Journal of Southeast University (English Edition) Dec. 2002 Vol.18 No.4 ISSN 1003—7985

Evolving Neural Networks Using an Improved Genetic Algorithm*

Wen Xiulan[1,2]** Song Aiguo[1] Duan Jianghai[1] Wang Yiqing[1]

([1] Department of Instrument Science and Technology, Southeast University, Nanjing 210096, China)

([2] Mechanical Engineering College, Inner Mongolia Polytechnic University, Huhot 010062, China)

Abstract: A novel real-coded improved genetic algorithm (GA) of training feed-forward neural network is proposed to realize nonlinear system forecast. The improved GA employs a generation-alternation model based the minimal generation gap (MGP) and blend crossover operators (BLX-α). Compared with traditional GA implemented in binary number, the processing time of the improved GA is faster because coding and decoding are unnecessary. In addition, it needn't set parameters such as the probability value of crossover and mutation by experiences. Therefore, it has the advantages of simple algorithms, strong robustness and high optimization efficiency. Then forecasting nonlinear system using feed-forward neural network is presented. Simulation shows the method is rapid and effective.

Key words: genetic algorithms, neural network, nonlinear forecasting

Modeling and control of nonlinear systems are research hot points in control theory and application areas in recent years. Because neural network was proved to be a universal approximation[1], a 3-layer feed-forward neural network can approximate any nonlinear continuous function to an arbitrary accuracy. Neural networks are widely applied in areas such as prediction, system modeling and control.

Back propagation (BP) learning is widely known as learning algorithm in neural networks. Given the set of teaching signals to the network, connection weights are adjusted in the direction of decreasing the differences between output activity and teaching signal, but the training procedures are easy to trap in local minimum and the training time is very long. A genetic algorithm is a directed random search technique invented by Holland[2] in 1975, which is widely applied in optimization problems. It is especially useful for complex optimization problems where the number of parameters is large and the analytical solutions are difficult to obtain. GA can help to find out the optimal solution globally over a domain. So it can be used to train neural network.

In this paper, a novel real-coded improved genetic algorithm is proposed to train BP neural network. Then using trained neural network to forecast nonlinear system is presented. Simulation shows the method is effective.

1 Improved Genetic Algorithm

GA is a powerful searching algorithm. The traditional GA processes are as follows[2]. First, a population of chromosomes is created. Second, the chromosomes are evaluated by a defined fitness function. Third, some of the chromosomes are selected for performing genetic operations. Forth, genetic operations of crossover and mutation are performed. The produced offspring replace their parents in the initial population. This GA process repeats until a user defined criterion is reached. However, a superior offspring is not guaranteed to produce in each reproduction process. In this paper, our improved GA is implemented in floating-point numbers, and the processing time is shorter than GA implemented in binary numbers as the coding and decoding processes are not needed and the probabilities of crossover and mutation are no longer needed. The improved GA adopts minimal generation gap (MGG) model proposed by Satoh et al.[3] and blend crossover (BLX-α)[4] strategy. Its details will be given as follows.

1.1 Minimal generation gap model

MGG model is shown in Fig.1. It is one of effective generation alternation models. In the MGG model, a generation alternation is done by applying a crossover operation n times to a pair of parents

Received 2002-09-15.

* The project supported by the Natural Science Foundation of Jiangsu Province (BK2001402); Dissertation Foundation of Southeast University — Nanrui Jibao Corporation.

** Born in 1966, female, associate professor.

randomly chosen from the population. From the parents and their children, we select the best and second individuals. With them the original parents are replaced.

1.2 Blend crossover

In BLX-α, offspring are generated as follows:

1) Choose two parents $\boldsymbol{x}^1, \boldsymbol{x}^2$ randomly from the population;

2) A value of each element x_i^c of the offspring vector $\boldsymbol{x}^c$ is chosen randomly from the interval $[X_i^1, X_i^2]$ following the uniform distribution

$$\left.\begin{aligned} X_i^1 &= \min(x_i^1, x_i^2) - \alpha d_i \\ X_i^2 &= \max(x_i^1, x_i^2) + \alpha d_i \\ d_i &= |x_i^1 - x_i^2| \end{aligned}\right\} \tag{1}$$

where x_i^1 and x_i^2 are the i-th elements of $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$; respectively, and α is a positive parameter.

