Evolving Artificial Neural Networks Using Simulated Annealing-based Hybrid Genetic Algorithms

Huawang Shi
Hebei University of Engineering, Handan, P.R.China
Email: stone21st@163.com

Wanqing Li
Hebei University of Engineering, Handan, P.R.China
Email: stone21st@sohu.com

Abstract—**Artificial neural networks are among the most effective learning methods currently known for certain types of problems. But BP training algorithm is based on the error gradient descent mechanism that the weight inevitably fall into the local minimum points. It is well known that simulated annealing (SA) and genetic algorithm (GA) are two global methods and can then be used to determine the optimal solution of NP-hard problem. In this paper, due to difficulty of obtaining the optimal solution in medium and large-scaled problems, a hybrid genetic algorithm (HGA) was also developed. The proposed HGA incorporates simulated annealing into a basic genetic algorithm that enables the algorithm to perform genetic search over the subspace of local optima. The two proposed solution methods were compared on Rosenbrock and Shaffer function global optimal problems, and computational results suggest that the HGA algorithm have good ability of solving the problem and the performance of HGA is very promising because it is able to find an optimal or near-optimal solution for the test problems. To evaluate the performance of the hybrid genetic algorithm-based neural network, BP neural network is also involved for a comparison purpose. The results compared with genetic algorithm-based indicated that this method was successful in evolving ANNs.**

Index Terms—**BP neural network, genetic algorithms, hybrid genetic algorithms, simulated annealing, global optimal**

I. INTRODUCTION

Neural network learning methods provide a robust approach to approximating real-valued, discrete-valued and vector-valued target functions. Artificial neural networks are among the most effective learning methods currently known for certain types of problems. But BP training algorithm is based on the error gradient descent mechanism that the weight inevitably fall into the local minimum points; Genetic Algorithm (GA) is good at global searching, and search for precision appears to be partial capacity Inadequate. Therefore, the present work intends to integrate ANN with GA[1-4] to determine properly the weights of neural network, making up for the defects of BP algorithm. The application of genetic algorithm into artificial neural network in this paper is regarded as the process of searching for optimum in the weight space. Genetic algorithm is a randomized search algorithm borrowing ideas from natural selection mechanism and genetic mechanism of living nature to acquire an optimal or sub-optimal solution. However, simple GA is difcult to apply directly and successfully to a larger range of difficult-to-solve optimization problems. So, in this paper, The genetic operators are carefully designed to optimize the neural network, avoiding premature convergence and permutation problems. And with the momentum to solve the slow convergence problem of BP algorithm. To evaluate the performance of the genetic algorithm-based neural network, BP neural network is also involved for a comparison purpose.The resualts indicated that Gas and with momentum were successful in evolving ANNs.

It is well known that Simulated annealing (SA) and genetic algorithm (GA) are two global methods and can then be used to determine the optimal solution of NP-hard problem. Genetic algorithm has a wide range of practicality, it can handle any form of objective function and constraints, whether it is linear or non-linear, continuous or discrete, in theory, have access to the optimal solution. Therefore you must keep in mind that genetic algorithms are not always the best choice. Sometimes they can take quite a while to run and are therefore not always feasible for real time use. They are, however, one of the most powerful methods with which to (relatively) quickly create high quality solutions to a problem.

However, in it's practical applications, genetic algorithm has the more serious problem such as : premature convergence, poor local optimization ability, and slow convergence and not convergence to global optimal solution. In recent years, many scholars try to improve genetic algorithms, such as improving the encoding scheme, fitness function, genetic operator design. However, these improvements are all make in internal of the genetic algorithm and it has been proved that it is unable to overcome these shotcomimgs effectively

doi:10.4304/jsw.5.4.353-360

To tackle this problem, a Simulated Annealing (SA) algorithm is proposed to solve the model. SA's major advantage over other methods is an ability to avoid becoming trapped at local minima[1-5]. The algorithm employs a random search which not only accepts changes that decrease objective function f, but also some changes that increase it. The algorithm works efficiently on a neighbourhood search within solution space, acceptance probability, and inferior solutions to escape from trap (i.e., local optimal solution). Numerical examples are solved to check for the efficiency and validity of the SA algorithm.

The rest of this paper is organized as follows: articial neural networks, simulated annealing (SA) and genetic algorithm (GA) are described in Section2. Section3 describes the defects of classic BP Algorithm and improvement of genetic algorithm then followed by the improved BP neural network algorithm. In Section3, The shortcomings and improvement of genetic algorithm with introduction of simulated annealing was introduced. The experimental test of results on using simulated annealing-based genetic algorithm to solve of classic optimization problems, the Rosenbrock function and Shaffer function global optimal problem has been studied in section5 and some discussions are presented in Section5. Finally, To test the power of simulated annealing-based genetic algorithm , we used the algorithm on an optimize problems. This work is described in a final section, followed by our conclusions.

