Nonlinear System Identification Based on a Novel Adaptive Fuzzy Wavelet Neural Network

Maryam Salimifard*, Ali Akbar Safavi**
*School of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran, msalimifard@aut.ac.ir
**Department of Power and Control, School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran, safavi@shirazu.ac.ir

Abstract: In this paper, an optimized adaptive Fuzzy Wavelet Neural Network (FWNN) is proposed for identification of nonlinear systems. The network combines Takagi-Sugeno-Kang fuzzy neural networks with the advantages of adaptive wavelet functions. Therefore, it provides an effective nonlinear mapping which can approximate the local as well as the global behaviour of nonlinear complex systems. Furthermore, an optimized constructive learning algorithm is proposed. In this regard, all network parameters including center and variance of membership functions, dilation and translation of wavelets, and weights are assumed to be adjustable which make the network structure quite flexible. In this method, the orthogonal projection pursuit (OPP) algorithm is invoked in the structure learning phase which can generate the fuzzy rules automatically. Then, the parameters of each rule are optimized based on a nonlinear global optimization method, here, the genetic algorithm (GA). To increase the performance of the optimization scheme, a nonlinear local optimization algorithm, the Levenberg-Marquardt (LM), is also applied with the initial point from the GA. This hybrid combination produces a more accurate optimal solution which provides better performance with a fewer number of required fuzzy rules. As some advantages of this approach, the proposed network is self-implemented, and there is no need to initialize the parameters or pre-determine the number of fuzzy rules. Finally, a nonlinear case study is provided which shows the efficiency of the proposed identification approach.

Keywords: Nonlinear System identification, fuzzy wavelet neural networks, orthogonal projection pursuit, genetic algorithm, Levenberg-Marquardt.

1. Introduction

Many real dynamic systems may have some uncertainties which cannot be well described by the deterministic models. In such cases, the use of alternative soft computing methods can be useful [1-2]. In this regard, Fuzzy Wavelet Neural Networks (FWNNs) have drawn a great deal of attention and become one of the powerful tools in modelling, identification, and control of nonlinear systems [3-8]. These networks combine the advantages of fuzzy logic to deal with the complexity and uncertainty in nonlinear systems with the benefits of localized wavelet basis functions and impressive learning ability of neural networks. In the literatures, several FWNNs structures with different learning algorithms have been presented for the purpose of function approximation, system identification, and controller design [9-15].

Optimum definition of the fuzzy If-Then rules is one of the most challenging issues in fuzzy inference systems. In the TSK-type fuzzy models, the consequent part of the fuzzy model is described by a linear combination of input variables and a constant term. Therefore, the TSK-type system cannot represent complex nonlinear systems with certain number of rules with a desired accuracy. These models are only able to describe the global behaviour of a system, and their convergence is generally slow [3]. To overcome this problem, the linear function has been replaced with the wavelet functions [3,4,6,10,12-14]. Unlike the linear function, wavelets have nonlinear local characteristics, and are able to provide an efficient nonlinear mapping. However, wavelets may not properly describe the overall behavior of complex nonlinear systems with reasonably number of fuzzy rules.

Accordingly, this paper proposes the combination of linear functions with adaptive wavelets in the FWNN structure which would greatly increase the network ability to describe the local as well as the global behavior of a nonlinear complex system with satisfactory performance.

There are different methods for training various types of neural networks, neuro-fuzzy networks, wavelet networks, and their combinations. These methods can be divided into two general approaches. The first one uses the local optimization techniques such as gradient based or backpropagation methods [3,6,7,16,17]. The second one is based on global optimization algorithms including evolutionary-type genetic and particle swarm methods [12,18,19].

Although the parameters of the networks can be adjusted with the use of any of these methods, however, the network structure must be specified with trial and error. Considering the balance between the number of fuzzy rules and optimal network performance is difficult. The large number of rules requires heavy computation which would not be suitable for practical applications.
While the network performance with the fewer number of rules may not be satisfactory enough for desired performance.

In this paper, an optimized learning algorithm is proposed which consists of two stages. First, the structure learning phase which automatically generates the fuzzy rules. Second, the parameter learning phase which optimizes the network parameters. For the structure learning phase, an orthogonal projection pursuit (OPP) algorithm is invoked. The projection pursuit algorithm and its extension to orthogonal type were previously proposed for regression problems [20–22], and have not been yet employed in FWNN literature. Using this constructive algorithm to determine the FWNN structure, the fuzzy rules are generated automatically during network training, and there is no need to specify the number of fuzzy rules with trial and errors. Once a fuzzy rule is generated, the corresponding parameters including center and variance of Gaussian membership functions, dilation and translation factors of wavelets and weight, are optimized by an efficient hybrid optimization approach. This algorithm combines the advantages of nonlinear local and global optimization schemes to achieve a more precise optimal solution.