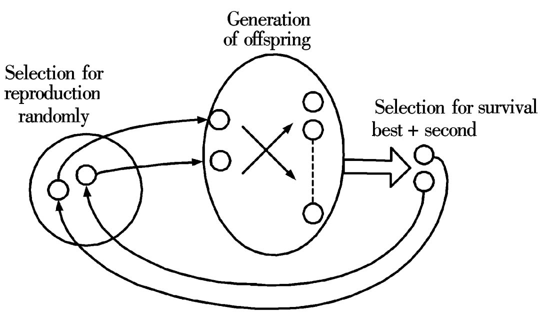

Fig.1 Mimimal generation gap model

2 Evolving Neural Networks Using the Improved GA

2.1 BP neural networks

In a three-layers BP neural networks, assume that the node numbers of input layer, hidden layer and output layer are N_i, N_h, and N_o, respectively. When the structure is determined in advance, we only consider optimizing weight and bias parameters. So a chromosome is an array of floating point numbers that corresponds to the weight and bias parameters of a neural network. In feed-forward neural networks, when the input pattern $X = [x_1, x_2, \cdots, x_{N_i}]$ is set to input neurons, the j-th neuron activity of hidden layer is given by

$$z_j = f_1\left(\sum_{i=1}^{N_i} w_{j,i} x_i + b_j\right) \qquad j = 1,2,\cdots,N_h \tag{2}$$

The k-th neuron activity of output layer is given by

$$y_k = f_2\left(\sum_{j=1}^{N_h} v_{k,j} z_j + c_k\right) \qquad k = 1,2,\cdots,N_o \tag{3}$$

where $w_{j,i}$ denotes the weight of the link between the i-th input and the j-th hidden node; $v_{k,j}$ denotes the weight of the link between the j-th hidden node and the k-th output node; b_j and c_k denote the biases for the hidden nodes and output nodes, respectively.

f_1 and f_2 are sigmoid functions expressed as

$$f_1(x) = \frac{1}{1 + \exp(-\varepsilon_1 x)} \tag{4}$$

$$f_2(x) = \frac{1}{1 + exp(-\varepsilon_2 x)} \tag{5}$$

where ε_1 and ε_2 are constants.

The objective function of training BP neural network is defined by

$$\boldsymbol{J}(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{b}, \boldsymbol{c}) = \frac{1}{2}\sum_{k=1}^{N_o}(\bar{y}_k - y_k)^2 \tag{6}$$

where $\bar{y}_k$ and y_k are the target and the actual output of the k-th output neuron. Obviously the objective is to minimize $\boldsymbol{J}$ by adjusting to weight and bias vectors $\boldsymbol{w}$, $\boldsymbol{v}$, $\boldsymbol{b}$, and $\boldsymbol{c}$.

2.2 Evolving neural networks using the improved GA

The algorithm of the evolving neural networks used in this paper is as follows.

Step 1 Generate N individuals randomly as initial population;

Step 2 Choose two individuals to be parents randomly from N individuals;

Step 3 Generate n offspring by applying the BLX-α to the parent n times;

Step 4 Evaluate offspring and parents: Create n neural networks from the offspring and 2 neural networks from the parents. Let each neural network perform the task to evaluate it and let the result be the evaluation value;

Step 5 Select survival: Select the best and the second individuals from the family consisting of the two parents and their offspring and replace the two parents in the population with the selected individuals;

Step 6 If a stop condition is satisfied, stop the algorithm. Otherwise, go to step 2.