II. Brief Introductions to Algorithms

A. Artificial Neural Networks

Articial Neural Networks(ANNs) are composed of simple elements that imitate the biological nervous systems. In the last few decades, signicant research has been reported in the field of ANNs and the proposed ANN architectures have proven the ireciency in various applications in the field of engineering. Articialneural Neural Networks focus primarily on computing and storing information with in a structure composed of many neurons. Because NN simitate the human brain in terms of learning, recall and generalization, they are usually designed to solve non-linear or ill-structured problems. The structure of a neural network of most commonly used type is schematically shown in Fig.1, Fig.2 and Fig.3. It consists of several layers of processing units (also termed neurons, nodes). The input values are processed within the individual neurons of the input layer and then the output values of these neurons are forwarded to the neurons in the hidden layer. Each connection has an associated parameter indicating the strength of this connection, theso called weight.

The NN model frequently used is multilayer perceptron learning with error back-propagation. In the present research work, the sequence with which the input vectors occur for the ANN straining is not taken into account, thus they are static networks that propagate the values to the layers in a feed-forward way. The training of the neural networks is performed through a back-propagation algorithm. Ingeneral, the back-propagation algorithm is a gradient-descent algorithm in which the network weights are moved along the negative of the gradient of the performance function.

B. Genetic Algorithms

Gas a search technique that imitates the natural selection and biological evolutionary process were first established on a sound theoretical basis by Holland [6-8]. Genetic algorithm has a wide range of, particularly in combinatorial optimization problems and they were proved to be able to provide near optimal solutions in reasonable time[9], it can deal with arbitrary forms of the objective function and constraints, whether it is linear or non-linear, continuous or discrete, in theory, have access to the optimal solution. However, in practical applications of genetic algorithm to demonstrate the more serious question is "premature convergence" problem, less capable local optimization, the late slow convergence and can not guarantee convergence to global optimal solution and so on. In recent years, many scholars try to improve genetic algorithms, such as improving the encoding scheme, fitness function, genetic operator design. However, these improvements are all make in internal of the genetic algorithm and it has been proved that it is unable to overcome these shortcoming effectively. The most common type of genetic algorithm works like this: a population is created with a group of individuals created randomly. The individuals in the population are then evaluated. The evaluation function is provided by the programmer and gives the individuals a score based on how well they perform at the given task. Two individuals are then selected based on their fitness, the higher the fitness, the higher the chance of being selected. These individuals then "reproduce" to create one or more offspring, after which the offspring are mutated randomly. This continues until a suitable solution has been found or a certain number of generations have passed, depending on the needs of the programmer[7].The basic structure of a GA is shown in Fig.4

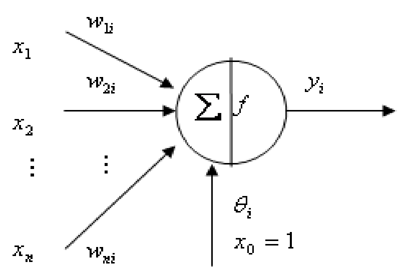

Figure1. Note how the caption is centered in the column.

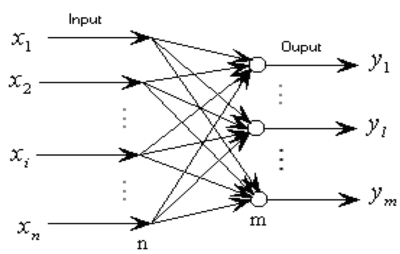

Figure 2. The single layer of feedforward networks

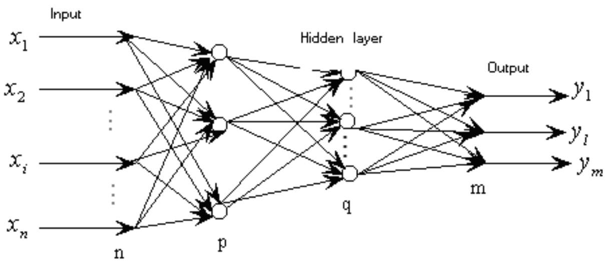

Figure3. The multi-layers of feedforward networks.

