

# Application of Genetic Algorithm and Neural Network in Construction Cost Estimate

Guangli Feng, Li Li

School of Computer science and technology, Henan Institute of Engineering  
Zhengzhou, 451191, China  
E-mail:feng\_jsj@163.com

**Abstract**—Genetic algorithm optimizing BP has been proposed to aim at handling locality minimum and low convergence speed. The method based on analyzing the basic fundamental states that how to use genetic algorithm to improve the ability of BP. After optimizing, the GA-BP model has been built up. The result of GA-BP model can get lower forecast error and iterations. For these reason, GA-BP model is appropriate for construction cost estimation.

**Keywords**—neural network; BP algorithm; genetic algorithm; construction cost estimate

## I. INTRODUCTION

Project cost estimation is the feasible research basis of the project, and whether the cost estimate is accurate or not affects project investment decision directly, and also affects the competition ability of project bidding. This shows it is important significant for project cost calculation to have a relatively more quickly and more accurate engineering cost estimation method. In recent years, the development of artificial neural network provides an effective way for solving this problem. Artificial neural network usually has two forms<sup>[1,2]</sup>. One is the engineering applications based on BP neural network, and the other is based on Hopfield neural network<sup>[3]</sup>. In the most of engineering applications, the researchers still use BP neural network more, while Hopfield neural network is often used in engineering structure optimization<sup>[4]</sup>.

## II. THE BP NEURAL NETWORK AND GENETIC ALGORITHM

### A. BP Algorithm and Its Shortcomings

BP neural network in the structure is similar to multilayer perceptions<sup>[5]</sup>. It is a kind of multilayer feed forward neural network, and also a kind of mapping function which has the ability of reverse transmission and error correction. It expresses the system through associative memory and learning input/output parameters of unknown system.

The main idea about BP is: input learning samples and spread information with competition<sup>[6]</sup>. Then repeatedly adjust and train the weights and bias of network by using the back propagation algorithm to make the output vector and expected vector as closer as possible. The adjusting and training don't finish until the sum of squares of network

output layer error is less than the specified error, the last save network weights and bias.

Because the BP network is derived from gradient method, it requires objective function is continuously differentiable, so it can cause some problems in the learning process, such as: slow convergence speed, easily falling into local optimum, finding the global optimal value difficultly, and causing oscillation effect, etc. what's more, because the weights and threshold values of BP network are random value on the choice, this is, the initial value is different from each other every time, which causes each training learning and prediction results different.

### B. Genetic Algorithms

Genetic algorithm has a strong macro search capability, and can find global optimal solution with the larger probability, so it can better overcome the shortcomings of BP algorithm using it to complete the search<sup>[7]</sup>.

Genetic algorithm is a kind of adaptive heuristic global search algorithm which imitates nature evolution. By simulating Darwin's biological evolution of genetic selection and natural selection, Genetic algorithm carries on the repeated operations on groups based on genetic(genetic, crossover and mutation). According to the scheduled target, fitness function evaluate each individual on the basis of the evolution rules of "survival of the fittest", and constantly get better groups, at the same time, it searches out the optimum individuals in optimization groups by the global parallel search mode to get the solution which meets the requirements. The organic integration between BP neural network and genetic algorithmic gives full play to the global search ability of genetic algorithm and the local search ability of BP neural network so that better prediction results can be gotten<sup>[8]</sup>.

## III. MATLAB IMPLEMENTATION OF USING GA - BP IN CONSTRUCTION ENGINEERING COST ESTIMATION

This study selects the engineering characteristics of construction engineering as the research object, and intends to use matlab to implement cost estimation function<sup>[9]</sup>. According to the book "the construction technology", written by ZhangXi and published in Chinese plan press March 2008, the classification criteria of related building engineering select 7 indexes as construction engineering classification characteristics indicators. The 7 indexes are: foundation

types, structure form, floor number, Windows and doors, outside wall adornment, wall body, and plane combination. The quantitative description is shown in table 1. The seven

characteristics index data are extracted as engineering sample data from in 20 building data of a certain construction enterprise's.