Though the classical local optimization methods are not able to search all the solution space and they may fall in local minima, but they have reasonable performance near the minimum and converge to it with a desirable gradient. On the other hand, the global optimization approaches can find the global optimum, however, their convergence near the optimal solution may not be as fast and precise as local optimization methods. Therefore, a hybrid combination of these two approaches can provide a powerful tool for fast and accurate search of the entire space. In this case, first, a global optimization method, the genetic algorithm (GA), is invoked to determine an approximate of the global optimal solution. Then, to achieve a favorable performance, a local optimization approach, the Levenberg-Marquardt (LM) algorithm, is employed with the initial point form GA. As the results verify, the proposed learning approach yields a more accurate optimal solution with the fewer number of required fuzzy rules.

The proposed network is self-implemented, since the entire parameters are adjusted automatically using the proposed learning algorithm. Therefore, it is not necessary to predetermine the number of fuzzy rules or initialize the parameters. As a result, this method is very flexible and desirable for nonlinear black-box system identification problems. The simulation results verify the efficiency of this approach.

The remainder of this paper is organized as follows. In Section 2, the FWNN structure is described. The optimized constructive learning algorithm of FWNN is presented in Section 3. To demonstrate the effectiveness of the proposed approach, a nonlinear case study is provided in Section 4. Finally the paper concludes in Section 5.

2. The FWNN Structure

In this section, a series-parallel FWNN structure for identification of NARX (nonlinear autoregressive with exogenous inputs) model as in Fig. 1 is introduced. This network consists of six layers which are described in the followings.

![Fig. 1: The series-parallel FWNN structure](image)

First layer: Each node in this layer represents an input variable. For identification of the NARX model, the input \( x_i \) at time \( t \) is defined as:

\[
x_i(t) = \begin{cases} 
  y(t - i), & 1 \leq i \leq n_y \\
  u(t - i + n_y), & n_y + 1 \leq i \leq n = n_u + n_y
\end{cases}
\]  

(1)

Second layer: Each node in this layer corresponds to a fuzzy membership function to describe the antecedent parts of the fuzzy rules. Here, the Gaussian membership functions are used:

\[
\mu_{ij}(x_i) = \exp \left( -\frac{(x_i - m_{ij})^2}{\sigma_{ij}^2} \right)
\]

(2)

where \( i = 1, \ldots, n \) and \( j = 1, \ldots, R \) represent the network inputs and fuzzy rules, respectively. The total number of inputs and trained rules are respectively \( n \) and \( R \). Here, \( m_{ij} \) and \( \sigma_{ij} \) are the center and standard deviation of fuzzy membership function corresponding to the input \( x_i \) in the \( j \)th rule.

Third layer: The number of nodes in this layer is equal to the number of fuzzy rules. In fact, each node corresponds to the antecedent part of a fuzzy rule. The output of these nodes can be obtained by:

\[
\rho_j = \prod_{i=1}^{n} \mu_{ij}(x_i), \quad j = 1, \ldots, R
\]

(3)

where \( \rho_j \) represents the firing strength of the \( j \)th rule.

Fourth layer: This layer corresponds to the consequent part of the fuzzy rules, and the number of nodes in this layer is equal to the number of rules. To improve the capability of nonlinear mapping in TSK-type fuzzy systems, the output of the \( j \)th node in this layer is described as a combination of wavelet functions and a linear function as follows:
\[ y_j = \sum_{i=1}^{n} c_{ij} \psi(x_i; a_{ij}, b_{ij}) + \sum_{i=1}^{n} w_{ij} x_i + w_{0j} \]  

(4)

where \[ \psi(x_i; a_{ij}, b_{ij}) = \psi(a_{ij} x_i - b_{ij}) \] is an adaptive one-dimensional wavelet function with adjustable dilation \((a_{ij})\) and translation \((b_{ij})\) parameters. By means of continuous families of wavelets with adjustable parameters, the required number of wavelets in each fuzzy rule is reduced to the number of inputs. In this equation, \(c_{ij}\), \(w_{ij}\), and \(w_{0j}\) are respectively correspond to the wavelet coefficient, weight of the linear function, and the constant term.

Fifth layer: In this layer, the firing strength of each rule is multiplied by the output of that rule.