C. Introduction to Simulated Annealing

The algorithm is based upon that of Metropolis et al. [1-5], which was originally proposed as a means of finding the equilibrium configuration of a collection of atoms at a given temperature. As its name implies, the Simulated Annealing (SA) exploits an analogy between the way in which a metal cools and freezes into a minimum energy crystalline structure (the annealing process) and the search for a minimum in a more general system. SA's major advantage over other methods is an ability to avoid becoming trapped at local minima. The algorithm employs a random search which not only accepts changes that decrease objective function f, but also some changes that increase it. The basic principle of SA is as follows

Given initial temperature $t_0 = t_{maz}$ and the initial point x , iterative steps $k = 0$ to calculate the function point value $f(x)$;

Randomly generated disturbance Δx , a new point $x' = x + \Delta x$ was got, to calculate the new point of function values $f(x')$, and function margin value; $\Delta f = f(x') - f(x)$

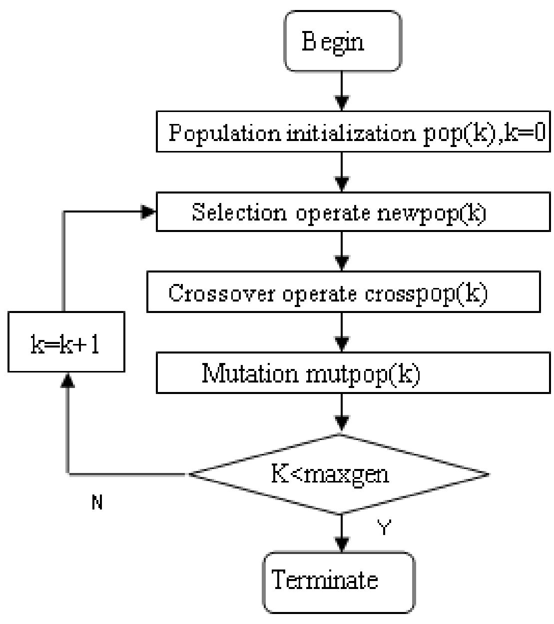

Figure 3. The Basic Structure of a Genetic Algorithm.

If $\Delta f \le 0$, then accept the new solution as the the initial point of next simulation, if $\Delta f > 0$, then calculate the probability of a new point $P(\Delta f) = \exp(-\Delta f / T_k)$, resulting evenly distributed random number r in interval $[0,1]$, r $[0,1]$, if $P(\Delta f) \ge r$, then accept new simulation points as the next initial point; otherwise, abandon the new point, still use the original point as the initial point of the next simulation.

The following three formulas are the main model for the simulated annealing algorithm.

$$P_{ij} = \begin{cases} G_{ij}(t)A_{ij}(t) & \forall j \ne i \\ 1 - \sum_{l=1, l \ne i}^{|D|} G_{ij}(t)A_{ij}(t) & j = i \end{cases} \tag{1}$$

$$G_{ij}(t) = \begin{cases} 1/|N(x_i)| & j \in N(i) \\ 0 & j \notin N(i) \end{cases} \tag{2}$$

$$A_{ij}(t) = \begin{cases} 1 & f(i) \ge f(j) \\ \exp(-\Delta f_{ij} / t) & f(i) < f(j) \end{cases} \tag{3}$$

III. BRIEF INTRODUCTIONS TO ALGORITHMS

A. Defects of Classic BP Algorithm

The neural network based on BP algorithm through a simple compound function of neurons, allowing the network with nonlinear mapping capability, such a network without feedback, are belonging to feed forward networks. Despite the improvement in theory and a wide range of practicality deciding its important position in the artificial neural network, but the algorithm itself has also an unavoidable flaws. The main problems can be summarized as follows:(1) local minimum point, (2)Slow convergence, (3)It is difficult to determine hidden layers and hidden layer nodes, (4)The poor learning and memory of network.

B. The Improved BP Neural Network Algorithm

From the BP neural network algorithms and genetic algorithms speak its own characteristics, BP training algorithm is based on the error gradient descent mechanism that the weight inevitably fall into the local minimum points; GA is good at global searching, and search for precision appears to be partial capacity Inadequate. So, in this paper, the GA was used to optimize the weights of neural network. Before the genetic algorithm is conducted, a group of solutionswere generated at random. Among them, the individuals with higher fitness were selected according to the principle of survival of fittest to do selection, crossover and mutation[10]. After evolution from generation to generation, it will converge to the fittest individual at last. That is the solution of the problem. In this paper, it was improved for four aspects.

1) According to the defect slow convergence of BP algorithm, in this paper, we adopt the measures of addind momentum to solve the problem, the principle is as follows: In practice, the choice of learning step is very important, in the other parameters remain unchanged, while learning rate big fast convergence, but too much may cause instability; small oscillation can be avoided, but the slow convergence to resolve this contradiction is the easiest way to join the "momentum of." Momentum is in each weight regulation to add a proportional to the weight of the previous regulation of the amount of value for:

$$\Delta w(n+1) = \eta \frac{\partial E}{\partial w(n)} + \phi \Delta w(n) \tag{4}$$

Where ϕ is the momentum coefficient, the general range is $[0,0.9]$. After adding momentum, the regulation of weight is makd toward in direction of the bottom of the average, that is, momentum is played the role of buffer and gently, so that network convergence speed is regulated.

