Chapter from the book *Particle Swarm Optimization*
Downloaded from: http://www.intechopen.com/books/particle_swarm_optimization
A Radial Basis Function Neural Network with Adaptive Structure via Particle Swarm Optimization

Tsung-Ying Sun, Chan-Cheng Liu, Chun-Ling Lin, Sheng-Ta Hsieh and Cheng-Sen Huang
National Dong Hwa University
Taiwan, R.O.C.

1. Introduction

Radial Basis Function neural network (RBFNN) is a combination of learning vector quantizer LVQ-I and gradient descent. RBFNN is first proposed by (Broomhead & Lowe, 1988), and their interpolation and generalization properties are thoroughly investigated in (Lowe, 1989), (Freeman & Saad, 1995). Since the mid-1980s, RBFNN has been used to apply on many applications, such as pattern classification, system identification, nonlinear function approximation, adaptive control, speech recognition, and time-series prediction, and so on. In contrast to the well-known Multilayer Perceptron (MLP) Networks, the RBF network utilizes a radial construction mechanism. MLP were trained by the error Back Propagation (BP) algorithm, since the RBFNN has a faster training procedure substantially and adopts typical two-stage training scheme, it can avoid solution to fall into local optima. A key point of RBFNN is to decide a proper number of hidden nodes. If the hidden node number of RBFNN is too small, the generated output vectors may be in low accuracy. On the contrary, it with too large number of hidden nodes may cause over-fitting for the input data, and influences global generalization performance. In conventional RBF training approach, the number of hidden node is usually decided according to the statistic properties of input data, then determine the centers and spread width for each hidden nodes by means of k-means clustering algorithm (Moddy & Darken, 1989). The drawback of this approach is that the network performance is depended on the pre-selected number of hidden nodes. If an unsuitable number is chosen, RBFNN may present a poor global generalization capability, as slow training speed, and requirement for large memory space. To solve this problem, the self-growing RBF techniques were proposed in (Karayiannis & Mi, 1997), (Zheng et al, 1999). However, the predefined parameters and local searching on solution space cause the inaccuracy of approximation from a sub-solution.

Evolutionary computation is a globally optimization technique, where the aim is to improve the ability of individual to survive. Among that, Genetic Algorithm (GA) is a parallel searching technique that mimics natural genetics and the evolutionary process. In (Back et al, 1997), they employed GA to determine the RBFNN structure so the optimal number and distribution of RBF hidden nodes can be obtained automatically. A common approach is applied GA to search for the optimal network structure among several candidates.
constructed initially by the unsupervised clustering method (Chen et al, 1999). However, its results depend on the pre-selected RBFNN structures which may not be appropriate. Another method is to fix the number of RBF nodes and adopted GA to search optimal network parameters, for example, centers and spread widths for RBF hidden nodes, and the weights connected to the output layer (Aiguo & Jiren, 1998). This method requires heavy computational cost while the number of RBF hidden nodes is too large, the dimension of each chromosome has to extend to corresponding length. It will spend too much time for training. GA based self-growing RBF network training method was proposed by (Yunfei & Zhang, 2002) to overcome the mentioned drawbacks. It searches single parameter, the cluster distance factor, which can avoid organizing a large dimension in a chromosome. It performs a fast training speed and well convergences while the GA operators (reproduction, recombination, and mutation, etc.) and fitness evaluation is properly applied. However, GA-based approaches are poorer in several aspects, as premature convergence and falling into local optima, than new evolutionary computation techniques.

The particle swarm optimization (PSO) is a novel and popular search algorithm based on the simulation of the social behavior of birds within a flock in evolutionary computation. As opposed to (Yunfei & Zhang, 2002), this paper proposes a PSO based RBFNN self-structure algorithm to overcome the drawbacks that mentioned above. PSO is a swarm intelligence method that roughly models the social behavior of swarms and has been proved to be efficient on many optimization problems in science and engineering. The social behavior of PSO allows particles to stochastically return toward previously successful regions in the search space. We propose a PSO-based approach for searching the optimal cluster distance factor to provide a suitable criterion on self-structure RBFNN training. The results of simulation experiments exhibit the rapid convergence and more better optimal solutions than other related approaches. Furthermore, it yields efficient training for constructing RBFNN.

