Using Genetic Algorithm to Optimize Artificial Neural Network: A Case Study on Earthquake Prediction

Qiuwen Zhang, Cheng Wang
College of Hydropower and Information Engineering,
Huazhong University of Science and Technology, Wuhan 430074, China
qwzhang_hust@163.com

Abstract

By integrating the global searching advantage of Genetic Algorithm(GA) and the local searching ability of BP Artificial Neural Network(BP ANN), this paper proposes a new model of BP ANN based on GA (called GA-BP ANN). Firstly, it applies GA to optimize the initial interconnecting weights and thresholds of BP ANN. Then, it utilizes the BP algorithm to train the neural network more accurately. This method can speed up the convergence and avoid local minimum of BP ANN. The experiments of earthquake prediction with general BP ANN and optimized GA-BP ANN are respectively conducted as a case study. The results show that the BP ANN optimized with GA can not only get better network configurations, but also improve the efficiency, precision and stability of earthquake prediction.

1. Introduction

Artificial Neural Network (ANN) is an inter-discipline of biology and computer science, which has been widely applied in signal processing, pattern recognition, computer vision, intelligent control, nonlinear optimization, data fusion and data mining, knowledge discovery, and so on. Back Propagation(BP) algorithm is one of the most effective methods of ANN. Any continuous function in a closed interval can be approximated by using a BP ANN with one hidden layer. For any complicated system, if its samples of input and output are enough, a BP ANN model that reflects the relationships between the input and output variants can be constructed after repeated learning and training. So, BP ANN has very strong capabilities of nonlinear modeling and analysis for huge and complex system[1].

However, since the initial interconnecting weights of BP ANN are often stochastically given, the learning times and final interconnecting weights of the network are therefore changed for different times of training. That is to say, the trained network is not unique and sometimes the network possibly plunges into partial minimum. In addition, the blindness of the determination of initial interconnecting weights always results in too many training times and slow convergence. These shortages of BP ANN seriously impact its precision of modeling and effects of application. It is quite necessary to optimize and improve BP ANN[1].

Genetic Algorithm(GA) is an iterative algorithm that is parallel and global. According to the theory of GA, the possible solution in the field of problem is considered to be an individual or a chromosome of the colony, and all the individuals are then coded to be symbol strings. By simulating the evolutionary processes of organisms such as natural selection and elimination, the colony is repeatedly selected, intercrossed and mutated. Based on the evolutionary rules of “survival of the fittest, and elimination of the unfittest”, as well as the adaptive estimation of every individual, better and better colony is gradually evolved. At the same time, the best adaptive individuals in the optimized colony are also searched by global and parallel ways[2].

Because the processed objects of GA are gene individuals that have been coded with parameter strings, GA can directly operate the structures of these objects. Especially, since GA evaluates multi-solutions in the searching space simultaneously, it has very strong ability of global searching and also easy to be parallelized[2]. It is possible to apply these advantages of GA to improve the shortages of BP ANN in partial optimization.

This paper proposes a technique to optimize BP ANN by using GA, and presents an integrated model that is called GA-BP ANN with the combined advantages of both BP ANN and GA. As a case study, the hybrid model is applied to predict earthquake.

978-0-7695-3334-6/08 $25.00 © 2008 IEEE
DOI 10.1109/WGEC.2008.96

2. Optimizing BP ANN with GA

During the learning and training of BP ANN, there are two key factors that influence the modeling effects and precisions. One is the initial interconnecting weights of the network, and another is their modified quantities. Because of the lack of evidences, the initial interconnecting weights of BP ANN are often stochastically and blindly produced, it is difficult to determine the initial interconnecting weights that are global, which might cause the network to run into partial optimization and therefore decrease the probability to obtain the best global solutions. In addition, because the Delta rule is always adopted to modify the interconnecting weights of BP ANN, the convergence velocity is always slow, or sometimes the network even does not converge. These shortages of BP ANN are quite necessary to be optimized and improved[1, 3].

The slow convergence velocity and oscillation effect of BP ANN are always resulted from the partial minimum at the end of learning and training. However, the problem of partial minimum of BP ANN can be solved by adjusting the initial interconnecting weights of the network. So the most important optimization of BP ANN can be focused on the optimization of its initial interconnecting weights[1, 3].