2) The Best numberof hidden nodes p : The network performance is impacted by the number of hidden nodes. when there is too many hidden nodes, it will lead to e-learning for too long, can not even convergence; and when there is less had hidden nodes, network has poor fault-tolerant capabilities. Best hidden nodes number p can be refered to the following formula:

$$p = (n+m)/2 + c \tag{5}$$

Where n is the number of input nodes; m is output nodes; c is constant between 1~10.

GA-based BP algorithms: BP network is one of the most widely used artificial neural network, is now widely used in signal processing, pattern recognition, system identification, adaptive control and other fields. But its easy in a local minimum point is the Achilles heel of BP algorithm, and the genetic algorithm as a global search algorithm, and BP algorithm, has many advantages, such as the most easily into the local advantages, in the error function can not be micro-or no gradient information, particularly ineffective However, the genetic algorithm is also unable to avoid the shortcomings, such as genetic algorithm local search capabilities, it is difficult to select parameters, such as defects. Based on the above theory, this paper GA algorithm BP algorithm and the respective strengths of the two organic combination of the completion of the common neural network weights and threshold adjustments.

C. Experimental Test

The substance of XOR is a function mapping problems, and easy to fall into local minimum point, Thus it is often used to test the performance of neural networks. In the paper, we select a fully connected single hidden three-layer feedforward articial neural network (ANN), with input nodes $n = 2$, hidden nodes $p = 5$, output nodes $q = 1$, and learning samples is shown in Table 1

TABLE I
LEARNING SAMPLES

x_1	0	0	1	1
x_2	0	1	0	1
y	0	1	1	0

Neural network control parameter analysis In the genetic algorithm algorithm , control parameters are the size of the main groups- $popsize$, cross-probability- p_c, mutation probability- p_m and evolution of algebra- $epoch$. the choice of parameters will affect the ultimate performance and efficiency of genetic algorithm.

1)Determining the groups size: Group size of the operation of genetic algorithm objects. If the groups are too small, the sample reflected information are insufficient. Individual groups help us to identify the diversity of the overall optimal solution, otherwise prone to the phenomenon of precocious puberty, that is the convergence of local optimal solution. But the group size is too large, would also lead to the calculation of each generation groups increased volume may make the convergence rate is extremely slow. As shown in Fig.5. When the population number increase, the effective and efficient getting better and better, but a lot of time is wasted. for XOR problem, the best groups size is adoptas $popsize = 40$.

2)Determining the crossover probability: If Crossover probability is big, individual update soon. But then it also brings disadvantages that destroyed excellent individuals. Under normal circumstances, the value of p_c is 0.5~0.8. As shown in Fig.6. , when the crossover probability was 0.5, the algorithm convergence rate is rather slow; when the crossover probability was 0.95, the network error obviously have a lot of oscillation, only when cross-check the probability was about 0.8 , the computing effect was best.

3) Determining the mutation probability: The value of mutation probability are often very small, typically 0.001~ 0.1. As shown in Fig.7. The mutation probability p_m =0.04 was the best .

D. Neural Network Control Parameter Analysis

1)Determining Learning rate: Learning rate is too large, although the error can decline rapidly, but there may be great oscillation, if too small learning process is very slow. As shown in Fig.8, the initial learning $\eta = 0.8$ was the best rate.

2)The training results of momentum impact: Fig.9 and Fig.10 are the results of without momentum and with the momentum To change the default, adjust the template as follows.

IV. HYBRID GENETIC ALGORITHM DESCRIPTION

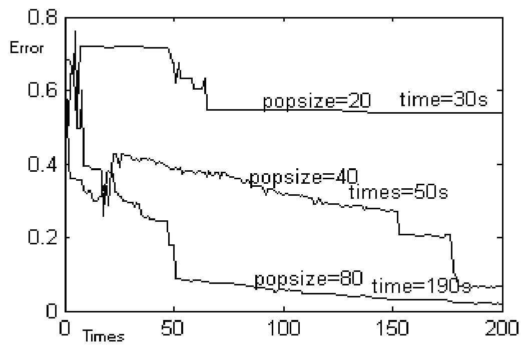

Figure 7. The analysis effective chart of population size

A. The shortcomings and improvement of Genetic Algorithm

Genetic algorithms are one of the best ways to solve a problem for which little is known. They are a very general algorithm and so will work well in any search space. Genetic algorithms tend to thrive in an environment in which there is a very large set of candidate solutions and in which the search space is uneven and has many hills and valleys. True, genetic algorithms will do well in any environment, but they will be greatly outclassed by more situation specific algorithms in the simpler search spaces[7].