The paper is organized as follows. Section II describes structure and the training of the RBF network. Section III describes the principle and procedures of the self-structure RBF algorithm. Section IV presents the application of a PSO to search the cluster distance factor. Section V evaluates our method for modeling nonlinear function and predicting time series by RBF Network and comparing the results with the GA-RBF Network and K-means methods. Section VI is the conclusion.

2. Radial Basis Function Neural Network

Generally, a RBFNN consists of three layers: the input layer, the RBF layer (hidden layer) and the output layer. The inputs of hidden layer are the linear combinations of scalar weights and the input vector \( \mathbf{x} = [x_1, x_2, \ldots, x_n]^T \), where the scalar weights are usually assigned unity values. Thus the whole input vector appears to each neuron in the hidden layer. The incoming vectors are mapping by the radial basis functions in each hidden node. The output layer yields a vector \( \mathbf{y} = [y_1, y_2, \ldots, y_m] \) for \( m \) outputs by linear combination of the outputs of the hidden nodes to produce the final output. Fig. 1 presents the structure of a single output RBF network; the network output can be obtained by

\[
\mathbf{y} = f(\mathbf{x}) = \sum_{i=1}^{k} w_i \phi_i(\mathbf{x})
\]
where $f(x)$ is the final output, $\phi_i(\cdot)$ denotes the radial basis function of the $i$-th hidden node, $w_i$ denotes the hidden-to-output weight corresponding to the $i$-th hidden node, and $k$ is the total number of hidden nodes.

$$f(x) = \sum_{i=1}^{k} w_i \phi_i(x)$$

where $\phi_i(x)$ is the radial basis function of the $i$-th hidden node, $w_i$ denotes the hidden-to-output weight corresponding to the $i$-th hidden node, and $k$ is the total number of hidden nodes.

A radial basis function is a multidimensional function that describes the distance between a given input vector and a pre-defined center vector. There are different types of radial basis function. A normalized Gaussian function usually used as the radial basis function, that is

$$\phi_i(x) = \exp\left(-\frac{\|x - \mu_i\|^2}{2\sigma_i^2}\right)$$

where $\mu_i$ and $\sigma_i$ denote the center and spread width of the $i$-th node, respectively.

Generally, the RBFNN training can be divided into two stages:

1. Determine the parameters of radial basis functions, i.e., Gaussian center and spread width. In general, k-means clustering method was commonly used here.
2. Determine the output weight $w$ by supervised learning method. Usually Least-Mean-Square (LMS) or Recursive Least-Square (RLS) was used.

The first stage is very crucial, since the number and location of centers in the hidden layer will influence the performance of the RBFNN directly. In the next section, the principle and procedure of self-structure RBF algorithm will be described.

### 3. Self-structure RBFNN

The hidden layer of an RBFNN acts as a receptive field operating on the input data space. The number of hidden node based on the distribution of the training data set. The proposed approach performs this task by defining a cluster distance factor, $\varepsilon$, which is the maximum distance between an input sample and a specific RBF node center and allowing the number of basis function to increase iteratively according to this factor.

The rationale of this learning is described as follows: the hidden layer starts with no hidden node and $\varepsilon$ is pre-determined by PSO to control the clusters production. The first RBF node center $\mu_i$ is set by choosing one data, $x_1$, randomly from $N_I$ input data sample. The value of Euclidean 2-norm distance between $\mu_1$ and the next input sample, $x_2$, is compared with $\varepsilon$. 

Figure 1. The structure of a RBFNN
If it is greater, a new cluster whose center location is $x_2$ is created as $\mu_2$; otherwise, the elements of $\mu_i$ are updated as

$$\mu_i(\text{new}) = \mu_i(\text{old}) + \alpha \|x_{2i} - \mu_i(\text{old})\|, i = 1, 2, \ldots, N$$ (3)

where $\mu_i$ and $x_{2i}$ are the $i$-th component of vectors $\mu_1$ and $x_2$ respectively, $\|\|$ denotes the Euclidean distance and $0 < \alpha < 1$ is the updating ratio. Thus, this procedure is carried out on the remaining training samples. The number of clusters grows or RBF nodes center self-adjust continuously until all of the samples are processed. The proposed self-structure RBFNN algorithm can be summarized as follows:

1. Assuming that there are $p$ clusters with their centers, $\mu_1, \ldots, \mu_p$, are generated from previous iterations. Taking a new input sample $x_n$ to calculate the distances with the each clusters $\|x_n - \mu_i\|$, where $i = 1, \ldots, p$.