Because the learning and training of BP ANN belongs to nonlinear problem, it is not easily to be realized by using general methods. GA is a nonlinear optimization method that has very strong ability of global searching. It can be adopted to optimize BP ANN.

After the structure and parameters of an ANN such as the number of layers and the number of neurons in every layer are determined, the approximate ranges of interconnecting weights and thresholds can be computed by using the algorithm of back propagation. Then, an initial colony in the field of solution can be randomly produced by adopting genetic algorithm. Each group of the interconnecting weights and thresholds of BP ANN is considered to be an individual of the colony and coded to be chromosome(Figure 1).

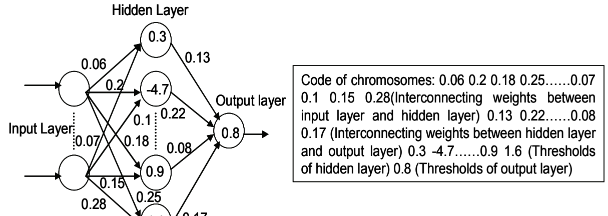

Figure 1. **Chromosomes coding of the interconnecting weights and thresholds of BP ANN**

With the errors of BP ANN as the adaptive function, all the chromosomes are operated by the selection, intercross and mutation of GA. The chromosomes that correspond to the best initial interconnecting weights and thresholds of BP ANN are gradually evolved. Finally, the initial interconnecting weights and thresholds optimized with GA are input into the learning and training of BP ANN again, and a better BP ANN model can be constructed.

Because the initial interconnecting weights and thresholds of BP ANN are globally optimized by utilizing GA, an appropriate search space can be located in the complex solution space, and the partial minimum or non-convergence can be avoided. The BP ANN optimized with GA is called GA-BP ANN.

3. Case study in earthquake prediction

3.1. Problem definition and sample data

In order to validate the method proposed in this paper, general BP ANN and GA-BP ANN are respectively adopted to predict earthquake, and their effects are compared and analyzed.

It is well known that the mechanism of earthquake is very complicated. Accumulation frequency of earthquake, accumulated release energy, b value, number of abnormal earthquake groups, number of earthquake zones, period of earthquake activity and magnitude of earthquake in neighboring area are the most important factors that influence the process of earthquake preparation and occurrence. Different combination of these factors may cause different earthquake risks[4].

The relationships between these factors(conditions) and their corresponding earthquake activity(results) can be described by a BP ANN model. Then, with the factors inputting into the BP ANN, the earthquake risks of these factors can be predicted by calculating the output of the constructed BP ANN.

So, the problem of earthquake prediction can be defined and described with a BP ANN model: the above 7 factors influencing earthquake are the input neurons, and the probability of earthquake risk is the output neuron of the network(Figure 2).

In the experiment of earthquake prediction, 17 groups of input and output data are selected to be the samples, among which 14 groups are the samples used to train the BP ANN, and other 3 groups are the validation samples.

3.2. Technical platform

The technical platform adopted to realize BP ANN and GA in this paper is MATLAB7.0 developed by

MathWorks. One of the most characteristics of MATLAB is its numerous toolboxes and simulation modules oriented to application. There are many different functions. In this paper, the toolboxes of neural network and genetic algorithm are utilized to model and analyze the problem of earthquake prediction[4-5].

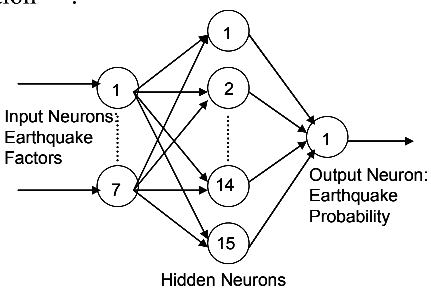

Figure 2. **Architecture of the BP ANN for earthquake prediction**

3.3. Training of the BP ANN

According to the results of repeated tests, the architecture of the BP ANN for earthquake prediction is setup to be three layers: one input layer, one hidden layer and one output layer. The number of their neurons is 7, 15 and 1, respectively. Each of the input neurons is corresponding to one factor related to earthquake, and the output is the earthquake probability(Figure 2). Tansig and Logsig are selected to be the activated functions, error precision is set to be 0.001, training function is trainlm algorithm, learning rate is 0.01, and momentum is 0.9[4].

By using the algorithm of back propagation, the ANN is trained, and the approximate ranges of interconnecting weights and thresholds of the network are obtained. The minimum and maximum of interconnecting weights and thresholds is U_{min} and U_{max}, respectively.

3.4. GA programming and parameterization

The toolbox of gatds in MatLab is applied to realize GA[5]. The interconnecting weights and thresholds of the trained BP ANN are considered to be the individuals of a colony. Since the BP ANN for earthquake prediction is a network with three layers, and the number of their neurons is 7, 15 and 1, respectively, the total number of interconnecting weights and thresholds of the BP ANN therefore is$(7\times 15+15)+(15+1)=136$. Among these 136 data, 105 of them are the interconnecting weights between input neurons and hidden neurons, 15 are the interconnecting weights between output neurons and hidden neurons, 15 are thresholds of hidden neurons and 1 is the output neuron(Figure 1). All of them are programmed to be chromosomes with real numbers.

In the interval of $[U_{min}-\delta_1, U_{max}+\delta_2]$, which is the field of solutions of interconnecting weights and thresholds in GA, an initial colony with 200 individuals is stochastically produced. In this colony, there are 20 sub-colonies(SUBPOP), and the number of individuals in each sub-colony is 10.

In the processes of GA evolution, generation gap(GGAP) is set to be 0.8, intercrossing rate(XOVR) is 1, mutating rate(MUTR) is 0.01, the maximum generations (MAXGEN) is 500, migration rate(MIGR) is 0.2 and insert rate(INSR) is 0.9. The reproducing operator is gear-bet. Other parameters of GA are set to be the defaults.

$SE=\sum_k\sum_m(Y_{km}-T_{km})^2$ is determined to be the adaptive function of GA, where SE is the square errors, Y_{km}, T_{km} is the actual value and output value, respectively, k is the number of samples and m is the number of neurons of the output layer.

After the operation of selecting, intercrossing and mutating in GA, the chromosomes with lower adaptation are gradually eliminated, and finally the best chromosomes(optimized solutions) that correspond to the interconnecting weights and thresholds of BP ANN are evolved.

3.5. Analyses of the results

According to the above parameter configurations of BP ANN and GA, both general BP ANN(called G-BP ANN) and optimized BP ANN with GA(called GA-BP ANN) are respectively applied to predict the earthquake in the test area. In order to compare and validate the precision and robustness of the method proposed in this paper, 5 times of predictions test with the same parameters are respectively calculated. The results are shown in Table 1.

It is indicated that the absolute errors, relative errors and average errors of the earthquake prediction with GA-BP ANN are smaller than those of G-BP ANN in all the three groups of samples(15,16 and 17). The decreases of averaged relative errors are up to 0.1110-0.4698. So it can be concluded that the BP ANN optimized with GA(GA-BP ANN) reduces the errors and therefore improves the precisions of earthquake prediction compared with general BP ANN(G-BP ANN).

In addition, the minimum and maximum of the averaged relative errors of GA-BP ANN is 0.0632 and 0.1585, respectively, their extent of variation is only 0.0953. While the range of the averaged relative errors of G-BP ANN is 0.2495-0.5330, their extent of variation is up to 0.2835. It shows that the training of

Table 1. **Comparison of the results of earthquake prediction with G-BP ANN and GA-BP ANN**