SA's major advantage over other methods is an ability to avoid becoming trapped at local minima. The algorithm employs a random search which not only accepts changes that decrease objective function f, but also some changes that increase it.

B Hybrid Genetic Algorithm with introduction of Simulated Annealing

Based on the above discussion, combining with the advantages of genetic algorithm and simulated annealing algorithm and avoid their shortcomings, in this paper, hybrid genetic algorithm (HGA) is presented. The hybrid genetic algorithm takes genetic algorithm computing processes as the main processes, integrating of simulated annealing mechanism in which to further adjust and optimize the groups.The algorithm is designed as follows:

- Given population size $popsize$ and initializing operating parameters, running algebra $k = 0$, the initial temperature is $t_k = t_0$, then randomly generates initial population $pop(k)$.

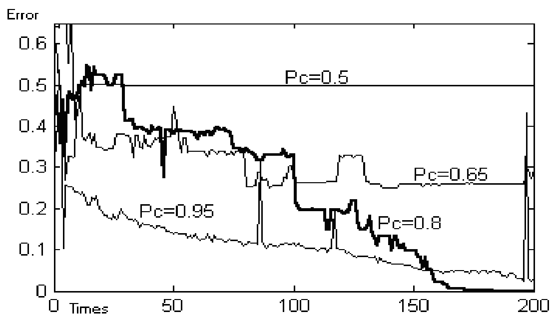

Figure 5. The optimized effect by crossover rate

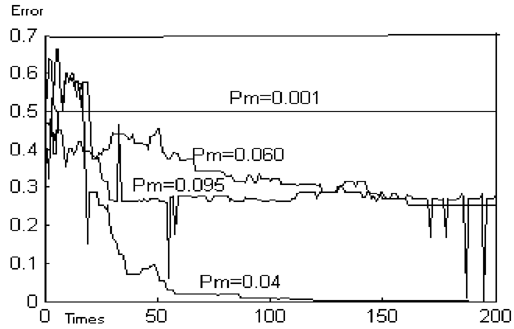

Figure 6. The optimized effect by mutation rate.

- In the domain of each chromosome $i \in pop(k)$ of groups $pop(k)$ the state $j \in N(i)$ was randomly selected according to acceptance probability of acceptance of simulated annealing:

$$A_{ij}(t_k) = \min\{1, \exp(-\frac{f(j) - f(i)}{t_k})\} \qquad (6)$$

- The fitness function $newpop1(k)$ was calculated as follows:

$$f_i(t_k) = \exp\{-\frac{f(i) - f_{\min}}{t_k}\} \qquad (7)$$

Where $f_{\min}$ is the minimum value in $newpop1(k)$. Through the probability distribution determined by the fitness function, population $newpop\ 2(k)$ was generated from of randomly selection $popsize$ chromosomes in $newpop1(k)$.

- Crossover operator. In accordance with a two-vector weighted average calculated as follows:. Were used to cross-convex, affine and linear cross-cross. After choosing the experiment in this article, the best stocks after cross. In this paper, when $\lambda_1 + \lambda_2 \le 1.1$, experiment show that the cross-linear results was best.after Crossover operator $cropop(k)$ was got.

Mutation operator. In accordance for non-uniform mutation, the specific operation, seen in[9]. After mutation, groups $mutpop(k)$ was got.

Let $t_{k+1} = \alpha t_k$, $t_{k+1} = \alpha t_k$, $t_{k+1} = \alpha t_k$, If meet the the

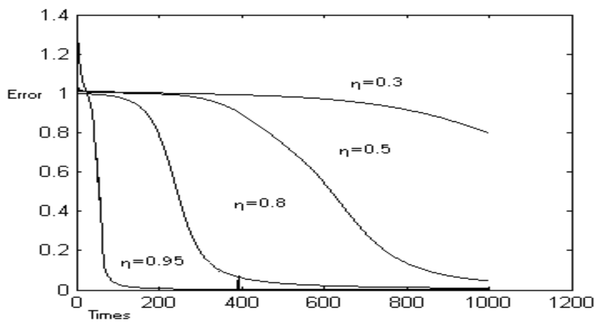

Figure 4. The optimized effect by leaning rate

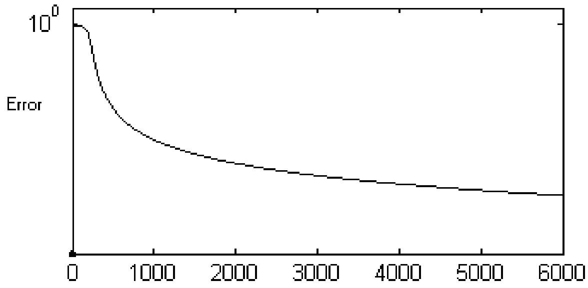

Figure 8. With momentum

rules of termination , end. otherwise return to (2). The calculation flow chart is shown in Fig.11.