2. The cluster whose center $\mu_q$ is $\arg \min_{\mu_i} \|x_n - \mu_i\|$ will be focused.

3. Comparing $\|x_n - \mu_q\|$ with the distance criterion parameter, $\varepsilon$. If it is greater than $\varepsilon$, then a new cluster center, $\mu_{p+1}$, is created at the position of the sample point, $x_n$.

   Otherwise the elements of $\mu_p$ are updated by (3).

4. Repeating the above steps until all of the samples are processed.

For $L$ clusters, a global spread width $\sigma$ can be derived by the average of Euclidean distance between each cluster center and its nearest neighbor as

$$\sigma = \left\langle \|\mu_i - \mu_j\| \right\rangle$$ (4)

where $\left\langle \right\rangle$ denotes the expression for the average value for $1 \leq i \leq L$, $1 \leq j \leq L$ and $i \neq j$.

In (Yunfei & Zhang, 2002), the cluster distance factor, $\varepsilon \in (0, \infty)$, is obviously a critical factor to determine input space partitioning and obtains the hidden node number and locations in RBFNN. An unduly large value of $\varepsilon$ does not reflect an enough number of cluster so it may cause a poor-generalized precision solution. On the contrary, an unduly small value of $\varepsilon$ will create redundant clusters; therefore, it may cause overlap between RBF neurons; moreover, it may lead to poor accuracy and slow convergence either. This paper proposes a PSO-based searching approach to determine the proper value of $\varepsilon$; further, the optimal structure of RBF network can be obtained. And, an objective function to evaluate the effectiveness of applying PSO is proposed. Following section will describe how to employ PSO technique to search a potential optimal value $\varepsilon$.

4. PSO-based Self-structure RBFNN

The PSO is a population based optimization technique that was proposed by Kennedy and Eberhart in 1995 (Eberhart & Kennedy, 1995), which the population is referred to as a swarm. The particles express the ability of fast convergence to local and/or global optimal position(s) over a small number of generations.
4.1 Evolution of PSO
A swarm of PSO consists of a number of particles. Each particle represents a potential solution of the optimization task. All of the particles iteratively discover the probable solution. Each particle generates a position according to the new velocity and the previous positions of the particle, and it is compared with the best position which is generated by previous particles according to the cost function. The best solution is then kept; i.e., each particle accelerates in the directions of not only the local best solution but also the global best position. If a particle discovers a new probable solution, other particles will move closer to it so as to explore the region more completely in the process (Gudise & Venayagamoorthy, 2003).

Let \( N \) denotes the swarm numbers. In general, there are three attributes, current position \( a_{ij} \), current velocity \( v_{ij} \) and past best position \( P_{bij} \), for particles in the search space to present their features. Each particle in the swarm is iteratively updated according to the aforementioned attributes assuming that the objective function \( f \) is to be minimized so that the dimension consists of \( n \) particles and the new velocity of every particle is updated by (5).

\[
v_{ij}(t + 1) = wv_{ij}(t) + c_1r_{ij}(t)[P_{bij}(t) - a_{ij}(t)] + c_2r_{2j}(t)[G_{bij}(t) - a_{ij}(t)]
\]

where \( v_{ij} \) is the velocity of the \( j \)-th particle of the \( i \)-th swarm for all \( i \in 1...N \), \( w \) is the inertia weight of velocity, \( c_1 \) and \( c_2 \) denote the acceleration coefficients, \( r_1 \) and \( r_2 \) are two uniform random values falling in the range between (0, 1), and \( t \) is the number of generations. The new position of the \( i \)-th particle is calculated as follows:

\[
a_{ij}(t + 1) = a_{ij}(t) + v_{ij}(t + 1)
\]

The past best solution of each particle is updated by:

\[
P_{bij}(t + 1) = \begin{cases} P_{bij}(t), & \text{if } f(a_{ij}(t + 1)) \geq f(P_{bij}(t)) \\ a_{ij}(t + 1), & \text{otherwise} \end{cases}
\]

The global best solution \( Gb \) will be found from all of particles during previous three steps are defined as:

\[
Gb = \underset{P_{bij}}{\arg \min} f(P_{bij}(t + 1)), \quad 1 \leq i \leq n
\]