No of test	No of sample	Actual value	Predicted value		Absolute error		Relative error		Averaged relative error		Changes of error
			G-BP ANN	GA-BP ANN	G-BP ANN	GA-BP ANN	G-BP ANN	GA-BP ANN	G-BP ANN	GA-BP ANN	
1	15	0.6250	0.7584	0.7235	0.1334	0.0985	0.2134	0.1576	0.2495	0.1385	0.1110
	16	0.7187	0.7462	0.7436	0.0275	0.0249	0.0383	0.0364			
	17	0.3750	0.1888	0.2913	-.1862	0.0837	0.4966	0.2331			
2	15	0.6250	0.8658	0.6651	0.2408	0.0401	0.3852	0.0642	0.3123	0.1585	0.1538
	16	0.7187	0.7951	0.7148	0.0764	-.0039	0.1063	0.0054			
	17	0.3750	0.5421	0.2229	0.1671	-.1521	0.4455	0.4057			
3	15	0.6250	0.9782	0.6234	0.3532	-.0016	0.5651	0.0025	0.5330	0.0632	0.4698
	16	0.7187	0.9530	0.6108	0.2343	-.1079	0.3261	0.1501			
	17	0.3750	0.1096	0.3889	-.2654	0.0139	0.7078	0.0371			
4	15	0.6250	0.4995	0.6382	-.1255	0.0132	0.2007	0.0211	0.3683	0.1563	0.2120
	16	0.7187	0.3809	0.5940	0.3378	0.1247	0.4700	0.1735			
	17	0.3750	0.5378	0.4778	0.1628	0.1028	0.4342	0.2742			
5	15	0.6250	0.5889	0.6718	0.0361	0.0468	0.0578	0.0749	0.3475	0.1551	0.1924
	16	0.7187	0.4881	0.6323	-.2306	-.0864	0.3208	0.1202			
	17	0.3750	0.6240	0.4764	0.2190	0.1014	0.6640	0.2704			

BP ANN optimized with GA is more stable than that of G-BP ANN, and the convergence velocity of GA-BP ANN is faster than that of G-BP ANN.

So, it can be concluded that compared with G-BP ANN, GA-BP ANN not only has greater efficiency and precision, but also has higher stability in the earthquake prediction.

4. Conclusions

BP ANN is a very effective method of nonlinear modeling and analysis. It has many advantages such as memory, learning and self-adaptation, but the convergence of the algorithm is slow, especially, for the problems of large searching space, multi-peak and non-differential function and very complex multi-dimensional curve, in which multi-local extremum points exist, the BP ANN is easily plunged into local extremum points. In addition, the selections of the initial interconnecting weights, thresholds and the structure of the network are always random. It is difficult to choose their global values. So the probability to obtain the best global solution is not great. All these shortages of BP ANN impact the functionality and therefore restrict its application.

GA is globally convergent and independent of the initial values. By integrating the advantages of global searching of GA and the instructive searching of BP ANN, a hybrid BP ANN model based on GA that is called GA-BP ANN can be constructed.

In this integrated model of BP ANN and GA, the initial interconnecting weights and threshold of BP ANN are optimized with GA so that the learning and training velocity of BP ANN is increased and the best network architecture is obtained. The experiments show that both the precision and robustness of the optimized GA-BP ANN are obviously improved.

Acknowledgments

The authors would like to appreciate the financial supports for this work from the National Natural Science Foundation of China (Grant #40672179), the Program for New Century Excellent Talent of Ministry of Education of China (Grant #NCET-07-0340), the National Basic Research Program of China (Grant #2007CB714107) and the Special Prophase Project of Key Basic Researches of Ministry of Science and Technology of China (Grant #2004CCA02500).

References

[1] J. G. Yang, S. Y. Weng, and H. Zhao, *Applied Textbook of Artificial Neural Network*, Zhejiang University Press, Hangzhou, 2001.

[2] X. P. Wang, and L. M. Cao, *Genetic Algorithm: Theory, Application and Software*, Xian Transportation University Press, Xian, 2002.

[3] J. G Yang, S. Y Wen, and H. Zhou, "An Optimized BP Network Model Using Genetic Algorithm for Predicting the Ignition-Stability Index of Pulverized Coal", *Dynamic Engineering*, 2006, vol.26, no.1, pp. 81-83.

[4] J. Guo, and W. J. Sun, *Theory of Neural Network and its Implementation with MatLab*, Electronic Industry Press, Beijing, 2005.

[5] Y. J. Lei, S. W. Zhang, and X. W. Zhang, *Genetic Algorithm Toolbox of MatLab and its Application*, Xian University of Electronic Science and Technology Press, Xian, 2005.