V. HYBRID GENETIC ALGORITHM EXPERIMENTAL TEST

A Example1

Considering the global maximum value calculation of Rosenbrock function:

$$f(x_1, x_2) = 100(x_1^2 - x_2)^2 + (1 - x_1)^2$$
$$\text{s.t. } -2.048 \le x_i \le 2.048 \quad i = 1,2 \qquad (8)$$

The global Optimal value of the function in the defined region is 3904.819, the Optimal point is (-2.048,-2.048). Comparing SGA with improved genetic algorithm in this article , the results are shown in Table 2.

Through the experiment, the best function values from each generation are increased with the algebra and its trends are shown in Fig. 12 and Fig. 13.

B Example2

Considering the global minimum value calculation of Shaffer function:

$$f(x_1, x_2) = 0.5 + \frac{\sin^2\sqrt{x_1^2 + x_2^2} - 0.5}{[1.0 + 0.001(x_1^2 + x_2^2)]^2} \qquad (9)$$

$$s.t. \quad -100 \le x_i \le 100 \qquad i = 1,2$$

The global optimal value of the function in the defined region is 0, the Optimal point is (0,0). Comparing SGA with improved genetic algorithm in this article , the results are shown in Table 3.

Through the same experiment, the trend of Optimal objective function value searched in each generation

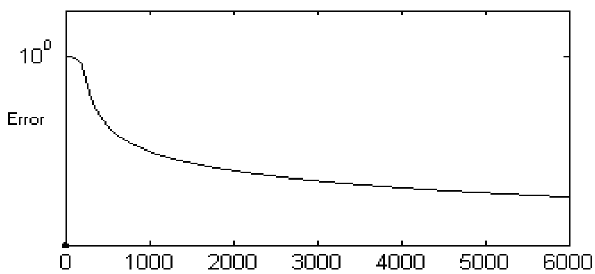

Figure 9. Without momentum

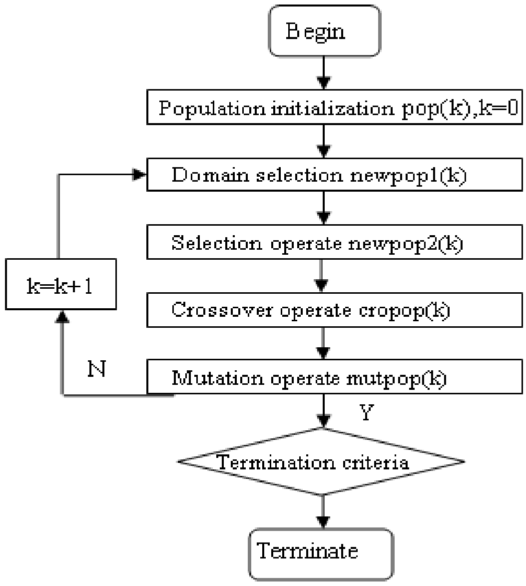

Figure 11.The Basic Structure of HGA

changing with the increasing generation are shown in Fig.14 and Fig.15. The goal values in Fig.14 and Fig.15. are the opposite number of the original goal function.

VI. APPLICATION

Using the algorithm proposed in this paper, calculate example in[22]. The optimize results of the objective function was 159.34, is superior to the result of 185.67 in literature [23][24].

A. Simulation

BP algorithm with the classic, traditional genetic algorithms and genetic With momentum improvement of hybrid neural network algorithm to train the network, the learning rate was 0.01, and expectative error was 0.001. Based on GA with momentum, the network training parameters are as follows:

GA parameters:

$popsize = 40$

$W = [-10,10]$

$pc = 0.8$

$pm = 0.04$

$epoch = 6000$

BP algorithm parameters

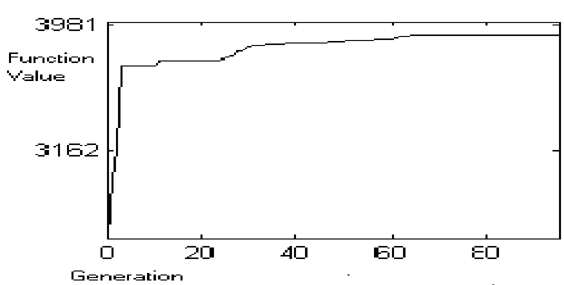

Figure 12. The evolution course of HGA

TABLE I.
COMPARISON OF RESULT (GENERATION : $\max gen = 100$ TIMES OF TEST :10)

Algorithm	Optimal value	Average optimal value	Optimal point	Average calculating time(s)
SGA	3896.1806	3876.7192	（-2.0363，-2.0017）	60
HGA	3905.8427	3903.7366	(-2.04798，-2.04778)	72