Sixth layer: In this layer, the defuzzification is done and the output of the overall FWNN is expressed as:

\[ \hat{y}(t) = \sum_{j=1}^{R} \rho_j y_j(t)/\sum_{j=1}^{R} \rho_j \]  

(5)

The proposed learning method of FWNN is described in the following section.

3. Learning Algorithm

In this paper, the orthogonal projection pursuit (OPP) algorithm is invoked in the structure learning phase of FWNN which automatically generate the fuzzy rules. In this algorithm, at first, a fuzzy rule is trained in the network. Then, the orthogonal error between the network and the real system is calculated. If this error is larger than a threshold value, a new fuzzy rule is also trained. As such, to achieve an acceptable error level, more fuzzy rules can be trained in the network. Therefore, unlike the common neuro-fuzzy networks, there is no need to determine the number of rules beforehand. The proposed learning method is described in the followings.

In the proposed network, to provide more flexibility, all parameters are assumed to be adjustable. The vectors of network parameters for the \(j\)th fuzzy rule are defined as:

\[ \alpha_j = [m_{1j}, ..., m_{nj}, \sigma_{1j}, ..., \sigma_{nj}, \alpha_{1j}, ..., \alpha_{nj}, b_{1j}, ..., b_{nj}]^T \]

\[ \theta_j = [c_{1j}, ..., c_{nj}, w_{1j}, ..., w_{nj}, w_{0j}]^T \]

Observation vector of the \(k\)th input, and the output vector for \(N\) samples are expressed as:

\[ x_k = [x_k(1), x_k(2), ..., x_k(N)]^T, k = 1, 2, ..., n \]

\[ y = [y(1), y(2), ..., y(N)]^T \]  

(6)

Where \(n\) is the total number of network’s inputs. For the \(j\)th fuzzy rule let:

\[ g(X; \theta_j) = \rho_j y_j, X = [x_1, x_2, ..., x_n], \theta_j = [\alpha_j, \theta_j] \]  

(7)

The main idea of this algorithm for adding each rule is successive approximation of arbitrary nonlinear systems based on minimizing the approximation error [20,22].

The OPP algorithm is implemented with a step-by-step process. At each step, a new fuzzy rule which minimize the error vector is generated and added to the network. In the first step of this algorithm, starting from the residual \(r_0 = y\), the first fuzzy rule \(g_1 = g(X; \theta_1)\) is generated so that \(\theta_1 = \arg \min_{\theta} ||r_0 - g(X; \theta)||^2\). Now, in the parameter learning phase, the proposed combination of GA and LM optimization techniques is applied to find the global optimal solution for \(\theta_1\). For this purpose, first the GA is invoked to search the entire space of solution. After that, the LM algorithm is employed to tune the parameters more accurately. In this case, the initial guess of iterative LM technique is obtained from GA. Once the parameters of the first rule are optimized, the orthogonal error vector \(r_1 = r_0 - \alpha_1 g_1\) corresponding to this rule can be used to generate the second rule where orthogonalization coefficient \(\alpha_1\) is defined as \(\alpha_1 = <r_0 g_1>/||g_1||^2\). Now, the parameters of the second rule are optimized so that the difference between \(r_1\) and \(g_1\) become minimized. The sum of squares error \(||r_m||^2\) in \(m\)th step (rule) can be used as a criterion for stopping the OPP algorithm. In this step, the parameters are obtained from the minimization of the following objective function:

\[ J_{m-1}(\theta) = ||r_{m-1} - g(X; \theta)||^2 \]

\[ = \sum_{t=1}^{N} [r_{m-1}(t) - g(X(t); \theta)]^2 \]  

(8)

Where \(r_m = r_{m-1} - \alpha_m g(X; \theta_m)\) and \(\alpha_m = <r_{m-1} g_m>/||g_m||^2\). The stopping criterion is defined as the error-to-signal ratio \(ESR = ||r_m||^2/||y||^2\) [22]. If the ESR is smaller than a predetermined threshold, the OPP algorithm is stopped. The sequence \(||r_m||^2\) is strictly decreasing and positive, therefore, the residual \(r_m\) form a Cauchy sequence. As such, \(r_m\) converges to zero, and then, the algorithm will converge [23,24].