TABLE I. QUANTITATIVE DESCRIPTION

Quantitative Value	1	2	3	4	5	6
Foundation	Brick strip base	Steel bar tong strip base	The powder spurts the pile reinforcement place Basic reinforced concrete strip foundation	Piece raft foundation steel bar tong strip base	Steel bar tong strip base	Steel bar tong full house base
Structure	Brick-concrete	Frame precast block	Entire cast-in-place framework	Frame-shear Wall Structure		
Story Number	6	5,7	8,9,10	11,12,13	14,15,16	17,18,19
Doors and Windows	Wooden doors and windows	Wooden Doors and plastic window	Wooden doors and aluminum windows	Aluminum alloy window		
Decoration	External wall	Sticks the stone dryly	granitic plaster	Face brick		
Wall	Hollow brick	Standard brick	Block	Reinforced Concrete diaphragm		
Plane Combination	One room and one hall	Two room and one hall	two room and two hall	Three room and two hall Four room and one hall	Four room and two hall	

The data in table I shows the construction engineering samples of a certain area in our country. Each sample shows its characteristics with seven physical and chemical indexes, and determines the network's input mode as follow:

$$Pk = (P1k, P2k, \dots, Pnk) \quad k=1, 2, \dots, m, \quad n=7$$

It means there are m groups of construction engineering sample vectors in total, and each sample includes seven elements.

According to table 1, the quantitative description of any one of the architectural engineering mode can be given. It is

expressed as  $T_i = (ti1, ti2, \dots, tij)$ ,  $T_i$  is the serial number of No. I project ( $i = 1, 2, \dots$ );  $t_{ij}$  ( $j = 1, 2, \dots, 7$ ) is the quantitative value of the No. j characteristic in the No. i project. If a project (set its serial number to i) is brick base, brick and concrete structure, 5 floors, wooden door and aluminum alloy window, exterior wall dry stick stone, standard bricks, three rooms and one living room, then its quantitative description is  $T_i = (1, 1, 2, 3, 2, 2, 3)$ , according to this method, the training sample can be got like Table II.

TABLE II. TRAINING SAMPLE

No	Input							Output			
	X1	X2	X3	X4	X5	X6	X7	Y1	Y2	Y3	Y4
1	1	1	2	1	1	2	2	498.2	16.10	0.006	143.1
2	3	1	2	3	3	2	4	525.1	18.45	0.004	178.3
3	2	1	1	1	2	2	2	493.4	17.18	0.007	159.3
4	1	1	1	1	1	1	2	487.4	15.9	0.005	132
5	1	1	1	3	2	2	3	506.5	15.6	0.005	138
6	2	1	2	3	3	2	4	538.6	16.4	0.004	149
7	3	1	1	1	2	2	4	542.9	17.9	0.004	168
8	4	1	2	3	3	2	5	562.4	19.2	0.004	172
9	2	2	4	3	3	3	4	897.2	26.0	0.004	208
10	3	2	5	3	3	3	3	989.7	29.4	0.003	221
11	4	2	6	3	3	3	4	1045	27.9	0.004	223
12	5	2	4	2	4	3	4	876	33.2	0.006	236
13	5	4	6	3	4	2	4	857	35.2	0.003	239
14	5	2	4	3	3	3	4	923	28.9	0.004	216
15	6	2	3	3	3	3	4	948	30.4	0.005	211.3
16	6	3	4	3	3	3	3	747	36.0	0.003	242
17	6	2	4	3	4	3	4	689	32.0	0.004	232
18	6	4	6	3	4	2	3	936	38.4	0.003	247

A. Matlab Implementation of Using BP Network in Construction Engineering Cost

First, apply GA - BP algorithm in table 2 to neural network learning, and save the value of the results, then substitute the test value into the network to forecast it.