4.2 Disturbance
Since initial particles are generated by randomly, they may not uniform enough to distribute over the solution space. Therefore, it may trap particles into local optimal solution inevitably. To avoid solution falling into the local minimal and jumping it out to find the global minimal, this paper added a mutation-like disturbance strategy into the PSO process (Sun et al, 2005). The disturbance mechanism randomly activates under a disturbance probability. While the disturbance mechanism is active, the selected particle will be randomly placed at a new position (\( \varepsilon \) value in this paper), then this particle will keep following the PSO process to search a better solution. The other non-selected particle will keep following the PSO iteration as usual and trying to find a new solution.
4.3 Objective function
For searching a suitable $\varepsilon$ value for RBFNN training, a function of root mean squared error (RMSE) which evaluates discrepancies between the sampling data output $y_n$ and the predictive output $y_n^*$ is applied. Thus, the objective function for $N_T$ sample is defined as

$$f(\varepsilon, y_n^*) = \text{RMSE}(y) = \sqrt{\frac{1}{N_T} \sum_{k=1}^{N_T} (y_n(k) - y_n^*(k))^2}$$

where $y_n^*(k)$ is the predictive output of the $k$-th sample data which is obtained by $\varepsilon$ value during training.

In the section II, the relationship between self-structure RBF network training and cluster distance factor $\varepsilon$ was discussed. If (9) can be reduced to a sufficiently small value, a suitable value of $\varepsilon$ could be obtained to train the structure of RBFNN. Thus, the predictive RBFNN output would be closed to the sampling data output.

4.4 RBFNN structure determination by PSO
In this paper, our goal is to minimize the value of $f(\varepsilon, y_n^*)$. The objective function minimized by PSO and found potential optimal solution finally. Since we only search one parameter by PSO (i.e., the cluster distance factor $\varepsilon$), the swarm number $i=1$, and defined the particle number as $1 \leq j \leq m$. In the initial state of PSO, all the particles’ positions $a_j$ (i.e., initial cluster distance factor $\varepsilon$) were set as 0.02, $v_j$ were set as 0, and the $Pb_j$ and $Gb_j$ were initialized by a random number generator in the range of [0, 1]. After particles moved by (6), each particle will find a potential solution, the new past best position would be updated by (7), and the global best position would be updated by (8). The particle would keep moving to find a better solution until it reaches the goal or meets the termination condition (Lin et al, 2005). The pseudo code of our PSO-based cluster distance factor searching approach presented in Fig. 2.

Create and initiate an N-dimension PSO: $P$

Repeat:
- Execute PSO to update $P$ by (5) and (6)
  - for each particle $i \in [1...m]$
    - if $f(\varepsilon, y_n^*) < f(Pb_j, y_n^*)$
      - then $Pb_j = \varepsilon$
    - if $f(Pb_j, y_n^*) < f(Gb_j, y_n^*)$
      - then $Gb_j = Pb_j$
  endfor

Until Termination condition is met

Figure 2. The pseudo code of PSO-based cluster distance factor searching
5. Simulations and Results

5.1 Setting of simulation
The six nonlinear functions with different complexities are tested here. These tested functions are listed as follows:

<table>
<thead>
<tr>
<th>Ex.</th>
<th>Tested function</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( y = \frac{(x-2)(2x-1)}{(1+x^2)} )</td>
<td>( x \in [-8,12] )</td>
</tr>
<tr>
<td>2</td>
<td>( y = \frac{\sin(x)}{x} )</td>
<td>( x \in [-10,10] )</td>
</tr>
<tr>
<td>3</td>
<td>( y = 1.1(1-x^2+2x^2)e^{-x^2/2} )</td>
<td>( x \in [-5,5] )</td>
</tr>
<tr>
<td>4</td>
<td>( y = 0.5 \sin\left(\frac{x}{2}\right) + \frac{(1+\cos x)}{2} )</td>
<td>( x \in [-4.5,10.8] )</td>
</tr>
<tr>
<td>5</td>
<td>( y = \sin\left(\frac{2\pi x}{5}\right) + \sin\left(\frac{2\pi x}{10}\right) )</td>
<td>( x \in [0,10] )</td>
</tr>
<tr>
<td>6</td>
<td>( y = -\sin\left(\frac{\pi x}{10}\right) + \cos\left(\frac{2\pi x}{5}\right) )</td>
<td>( x \in [-10,10] )</td>
</tr>
</tbody>
</table>