3 Simulation

For simplicity, a nonlinear system described by Eq.(7) is considered.

$$y = e^{-1.9(x+0.5)}\sin 10x \tag{7}$$

In identification, a three-layers feed-forward network trained by using above method is applied to forecast (7). The neuron numbers of input layer and output layer are all 1 and the neuron number of hidden layer is 4. Population size N, offspring population size n and the maximum evolution iteration are 20,20 and 2?000, respectively. The value of variable $\boldsymbol{x}$ is from -0.50 to 0.45 uniformly and the interval is 0.02. Curves of identification are shown in Fig.2. The training error is 0.004?3 and the maximum identification error is 0.082?7. It took 150.17?s to finish 2?000 iterations evolution. Evidently, neural network training time using the improved GA is very short. Also, with the increase of iteration number, training error will decrease.

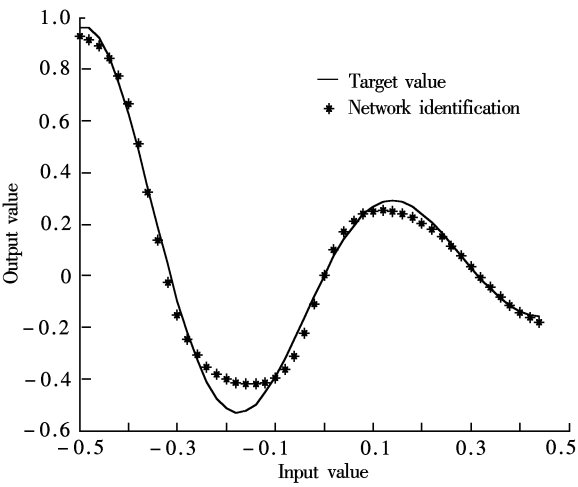

Fig.2 Curves of identification

4 Conclusion

In this paper, strategy on training a feed-forward neural network based on the improved genetic algorithm is expounded. Then forecasting nonlinear system using feed-forward neural network is presented. Simulation shows the method has rapid convergence speed and high optimization efficiency.

References

[1] Lam H K, Ling S H, Leung F H F, et al. Tuning of the structure and parameters of neural network using an improved genetic algorithm[A]. In: *Proc of IECON*[C]. 2001.25-30.

[2] Holland J H. *Adaptation in natural and artificial system*[M]. Ann Arbor, MI: University of Michigan Press, 1975.

[3] Satoh H, Yamamura M, Kobayashi S. Minimal generation gap model for gas considering both exploration and exploitation [A]. In: *Proc of IIZUKA*[C]. 1996. 494-497.

[4] Eshelman L J, Schaffer J D. Real-coded genetic algorithms and interval-schemata[J]. *Foundations of Genetic Algorithms*, 1993, **2**: 187-202.

[5] Ono I, Kobayashi S. A real-coded genetic algorithm for function optimization using unimodal normal distribution crossover [A]. In: *Proc 7th ICGA*[C]. 1997.246-253.

[6] Gong D, Xu S, Sun X. Research on fast training algorithm for recurrent neural network[A]. In: *IEEE Inter Symp on Industrial Electronics*[C]. 2001.446-448.

基于改进遗传算法进化神经网络

温秀兰[1,2] 宋爱国[1] 段江海[1] 王一清[1]

([1] 东南大学仪器科学与工程系, 南京 210096)
([2] 内蒙古工业大学机械学院, 呼和浩特 010062)

摘 要 本文提出一种新颖的基于实数编码的改进遗传算法用于前馈神经网络的训练,进而实现对非线性系统预测.该改进遗传算法采用基于代沟最小的代选择模型,选用 BLX-α 混合交叉算子.与经典的基于二进制编码的遗传算法相比较,该算法不需要编码和解码,所以计算速度快;且不需要根据经验设置交叉和变异概率,因而算法简单、鲁棒性强、优化效率高.同时给出了应用该算法对前馈神经网络进化时的计算流程.仿真结果证实该方法对非线性系统进行预测是快速有效的.

关键词 遗传算法, 神经网络, 非线性预测

中图分类号 TN911