Assume there are three layers in BP network including input layer hidden layer and output layer. Set the value from input layer to hidden layer as  $w_1$ , its valve value as  $v_1$ ; set the value from hidden layer to output layer as  $w_2$ , its valve value as  $v_2$ . The sample contains singular matrix, so first of all the sample should be normalized.

In the normalization process, assume the input layer neurons as 7, output layer neurons as 1, according to the kolmogorov theorem, the number of hidden layer neurons can be obtained between 4 and 14. Set the number of hidden layer neuron as 10, set neuron transfer function of hidden layer as *tansig*, neuron transfer function of output layer as *purelin*, BP network training function as *trainlm*. Set Training times as 1000, learning rate is 0.1, target value is 0.0002.

Set the number of input and output sample, then establish the BP network . Put the predicted value *input\_test* into the network which has been established. *h input\_test* has two vector groups, one is [2 1 1 3 2 2 2.5],the other is [5 4 7 3 4 3 4.1].

The calculation error is 0.1075, 0.1261. keep other parameters same, and set neuron transfer function of output layer as *logsig*, training function of BP network as *traingdx*, then the operation result is shown in figure 1:

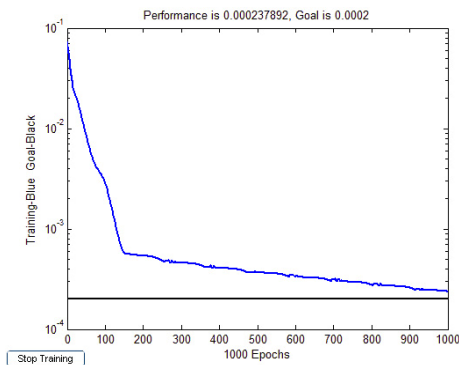


Figure 1. Logsig error change

Figure 1 shows that the BP network has gotten in local minimum values. How can the target value be reached in steps? In order to solve the problem, we introduce genetic algorithm to optimize it.

#### A. Matlab Implementation of Using GA - BP Mode in Construction Project Cost

Applying Genetic algorithm to neural network mainly plays two roles. One is to optimize artificial neural network (ANN) structure, the other is apply GA to learn the weight of ANN, this is, use genetic algorithm to replace some of the traditional learning algorithms [10]. Specific steps are as follows:

##### 1) Coding

Genetic algorithm doesn't operate the actual decision variable of optimization problem, so the first question of using genetic algorithm is that decision variable is expressed as string structure data through the coding.

##### 2) Initialization group

Set the number of group num as 80, then the generation function of initial group is:

```
Pop=initialize ga (num, bounds,'fitness');
```

3) Calculate the fitness of each individual in current group , and write objective function

Calculate the fitness of each individual in current group , and write objective function

```
function [sol, Val] =fitness(sol, options)
```

```
numv=size (sol, 2)-1;
```

```
x=sol (1: numv);
```

```
Eval=fitness decode(x);
```

4) Choose

The selection process: use the adaptive value of each individual which is obtained after decoding to eliminate some poor individuals for the crossover and mutation operation in the next step. The algorithm uses the idea of combining optimal preservation strategy with proportion choice method, this is, first to find out the individual whose adaptive value is maximum or minimum in the current group, and then be keep the best individual and use it to replace the worst individual . In order to ensure the best individual will not be destroyed by the operation of crossing and mutation, it is allowed to directly get into the next generation without participating in crossing and mutation Then the other individuals will be operated according to proportion choice method. Proportion method of choice refers to the probability of selecting the individual is in proportion to its adaptive value. The purpose of combining two methods together is: in the progress of genetic operation, we can not only improve the average adaptive value of the group but also ensure that the adaptive value of the best individual doesn't change small.