Table 1. The six tested nonlinear functions

In order to confirm the advantages of the proposed approach, the K-means algorithm (Moddy & Darken, 1989) and GA-based self-growing RBFNN training algorithm (Yunfei & Zhang, 2002) are also carried out in these tested functions. Due to (Yunfei & Zhang, 2002) adopted Simple Genetic Algorithm (SGA) which using binary coding to train RBF structure for saving computation time, but it will lose some accuracy compared to the real-valued, i.e., this method may not present the optimal solution. So we implemented it with Real-value Genetic Algorithm (RGA) to obtain accuracy results.

For every simulation, the training data set consists of 50 input-output data samples taken at random, and the testing data set includes 75 samples different from the training data set. For the definition of parameters in the proposed approach, \( w, c_1 \) and \( c_2 \) are given 0.12, 0.25 and 0.25 respectively, and the search range of \( \varepsilon \) is bounded between 0.2 and 1, the particle number is 10. For the GA-based self-growing RBFNN training algorithm the search range of \( \varepsilon \) in the input space is also in the range from 0.2 to 1, the crossover rate \( P_c \) is given 0.8, and mutation rate \( P_m \) is given 0.01, the population size is 10. For the K-means method, the optimal number of RBF neurons in the hidden layer is chosen to be 30 by experience.

5.2 Simulation results
After simulations, the RMSE of training data, RMSE of testing data, maximal error and number of hidden node will be presented in tables for each case. In these tables, the three involved algorithms are denotes as PSO-based, GA-based (Yunfei & Zhang, 2002) and K-means (Moddy & Darken, 1989). Additionally, the real data and approximated data will be shown in the same figure; meantime, the error from each approximation will be presented by figures. There three sub-figures in each figure, the results from the left sub-figure to the
right sub-figure are generated by PSO-based approach, GA-based approach and K-means approach, respectively.

**Example 1.**

<table>
<thead>
<tr>
<th></th>
<th>PSO-based</th>
<th>GA-based</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE for training data</td>
<td>0.0332</td>
<td>0.0584</td>
<td>0.1552</td>
</tr>
<tr>
<td>RMSE for testing data</td>
<td>0.0520</td>
<td>0.0786</td>
<td>0.1962</td>
</tr>
<tr>
<td>Maximal error</td>
<td>0.3852</td>
<td>0.2355</td>
<td>0.9508</td>
</tr>
<tr>
<td>Number of hidden node</td>
<td>33</td>
<td>29</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2. Comparison between the three approaches in example 1

![Curves of RBFNN output and real data in example 1](image1)

Figure 3. Curves of RBFNN output and real data in example 1. (solid-line represents the real data, dashed-line represents the output data)

**Example 2.**

<table>
<thead>
<tr>
<th></th>
<th>PSO-based</th>
<th>GA-based</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE for training data</td>
<td>0.0035</td>
<td>0.0046</td>
<td>0.0234</td>
</tr>
<tr>
<td>RMSE for testing data</td>
<td>0.0099</td>
<td>0.0113</td>
<td>0.0357</td>
</tr>
<tr>
<td>Maximal error</td>
<td>0.0460</td>
<td>0.0509</td>
<td>0.1006</td>
</tr>
<tr>
<td>Number of hidden node</td>
<td>28</td>
<td>28</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 3. Comparison between the three approaches in example 2

![The errors between the real data and approximations in example 1](image2)

Figure 4. The errors between the real data and approximations in example 1
A Radial Basis Function Neural Network with Adaptive Structure via Particle Swarm Optimization

Figure 5. Curves of RBFNN output and real data in example 2. (solid-line represents the real data, dashed-line represents the output data)

Figure 6. The errors between the real data and approximations in example 2

Example 3.

<table>
<thead>
<tr>
<th></th>
<th>PSO-based</th>
<th>GA-based</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE for training data</td>
<td>0.0056</td>
<td>0.0173</td>
<td>0.1271</td>
</tr>
<tr>
<td>RMSE for testing data</td>
<td>0.0057</td>
<td>0.0188</td>
<td>0.0803</td>
</tr>
<tr>
<td>Maximal error</td>
<td>0.0134</td>
<td>0.0432</td>
<td>0.2441</td>
</tr>
<tr>
<td>Number of hidden node</td>
<td>24</td>
<td>22</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 4. Comparison between the three approaches in example 3

Figure 7. Curves of RBFNN output and real data in example 3. (solid-line represents the real data, dashed-line represents the output data)
Example 4.