4. Simulation Results

Consider the following nonlinear dynamic system [25,26]:

\[ y(t) = \frac{y(t-1)y(t-2)y(t-3)(u(t-2)(y(t-3) - 1) + u(t-1))}{1 + y^2(t-2) + y^2(t-3)} \]  

(9)

The input signal \(u(t)\) is applied as:

\[ u(t) = \begin{cases} 
\sin(\pi t/25) & 0 < t < 250 \\
+1.0 & 250 \leq t < 500 \\
-1.0 & 500 \leq t < 750 \\
0.3\sin(\pi t/25) + 0.1\sin(\pi t/32) + 0.6\sin(\pi t/10) & 750 \leq t < 1000 
\end{cases} \]

In the FWNN structure, the Mexican Hat wavelet \(\psi(x) = (1 - x^2)e^{-x^2/2}\) is employed [27]. The network’s inputs for identification of the NARX model are selected as:

\[ [y(t-1), y(t-2), y(t-3), u(t-1), u(t-2)] \]  

(10)

For structure learning phase, the orthogonal projection pursuit algorithm is used while for parameter learning
phase, the genetic algorithm, the Levenberg-Marquardt technique, and the hybrid combination of GA and LM techniques are used. For comparison, the system is also identified with fuzzy wavelet network (FWN) [12]. In this network, the antecedent part of the fuzzy rules is described by Gaussian membership functions, and the consequent part is described by the wavelet functions. All the parameters of membership functions (center and variance), wavelet parameters (dilation and translation), and the network weights are optimized by the genetic algorithm. Table I provides a comparison between these approaches where the performance criteria MSE and $I$ are defined as:

$$\text{MSE} = \frac{1}{N} \sum_{t=1}^{N} (y(t) - \hat{y}(t))^2$$

(11)

$$I = \frac{\sum_{t=1}^{N} (y(t) - \bar{y}(t))^2}{\sum_{t=1}^{N} (y(t) - \bar{y})^2}, \bar{y} = \frac{1}{N} \sum_{t=1}^{N} y(t)$$

(12)

In Fig. 2, the estimation error of network using the mentioned algorithms are compared, while for a closer vision, the interval [745,800] of error is expanded in Fig. 3. It is clear from Table I, Figs. 2 and 3 that the hybrid combination of genetic algorithm and Levenberg-Marquardt optimization techniques yields a better performance with the fewer number of fuzzy rules. Although the number of rules in the hybrid combination GA and LM is approximately equal to the number of fuzzy rules in the FWN [12], but the computation time and the performance are quite different. In fact, in the fuzzy wavelet network [12], the structure of network is determined by trial and error, and the number of fuzzy rules is set to be fixed beforehand. While in the proposed approach, by using the OPP algorithm, the number of required rules is determined during the structure learning. Moreover, since these rules are generated one-by-one, the parameters of only one rule should be optimized in the parameter learning phase while in [12], the entire network parameters related to all rules should be optimized at once. The approximation quality of the proposed FWNN using the OPP and hybrid combination of GA and LM techniques is illustrated in Fig. 4 where a good agreement between the original output and the network output is seen. For more validation, this network is also examined for other two cases. In the first case, the wavelet functions are removed from the network while in the latter case, the linear function is removed from the consequent part of fuzzy rules. In simulations, for the first case, there may be large local errors while in the latter case, the number of generated rules is more than before. Therefore, the combination of linear functions and wavelets can provide better performance in the structure of fuzzy wavelet neural networks to approximate the local as well as the global behavior of nonlinear complex systems.

<table>
<thead>
<tr>
<th>Approach</th>
<th>MSE</th>
<th>$I$</th>
<th>No. of rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPP+GA</td>
<td>0.0083</td>
<td>0.0533</td>
<td>9</td>
</tr>
<tr>
<td>OPP+LM</td>
<td>0.0192</td>
<td>0.1950</td>
<td>15</td>
</tr>
<tr>
<td>OPP+GA+LM</td>
<td>0.0010</td>
<td>0.0272</td>
<td>5</td>
</tr>
<tr>
<td>FWN [12]</td>
<td>0.0152</td>
<td>0.1858</td>
<td>6</td>
</tr>
</tbody>
</table>
5. Conclusions

This paper proposes a novel optimized fuzzy wavelet neural network for identification of nonlinear complex systems. The network structure provides a powerful nonlinear model based on the capabilities of TSK-type fuzzy models and adaptive wavelet functions. The learning algorithm for this network consists of two phases, structure determination and parameter estimation. In the structure learning phase, the orthogonal projection pursuit algorithm was invoked. Using this method, once a rule is generated in the network, the parameters are tuned by an efficient hybrid combination of genetic algorithm and Levenberg-Marquardt optimization methods. As the results verify, the proposed method yields more accurate models with fewer number of required fuzzy rules.

References