the outlet of the hose, which is at the bottom of the tank, causes a change in the pump flow depending on the level of the. There are also non-linearity into the hole at the base of the tank whose flow varies depending on the level. Besides these, there is the read noise, the delay and inaccuracy of the sensor. In this way, so that the handle is successful it is necessary to consider these nonlinearities.
6. RESULTS

This tank presented in the previous section was used to prove the efficiency of FWNN. Thus, at first were collected training data obtained from a PRBS signal, which varies the voltage level applied to the pump. With this signal, the water level varied passing all sections, regardless of the levels where the transitions occur from one section to another. Thus, we obtained a group of training data for the FWNN of 1,200 samples, containing the values of voltage and liquid level. The range of the PRBS signal, which can be seen in Fig. 4, is 12 volts. With this signal, the liquid level in the tank varies from 0 to 30 cm.

![Figure 4: PRBS Signal](image)

Thus, the inputs used in the training were the previous output of the plant and the voltage applied to the water pump. Each of these inputs have three membership functions (bell-shaped type, as previously mentioned), thus forming nine rules. This number of membership functions was determined intuitively, so it had not been a high number of rules, which would result in a greater number of parameters to be adjusted. In the FWNN presented it has 63 parameters to be updated by backpropagation algorithm.

Observing Fig. 5, we can notice that the training of the FWNN was satisfactory, having learned accurately even in regions with more abrupt level changes. As a performance criterion, the mean square error (MSE) is used. After 10,000 epochs, the MSE was 0.00017409, ie, the output estimated by FWNN is almost equal to actual output, as shown in Figure 5.

![Figure 5: Training - desired output (blue line) x Real output (red line)](image)
Despite the training has been quite satisfying, the validation was performed to confirm the accuracy of the model obtained, using 400 points, which are different from those used in training. The FWNN was replenished with the output of the plant, as shown in Fig. 6.

![Figure 6. Validation Scheme](image)

The result can be seen in Fig. 7, shown below. The blue line represents the desired output, the red line represents the real output.

![Figure 7. Validation - Desired output (blue line) x Real output (red line)](image)

It is noticed that the output of FWNN (red line) was close enough to the desired output (blue line), using data different from those used in training, which confirms that the model nonlinear is accurate, able to represent the real model of the tank satisfactorily.

5. CONCLUSIONS

In this work, FWNN was used for identification nonlinear of a real system. This technique incorporates the advantages of wavelet function, neural networks and fuzzy logic, resulting in an adaptive structure with fast learning capability that can identify nonlinear dynamic models characterized with uncertainties. Thus, the mainly objective of this work was demonstrating the potential and the utility of FWNN to identify a nonlinear level tank formed by multiple sections. The algorithm backpropagation was used for adjust parameters. The results obtained were satisfactory, obtaining an accurate model and opening the possibility to study the use of another wavelet type, for instance, Morlet, besides the possibility of increasing the system order.

6. REFERENCES


Palit, A. K., Popovic, 2005, Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications (Advances in Industrial Control), Springer, 1ª edição.


7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.