TABLE II.
COMPARISON OF RESULT (GENERATION: $\max gen = 300$,TIMES OF TEST:10)

Algorithm	Optimal value	Average optimal value	Optimal point	Average calculating time(s)
SGA	0.2327	0.4796	（6.2801，17.8602）	80
HGA	0.0000	0.0151	(0.0098，-0.0077)	96

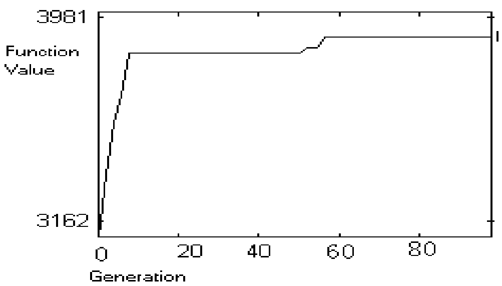

Figure 13. The evolution course of SGA

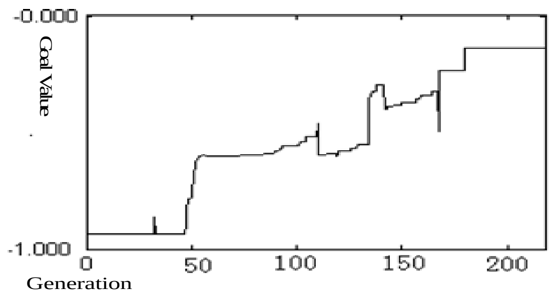

Figure 15. The evolution course of HGA

$$\eta = 0.8$$
$$\gamma_1 = 1.05$$
$$\gamma_2 = 0.7$$
$$\phi = 0.5$$

The results of the simulation algorithm comparing with HGA are shown as Fig.16. And the optimize evolution curves is also shown in Fig.16.

The algorithm presented in this paper increases the choosing scope of operations, to avoid in the early emergence of the phenomenon of super-chromosome of a simple genetic algorithm and increased the local search capabilities, and can escape from local minimum points. At the same time it can speed up the fitness function in the latter which increasing the speed of convergence. From comparison results in Fig.16, fully describes the advantages of improved algorithm in this article.

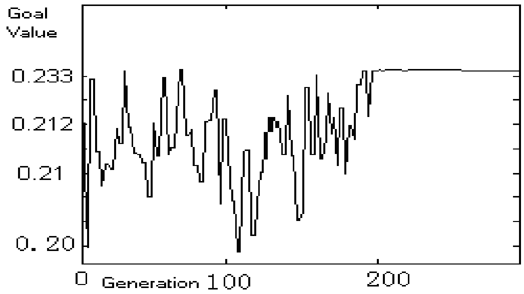

Figure 14. The evolution course of SGA

VII. ONCLUSIONS

Artificial Neural Network (ANN) has outstanding characteristics in machine learning, fault, tolerant, parallel reasoning and processing nonlinear problem abilities. It offers significant support in terms of organizing, classifying, and summarizing data. It also helps to discern patterns among input data, requires few one, and achieves a high degree of prediction accuracy racy. These characteristics make neural network technology a potentially promising alternative tool for recognition, classification, and forecasting in the area of construction, in terms of accuracy, adaptability, robustness, effectiveness, and efficiency. Therefore, quality application areas that require assessment and prediction could be implemented by ANN.

However, in practical applications of genetic algorithm

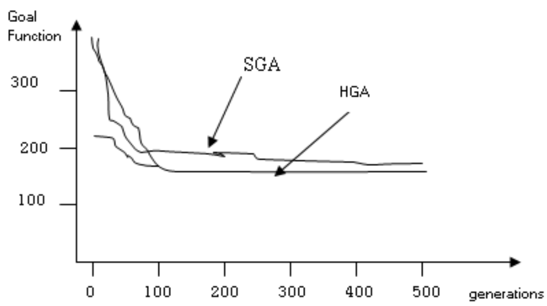

Figure 16. Evolution curves compared with SGA

to demonstrate the more serious question is "premature convergence" problem, less capable local optimization, the late slow convergence and can not guarantee convergence to global optimal solution and so on. In recent years, many scholars try to improve genetic algorithms, such as improving the encoding scheme, fitness function, genetic operator design. However, these improvements are all make in internal of the genetic algorithm and it has been proved that it is unable to overcome these shortcoming effectively.

This article first introduced in neural networks and genetic algorithms briefly, and pointed out the respective merits of two algorithms, as well as the combination of two algorithms possibilities. After pointing the relative limitations in its application, overcome the genetic algorithm limitations in accordance with the idea of simulated annealing to. proposed HGA incorporates simulated annealing into a basic genetic algorithm that enables the algorithm to perform genetic search over the subspace of local optima. The two proposed solution methods were compared on Shaffer function global optimal problems, and computational results suggest that the HGA algorithm have good ability of solving the problem and the performance of HGA is very promising because it is able to find an optimal or near-optimal solution for the test problems. Finally, examples of comparing simple genetic algorithm and improved genetic algorithm, the results proved that the optimization effect of the genetic algorithm with introduction of annealing mechanism significantly was better than simple genetic algorithm and improved BP neural network algorithm.

REFERENCES

[1] Jesus Fraile-Ardanuy, P.J.Zufiria. Design and comparison of adaptive power system stabilizers based on neural fuzzy networks and genetic algorithms. Neuro computing 70 (2007) 2902-2912.

[2] K.M.Saridakis , A.J.DentsorasJ. Integration of fuzzy logic, genetic algorithms and neural networks in collaborative parametric design. Advanced Engineering Informatics 20 (2006) 379-399.

[3] J.Hertz,A.Krogh,R.G.Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley, New York, 1991.

[4] J.H. Holland, Outline for a logical theory of adaptive systems, J. Assoc. Comput. Mach. 3 (1962) 297–314.

[5] KasabovKN. Foundation of neural networks, fussysystems and knowledge engineering. MIT Press; 1996.

[6] K.M.Saridakis, A.C.Chasalevris, C.A.Papadopoulos , A.J.Dentsoras. Applying neural networks, genetic algorithms and fuzzy logic for the identication of cracks in shafts by using coupled response measurements. Computers and Structures 86 (2008) 1318-1338..

[7] Masato Sasaki, Mitsuo Gen, Fuzzy multiple objective optimal system design by hybrid genetic algorithm, Applied Soft Computing3 (2003) 189-196.

[8] GoldbergD. Genetic algorithms in search optimization and machine learning. Massachusetts: Addison-Wesley;1989.

[9] DimarogonasAD. EXPERTS, a fuzzy logic, neural network structured expert system shell for diagnosis and prognosis users manual. StLouis, Missouri, USA: Clayton Laboratories.1988.

[10] Cheng Xiangjun. Neural network principle and application[M]. Beijing: Defence Industrial Press, 1995.

[11] Wang Yongqing. Artificial intelligent principle an dmethod [M]. Xi'an: Xi'an Jiaotong University Press. 2002.

[12] SHI Huawang. The Risk Early-warning of Hidden Danger in Coal Mine Based on Rough Set-neural network. Proceeding of the 2nd International Conference on Risk Management and Engineering Management.November 4-6,2008.pp314-317

[13] Wenjuan Liu, Qiang Liu , Feng Ruan , Zhiyong Liang , Hongyang Qiu. Spring back prediction for sheet metal forming based on GA-ANN technology. Journal of Materials Processing Technology .187-188 (2007) 227-231.

[14] Odom M D, Sharda R A. Neural network for bankruptcy prediction. International Joint confrence on Neural Network,pp17-70,june,1990.

[15] Zhijian Hou, Zhiwei Lian, YeYao, Xinjian Yuan. Cooling-load prediction by the combination of rough set theory and an articial neural-network based on data-fusion technique[J]. Applied Energy 83(2006)1033-1046.

[16] Fletcher, D & Goss, E. (1993). Forecasting with neural networks: an application using bankruptcy data. Information and Management, 24,159-167.

[17] Computational Intelligence in Scheduling (SCIS 07), IEEE Press, Dec. 2007, pp. 57-64, doi:10.1109/SCIS.2007.

[18] SHI Shi-liang , LI Run-qiu , XIE Jian-xiang .Theoretical study on safety assessment indexes system of coal mines. journal of coal science & engineering Vol.9 No.2 Dec. 2003, PP63—69.

[19] Saaty, Thomas L. Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World. Pittsburgh, Pennsylvania: RWS Publications.

[20] Michael Skinner. Genetic Algorithms Overview http://geneticalgorithms.ai-depot.com/Tutorial/Overview.html.

[21] Schwefel, H-P., Evolution and Optimum Searching, Wiley Interscience, John Wiley & Sons, New York, 1995.

[22] Booker, L. B., Improving Search in Genetic Algorithms, pp. 61-73, Genetic Algorithms and Simulated Annealing (L. Davis, editor), Pitman, London, 1987

[23] J.Tian, J.S.Kou, M.Q.Li. Optimizing the construction arrow network techniques using Genetic AlgorithmsJ. Systens Engineering-Theory&Practice1999(5)

[24] Huawang Shi. Evolving artificial neural networks using an improved GA. Proceeding of the 2nd ISECS, 2009:475-478.

[25] Wu T H, Chang C C, Chung S H.A simulated annealing algorithm for manufacturing cell formation problems. Expert Syst Appl 2008; 34(3): 1609-17.