5) Crossing

Now we use the method of single-point crossing to realize crossover operator, namely, randomly set an intersection according to individual coding strings which are paired up on the basis of choice probability PC, then exchange part genes of two matching individuals at the point, thus form two new individuals. The realization of the crossover operator as follows:

```
Function [c1, c2] = earthmover (p1, p2, bounds, ops)
```

```
a = rand; % pick up a random number
```

```
% create children
```

```
c1 = p1*a + p2*(1-a);
```

```
c2 = p1*(1-a) + p2*a;
```

6) Variation

For binary genetic string, mutation operations are to randomly selection variation point in accordance with mutation probability pm and invert its bit at the mutation point.

According to the method mentioned above, get the operation result which is shown in figure 2:

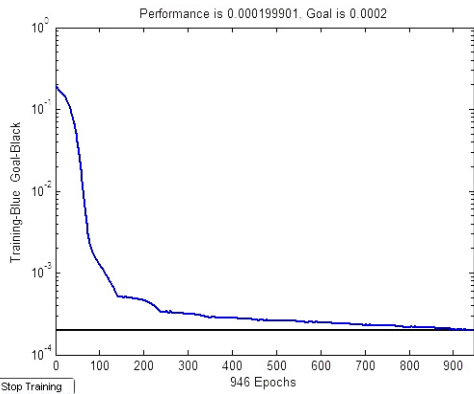


Figure 2 shows the same questions as those in figure 1, that is, genetic algorithm can be very good at speeding up convergence speed and solving the problem of local minimum to realize the global search. The calculation error is 0.0069, 0.0263. In comparison to the error generated by using BP, the error is much lower and its generalization ability is stronger.

The error curve in BP network training decreases more quickly, while that in GA - BP network training declines gently at the beginning and reduces rapidly after getting relative optimal solution, and its error is 0.0492, 0.1264 , which is smaller than BP network's.

#### IV. CONCLUSION

Aiming at the existing deficiency of BP network mode, this paper puts forward the improved network mode (GA - BP). GA - BP is a kind of fast and effective method. it conducts genetic selection for many points in solution space, and searches for optimization point by the improved BP algorithm, which can not only overcome the problems of easily falling into local minimum point and the poor

generalization ability in BP algorithm, but also overcome the long-time training which is caused by searching for optimal solution in a similar random form.

The simulation results show that GA - BP model gets great improvement in generalization ability, and has higher reliability. It is more suitable for the practical application of engineering cost estimation than BP model.

#### REFERENCES

- [1] Stanley, K. O. Miikkulainen, R. Evolving neural networks through augmenting topologies. *Evolutionary Computation*, pp. 99–127, 2002.
- [2] Hai Jin, Shi Zhongzhi, Ye Shiwei. *Neural network principle [M]*, Mechanical industry press, 2004.
- [3] C.T. Lin, C.S.G. Lee, Neural-network-based fuzzy logic control and decision system, *IEEE Transactions on Computers*.Dec.1991, pp.1320–1336.
- [4] Yuan Qing, Application of Neural Network in Missile Test System Fault Diagnosis, *Journal of Test and Measurement Techol*, pp.702-706, June,2002.
- [5] Chen Xiaoqu,Lu Congda,Liao Zhiping,Improvement of BP Learning Algorithm in Matlab,*Control Engineering of China*, pp.96-98 May,2005,.
- [6] Shen Jiashan, A quick cost-estimating method for the civil engineering based on BP neural network, *HeNan Science*, pp.479-482 August,2003.
- [7] David J. Montana and Lawrence Davis, Training Feedforward Neural Networks Using Genetic Algorithms,*The International Joint Conferences on Artificial Intelligence* ,1989, pp.762-767
- [8] A. F. SHAPIRO. The merging of neural networks, fuzzy logic, and genetic algorithms.*Insurance:Mathematics and Economics* ,vol.31, pp.115-131,2002.
- [9] Ge Zhexue, Sun ZhiJiang,Neural network theory and matlabR2007 implementation , *Electronic Industry Press* ,2007.
- [10] Lei Yingjie , Zhang Shanwen, Li Xuwu, Zhou Chuangming. *Matlab genetic algorithm toolbox and application*,Xian university of electronic science and technology Press, 2006.