<table>
<thead>
<tr>
<th></th>
<th>PSO-based</th>
<th>GA-based</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE for training data</td>
<td>0.0079</td>
<td>0.0112</td>
<td>0.0337</td>
</tr>
<tr>
<td>RMSE for testing data</td>
<td>0.0192</td>
<td>0.0292</td>
<td>0.0740</td>
</tr>
<tr>
<td>Maximal error</td>
<td>0.1217</td>
<td>0.0939</td>
<td>0.1854</td>
</tr>
<tr>
<td>Number of hidden node</td>
<td>19</td>
<td>31</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 5. Comparison between the three approaches in example 4
Example 5.

<table>
<thead>
<tr>
<th></th>
<th>PSO-based</th>
<th>GA-based</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE for training data</td>
<td>0.0460</td>
<td>0.0519</td>
<td>0.0637</td>
</tr>
<tr>
<td>RMSE for testing data</td>
<td>0.0546</td>
<td>0.0564</td>
<td>0.0867</td>
</tr>
<tr>
<td>Maximal error</td>
<td>0.3056</td>
<td>0.3236</td>
<td>0.2208</td>
</tr>
<tr>
<td>Number of hidden node</td>
<td>20</td>
<td>21</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 6. Comparison between the three approaches in example 5

![Curves of RBFNN output and real data in example 5.](image)

Example 6.

<table>
<thead>
<tr>
<th></th>
<th>PSO-based</th>
<th>GA-based</th>
<th>K-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE for training data</td>
<td>0.0079</td>
<td>0.0092</td>
<td>0.0690</td>
</tr>
<tr>
<td>RMSE for testing data</td>
<td>0.0439</td>
<td>0.0509</td>
<td>0.0859</td>
</tr>
<tr>
<td>Maximal error</td>
<td>0.2159</td>
<td>0.2486</td>
<td>0.1855</td>
</tr>
<tr>
<td>Number of hidden node</td>
<td>29</td>
<td>27</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 7. Comparison between the three approaches in example 6

![The errors between the real data and approximations in example 5](image)
5.3 Discussion
In the simulation results in tables, PSO-based approach has lower RMSE for training data and testing data. It means that over fitting does not happen in the proposed approach. From figures of the curves of RBFNN output and real data, the approximated curves by PSO-based approach is closer to the real data than these by others. From figures of the approximated errors, it could be shown that PSO-based approach results small error in most of sample, whereas the K-means approach has largest error.

We know that RBFNN needs different number of hidden node and cluster radius for different complexities. K-means approach usually performs a larger error because it is not able to decide a suitable number of hidden node. Though GA-based approach decides a suitable number of hidden node, its cluster radius is not good enough to classify whole data. The proposed approach is able to find out the optimal cluster radius to further decide a number of hidden node because PSO has better capacity of global searching than GA.

6. Conclusion
This paper has presented a novel approach for self-structure RBFNN. A very important step for the RBFNN training is to decide a proper number of hidden node. If the number of hidden node does not chosen properly, the RBFNN may present poor global generalization capability, slow training speed, and the requirement of large memory space. Therefore, to
decide a suitable cluster distance factor ($\varepsilon$) is the crucial condition for creating an optimal self-structure RBFNN. This paper proposed a PSO-based approach for searching the optimal $\varepsilon$; further, RBFNN is able to determine the optimal number of hidden node automatically.

For proving benefits of the proposed PSO-based approach, the simulations consisting of six nonlinear system modeling were tested; meanwhile, GA-based approach and K-means approach were also carried out for comparison. Simulation results show that the PSO-RBFNN algorithm outperforms the GA-RBFNN and K-means methods by the minimal training RMSE and the minimal testing RMSE.

7. References


Particle swarm optimization (PSO) is a population based stochastic optimization technique influenced by the social behavior of bird flocking or fish schooling. PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. This book represents the contributions of the top researchers in this field and will serve as a valuable tool for professionals in this interdisciplinary field.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: