CIVIL DEVELOPMENT PROJECT COST ASSESSMENT BY NEURAL NETWORK WITH GENETIC ALGORITHM

Arvind Jaiswal

Gaurav Dubey

Abstract

Cost assessment is most important part when new civil development is start, if cost assessment is not correct then project may not complete within a budget. So many times budget is over and development is stopped. To solve this problem we are using the concept of neural network and genetic algorithm, and designing a model by which we can assess the cost of any civil project very efficiently. In this paper we are using Backpropagation Method of Neural Network and Genetic Algorithm. Backpropagation method will give some solution of problem and our genetic algorithm will give the optimize solution. During the process of development cost assessment, back propagation neural network has a very powerful application. Genetic algorithm optimizing backpropagation has been proposed to aim at handling local minimum and low convergence speed. The method based on analyzing the basic fundamental tells that how to use genetic algorithm to improve the ability of back propagation. After optimizing, the Genetic Algorithm Back Propagation model has been built up. Some many project cases and some testing samples are put into the model to observe generalize ability. The result of Genetic Algorithm Back Propagation model tangle lower forecast error and iterations. For these reason, Genetic Algorithm Back Propagation model is appropriate for construction cost assessment.

1 Introduction: Cost assessments are essential to project related engineering and greatly influence construction management. Such assessments allow owners and planners to evaluate project feasibility and control costs effectively in detailed project design work. Cost overruns are a major problem, especially with today's emphasis on tight budgets. Overruns may lead to cancellation of a project. In some cases, a potential overrun may result in modifying a project to a design-to-cost task. Our goal was to develop a model that shows the relationship between the estimated cost of a project and the amount that should be spent on doing an assessment. We hope that such a model may help prevent or at least reduce overruns due to inaccurate cost assessments. Cost assessment models have been developed using statistical methods. Regression analysis represents a traditional alternative, an inherent disadvantage of which is the requirement of a defined mathematical form for cost functions. Many previous researches have proved that the neural network cost assessment models are superior to the traditional regression estimation model. The purpose of this study was to apply genetic algorithms to handle the general local ability problem from improving the

---

Student M.Tech IV sem (CSE), Amity University, Noida (UP)

Asst Professor, Amity School of Computer Science, Amity University, Noida (UP)

accuracy of cost assessment and the performance and validity of optimizing both the neural network size and parameters using Genetic Algorithms.

Genetic Algorithm: Genetic Algorithm (GA) is an artificial intelligence procedure. It is based on the theory of natural selection and evolution. Genetic algorithms provide computers with a method of problem-solving which is based upon implementations of evolutionary processes. The computer program begins with a set of variables which internally resemble the chromosomes which store the genetic information in humans. Each genome of these digital chromosomes represents a trait of whatever the data structure is supposed to represent. This information can be stored in the binary bit form in, which each genome is classified as being on or off (0 or 1, respectively). Alternatively, they can be stored in a character string in which each character represents an integer value which describes the magnitude of a trait. The computer program first creates these digital chromosomes through stochastic (random) means, and then tests their "fitness". This can be done through one of two methods:

Fitness proportionate selection: This is a kind of genetic selection in which the computer would use a model or procedure to test the fitness of the chromosome and assign some kind of numeric value of fitness in comparison to other chromosomes.

Tournament selection: This form of selection involves "pitting" the chromosomes against each other in some kind of modeled environment. Those which survive the competition are deemed to be the fittest.

The computer program then takes the fittest chromosomes and creates another generation through the use of some kind of genetic operator. That is, the new generation of chromosomes can be created in either (or both) of two ways:

Genetic recombination: This is analogous to sexual reproduction, new offspring are created from the fittest chromosomes of the previous generation.

Mutation. This is analogous to the simulation of genetic mutation, which the offspring are identical to the parents but have random stochastic changes in their structure (and thus their fitness somewhat modified).

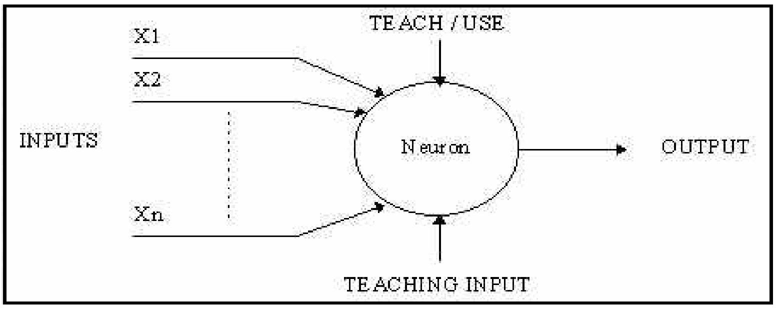
These two genetic operators can be used in different combinations, all of them producing different results. Using both would simply be genetically recombining the chromosomes and then mutating them and would most closely approximate the natural reproduction pattern of humans.

Uses of Genetic Algorithms: Genetic algorithms begin with a stochastic process and arrive at an optimized solution. Because of this, it will probably take much longer to arrive at a problem's solution through the use of a genetic algorithm than if a solution is found through analytical means and hard wired into the code of the computer program itself. For this reason genetic algorithms are best suited for those tasks which cannot be solved through analytical means or problems where efficient ways of solving them have not been found.

Back Propogation Neural Network : Backpropagation is a common method of training artificial neural networks so as to minimize the objective function. It is a multilayer structure. It is a multilayered feed-forward neural network backpropagation calculates the gradient of the error of the network regarding the network's modifiable weights. This gradient is almost always used in a simple stochastic gradient descent algorithm to find weights that minimize the error. Often the term "backpropagation" is used in a more general sense, to refer to the entire procedure encompassing both the calculation of the gradient and its use in stochastic gradient descent. Backpropagation usually allows quick convergence on satisfactory local minima for error in the kind of networks to which it is suited. The backpropagation algorithm for calculating a gradient has been rediscovered a number of times, and is a special case of a more general technique called automatic differentiation in the reverse accumulation mode.

Engineering approach

An artificial neuron is a device with many inputs and one output. The neuron has two modes of operation; the training mode and the using mode. In the training mode, the neuron can be trained to fire (or not), for particular input patterns. In the using mode, when a taught input pattern is detected at the input, its associated output becomes the current output. If the input pattern does not belong in the taught list of input patterns, the firing rule is used to determine whether to fire or not.



A simple neuron

Back Propagation Algorithm : Back Propagation algorithm has been proposed by Rumelhart as a controlled learning algorithm. Back Propagation neural network has a great application because of the simple structure, multiple adjustable parameters, many kinds of training algorithm and well manipulation. According to statistics, 80% -90% neural networks adopt Back Propagation network or the transformation of Back Propagation. Back Propagation network is the heart of the feed-forward network. It reflects the perfect content of neural network. Back Propagation algorithm has been made up of two sections: forward-propagation of information and counter-propagation of error. During the process of forward-propagation, the input information transmits to output layer through input layer and hidden layer. Neural condition in every layer can affect the neural condition in the next layer. If the expected output result can not get in the output layer, the error calculation should be worked out. Then the process must add counter-propagation.

Error original counter transform into the neural net every layer along the primary connection in order to get the expected goal.

The main ideology of Back Propagation can be conducted as follows:

Input the samples, forward-propagate the information, adjust the weight and error through counter-propagation to make the input data approach the expected data as soon as possible. The training can be over when the error sum-of-squares is lower than the specified error in the output layer at the end, the new weight and error should be saved.

Concrete steps can be stated as follows:

(1) Initialization. Weight [w][v] and threshold value [θ][r] can be given in a random way.
(2) forward-propagation of information. The output data in every layer neural should be computed by input and output specified model.

$$b_j = f(\sum_{i=1}^{n} w_{ij} x_i - \theta_j)$$
$$c_t = f(\sum_{i=1}^{n} v_{jk} b_j - r_t)$$

As on the above, b_j is the actual output result of the j th neural in the hidden layer; c_t is the actual output result of the unit in the output layer; w_j is the connecting weight from input layer to the hidden layer; v_{jk} is the connecting weight from the hidden layer to the output layer.

(3) Counter-propagation of error. The error counter-propagation in the process of Back Propagation network learning is affected by the following equation.

$$E = \frac{1}{2} \sum_{k=1}^{m} (u^k - y^k)^2$$

Among the expression, yk is the expected output data of the network.

Genetic algorithm: Genetic algorithms were first proposed by Holland (1975). Genetic algorithms are theoretically and empirically proven to provide a robust search in complex spaces, thereby offering a valid approach to problems requiring efficient and effective searches. When Genetic algorithms have been used to solve the problem, the model structure and parameters should be encoded generally by character strings. Solving a particular optimization task using Genetic algorithm requires the designer to address the five following issues.

(1) A genetic representation of candidate solutions,

(2) A way to create an initial population of solutions;
(3) An evaluation function which describes the quality of each individual;
(4) Genetic operators that generate new variants during reproduction; and
(5) Values of the parameters of the Genetic Algorithm such as population size, number of generations and probabilities of applying genetic operators.

Genetic algorithm and Back Propagation in construction cost estimate: This study investigates the use of genetic algorithm in the design and implementation of neural network controller. According to building operation technology, 7 features were chosen as classified standards in constructions such as base type, architecture form, number of piles, door and window, siding ornament, wall and plane assemble. The quantity description can be stated as follows: m kinds and building project samples have been got. Each sample has 7 features. So one network input node can be selected as follows.

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

It can be seen that m kinds of vectors of building project samples have been built up. There are 7 features in every sample. And then quantity description of any constructional engineering can be stated T_j $(j=1,2,\cdots,7)$ indicates quantity values of jth feature in project. Taking a project for example, the project has 7 features such as brick foundation, 5 floors, timber door and aluminum alloy window, siding rock dash, standard brick, three chambers and one hall. So the quantity description can be expressed in $T_i=(1,1,2,3,2,2,3)$. We can choose 20 premises from a constructional operations company, 7 features can be considered as samples. On the basis of the above method, training samples can be stated as table.

SAMPLES

N	input data							output
	x1	x2	x3	x4	x5	x6	x7	y1
1	1	1	2	1	1	2	2	498
2	3	1	2	3	3	2	4	525
3	2	1	1	1	2	2	2	493
4	1	1	1	1	1	1	2	487
5	1	1	1	3	2	2	3	506
6	2	1	2	3	3	2	4	538
7	3	1	1	1	2	2	4	542
8	4	1	2	3	3	2	5	562
9	2	2	4	3	3	3	4	897
10	3	2	5	3	3	3	3	989
11	4	2	6	3	3	3	4	1045
12	5	2	4	2	4	3	4	876
13	5	4	6	3	4	2	4	857
14	5	2	4	3	3	3	4	923
15	6	2	3	3	3	3	4	948
16	6	3	4	3	3	3	3	747
17	6	2	4	3	4	3	4	689
18	6	4	6	3	4	2	3	936

TABLE I.

The data of table 1 can be loaded, and then we can use Back Propagation algorithm and Genetic Algorithm - Back Propagation algorithm to learn the neural network with the

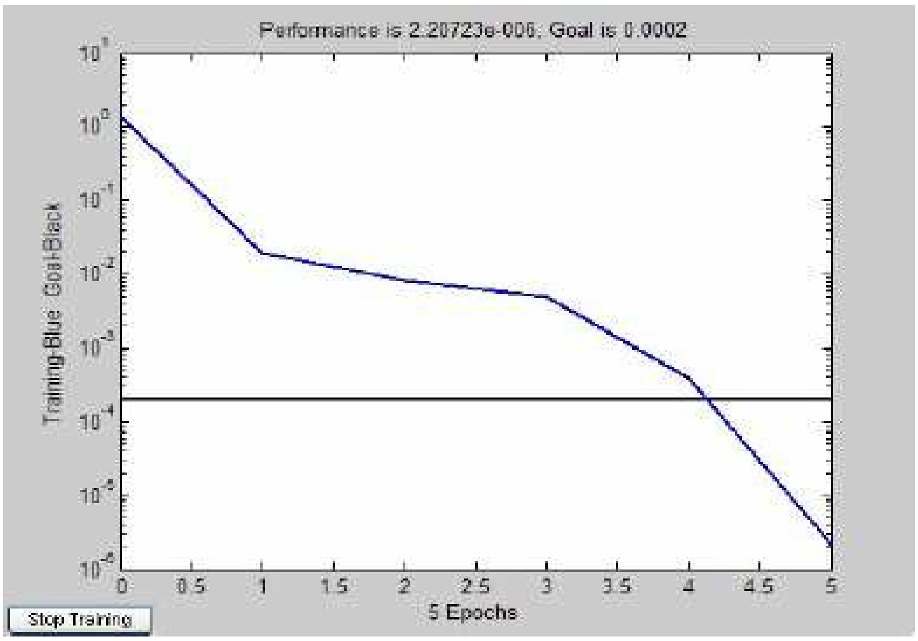
purposeofpredictingtestvalues.Therealhave otherexamplesofconstructioncost estimationwecanimplementityMATLAB.

Inourmodelwesetthehiddenlayernumberis10w Inspecialmethod,deleneuron functionisansiginthehiddenlayer,andtheout putlayer'sdeleneuronfunctionis purelinBackPropagationnetworktrainingfunction hasbeensetastrainlm.Weplanto train1000times,andthelearningrateis0.01.The goalvalueis0.0002.Thenwecanbuild upourBackPropagationnetwork.Theconcretecode canbewrittenasthefollowing:

net=newff(minmax(input_train),[10,1],{'tansig','purelin'},'trainlm')'

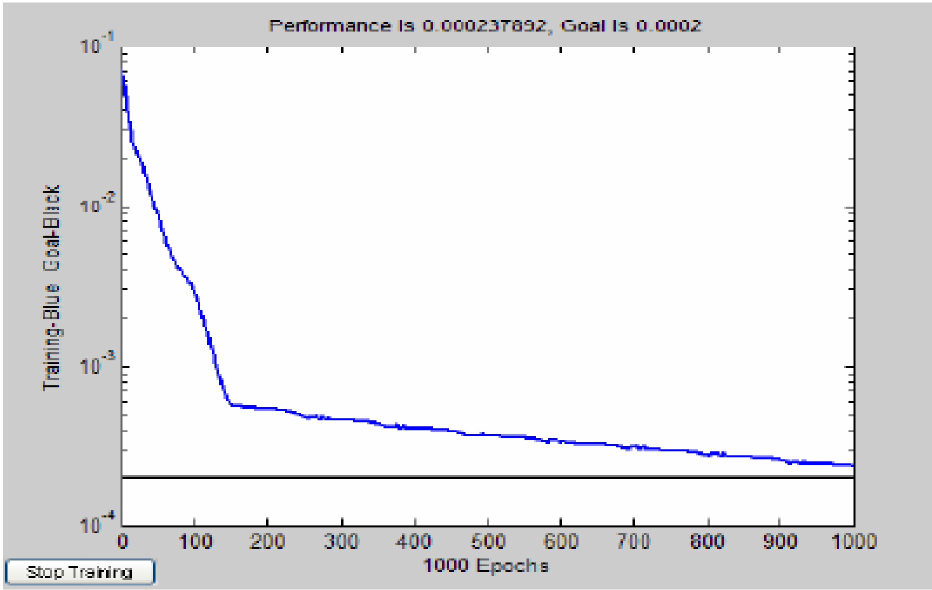
Theinput_testvaluethatwewanttopredictshould beputintothenetworkwhichhas beenbuiltup.Input_test hastwovectors:[2113222.5]and[547343 41]

Theconcretecodecanbewrittenasthefollowing:
an=sim(net,input_test)



Puretheerroracceleration

Errorsare0.1075,and0.1261.Ifdelevingneuron functionofoutputlayerbetogsig,andtraining functionistraingdx,whiletheotherparametersar enotchanged.



Logsig error acceleration

Genetic algorithm applied in neural network on the one hand is used for optimizing the structure of ANN, on the other hand to learn the weight of ANN; genetic algorithm has taken over traditional algorithms. The concrete steps can be written as the following:

(1) Coding. The problem we should solve is that decision variables must be changed into string structure data, because actual decision variables cannot be handled in genetic algorithm.

w1(k)=x(inputnum (1)+k)

w2(k)=x(hiddennum (1)+k+inputnum hiddennum)

B1(1)=x(inputnum hiddennum +hiddennum outputnum)+1) ;

B2(1)=x(inputnum hiddennum +hiddennum outputnum +hiddennum)+1)

(2) Initialize the population. We set the number of population is 30.

(3) Compute fitness of individuals in present population. The goal value can be stated as follows

$$E = \frac{1}{2}\sum_{k=1}^{m}(u^k - y^k)^2$$

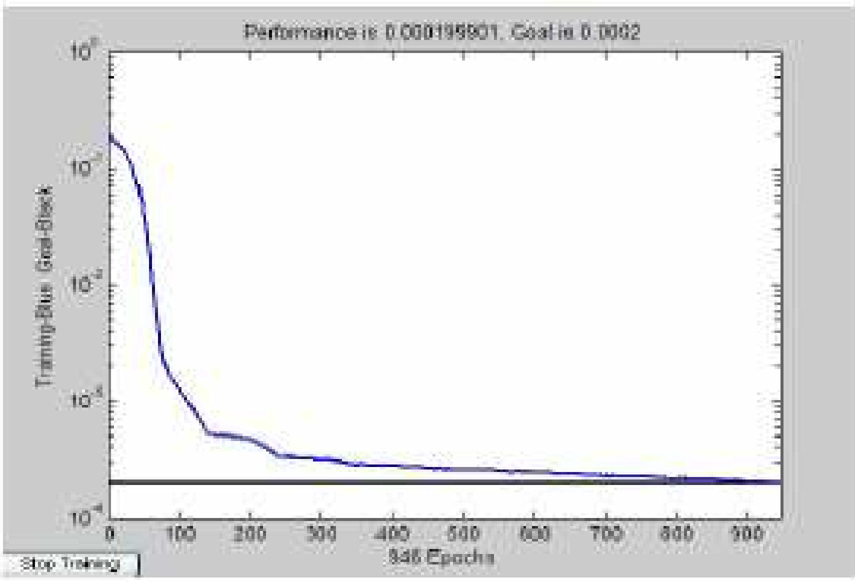
(4) Selection. In the first step, the selection mechanism computes a selection probability according to which the sampling algorithm generates an intermediate population of parents in order to generate new variants by recombination and mutation. The main method is to put the current population onto a roulette wheel, such that the size of each chromosome corresponds to its selection probability. The intermediate population is

generated from identical copies of chromosomes sampled by spinning the roulette wheel Nitin else Here we use genetic toolbox.

(5) Recombination. Instead of every parent chromosome, only a fraction of parent population is selected at random according to the crossover probability Pc is subjected to recombination. Genetic toolbox has the concrete code.

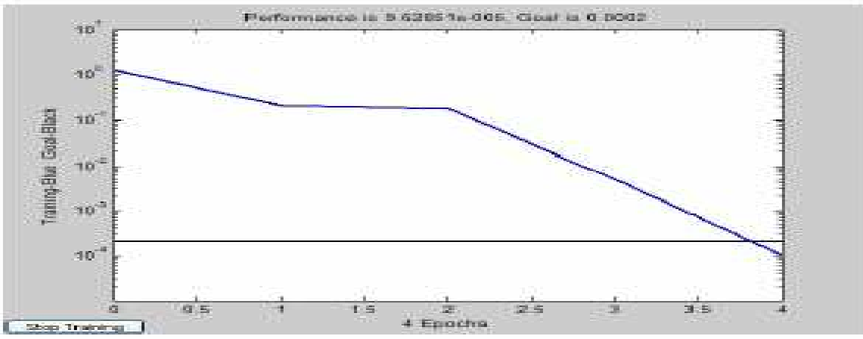
(6) Mutation. In Genetic Algorithm s, mutation plays the role of a background operator which arbitrarily inverts one or more bits of a parent chromosome to increase the structural variability of the population. Mutation provides a means to restore lost or unexplored genetic material and prevent the Genetic Algorithm from premature convergence to sub-optimal solutions. In this sense, mutation performs an exploration of the search space. The evolution process can be seen as follows picture.

Due to the steps that we have stated in the above, we can get the results by the parameters which we have chosen from the above. Clearly, it is sink into local minimal value while use the Back Propagation only, even if the parameters are totally different. However, the combination of Genetic Algorithm and Back Propagation can handle this problem which can be seen as figure



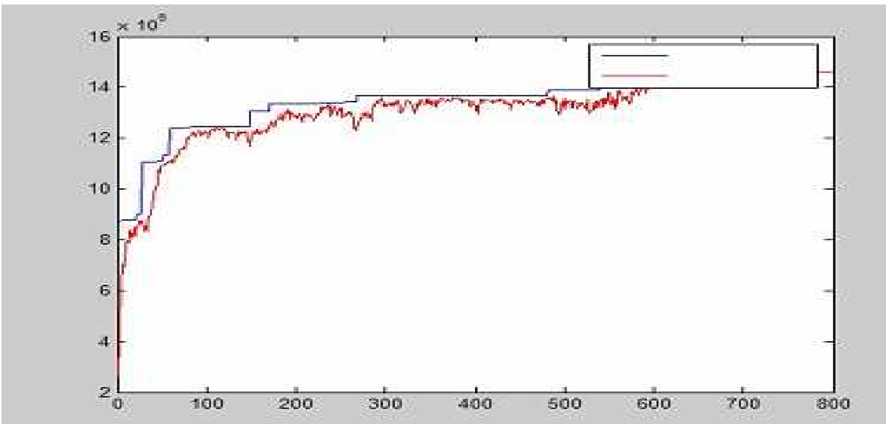
Logsig error acceleration by Genetic Algorithm optimization

It can be seen from figure 3 that genetic algorithm can speed up convergence and handle local minimum problem. The errors are 0.0069, and 0.0263. At this moment, compared with BACK PROPAGATION, the error is low er and the generalization ability is better. Therefore, Genetic Algorithm - Back Propagation model is fit for Construction Cost Estimate. Even though the parameters of figure have been used for comparing, we can see that the figure.



Pure the error alteration by Genetic Algorithm

Error curve in Back Propagation network descends qu ickly, while error curve in Genetic Algorithm -Back Propagation network descends gently. But after finding the optimum number, the error curve in Genetic Algorithm -Back Propagation nd escends rapidly, and the error is lower than Back Propagation network. The errors are 0.0492, and 0.1264. Genetic algorithm optimizing figure is figure



Genetic algorithm optimizing

Blue line stands for optimum path, and red line sta nds for mean value path of each generation. It tells that each generation constant ly rise the value that survival of the fittest, consequently the effect of optimal would b e obtained.

Conclusions: An optimization method for construction cost estim a te is presented. The method is called Genetic Algorithm -Back Propagation . Genetic Algorithm -Back Propagation model adopts the algorithm that chooses the optimized of form multiply dot in solution space simultaneously. Then a model of B ack Propagation algorithm is used to search the result. This method not only can avoid l ocal minimum and bad generalize ability problem, but also can overcome the weakness of long training time when Genetic

Algorithm find the optimum in random way. It is an effective and efficient method. The simulate results show that the system model of GA-BP model has a great improving generalization ability, and it can give a reasonable response when the new input data is coming. Therefore Genetic Algorithm-Back Propagation model has got a higher reliability, and it can be used in construction cost estimate. However, this method has shortcoming, it runs a long time, and the error has not attained the ideal degree. How to raise speediness and efficiency of cost estimate is problem that we need to investigate.

References

1. H.R.B.Donald, S.Remer, "Estimating the cost of ordering a cost estimate," Int. Production Economics, vol. 66, no. 1, pp. 101-104, 2000.
2. E.Y.Gwang-Hee Kim, Sung-Hoon An, Hun-Hee Cho, Kyung-In Kang, "Neural network model incorporating a genetic algorithm in estimating construction costs," building and environment, vol. 39, no. 3, pp. 1333-1340, 2004.
3. Zhe-Xue Ge, "neural network and realization by matlab 2007[M]," Electronic Industry Press, 2007.
4. Ji-Hai and Shi-Wei Ye, "neural network theory," machine Industry Press, 2004.
5. M Zhang, Y Hu, T. Wang et al, "An optimal predicting method based on improved genetic algorithm embedded in neural network and its application to petroleum analysis," Sheng Wu Yi Xue Gong Cheng Xue Za Zhi, vol. 26, no. 6, pp. 1186-90, Dec. 2009.
6. R.E.D. Randall S. Sexton, John D. Johnson, "Toward global optimization of neural networks: A comparison of the genetic algorithm and backpropagation" Decision Support Systems, vol. 22, no. 2, pp. 171-185, 1998.
7. Jing-Wen Tian, "artificial neural network algorithm research and application[M]," Beijing Institute of Technology press, 2006.
8. Ying-Jie Lei, "matlab genetic algorithm toolbox and application[M]," Xidian University, 2006.
9. F.H.O. scar cordon, frank hoffman, luis magdalena, "genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases," advances in fuzzy system-applications and theory, vol. 19, pp. 488 pages, 2001.
10. Xiao-Ping Wang, "genetic algorithm-theory, application, and software realization," Xi'an Jiaotong University, 2002.
11. L.D.David Montana, "Training Feedforward Neural Networks Using Genetic Algorithms," Machine Learning, vol. 89, no. 1, pp. 762-767, 1989.
12. F.G.O.Cordón, F.Herrera, F.Hoffmann, and L. Magdalena, "Ten years of genetic fuzzy systems: current framework and new trends" Fuzzy Sets and Systems, vol. 141, no. 1, pp. 5-31, 2004.
13. M.S.Wen-der Yu, "A neuro-fuzzy computational approach to constructability knowledge acquisition for construction technology evaluation" Automation in Construction, vol. 8, no. 5, pp. 539-552, 1999.