A Study on Genetic Algorithm Optimization of Artificial Neural Networks

Hongjun ZHONG *, Guiying HE, Ying HUO, Chuiyi XIE

Education Department, Shaoguan Radio and Television University, Shaoguan, Guangdong , 512026, China.

***Abstract* — Artificial neural networks and evolutionary computation techniques have rapidly developed in recent years and these methods represents a trend of mutual integration. A genetic algorithm is a global, optimized and 'probing' algorithm. There are many simple generic and robust features in parallel search methods for global optimal solution with high convergence precision and speed. Genetic algorithms and neural networks can be used in combination to improve the intelligibility of the model at a higher level. The combination of genetic algorithm optimization and neural networks can extend greatly the ability of neural networks by using the fast convergence properties of genetic algorithms. They supplement and strengthen each other, which can improve their ability to solve practical problems. This paper proposes the use of genetic algorithms to optimize neural networks. An optimization strategy is proposed to solve some engineering optimization problems by combining neural networks and genetic algorithms. Back Propagation, BP, networks are used for function approximation to obtain implicit function expressions to the problem, then we optimize the output of the network by a genetic algorithm.**

Keywords - Genetic algorithms, neural networks, optimization

I. INTRODUCTION

Genetic algorithms and neural networks are the key technologies of computational intelligence techniques in 21st century. The speed of modern computer's calculation is hundreds of times as much as the speed of human's brain. Problems of programmable computing can be solved rapidly and effectively for those clear features and clear rules of inference. Its values, concluded in terms of arithmetic and logic operations, are accurate, and then it can greatly expand the capacity of the human brain. But for a long time, people are trying to understand the nature of the working mechanism of the human brain and thinking about how to construct artificial intelligence system to meet human's yearning issues so as to mimic the human's brain function and complete the work which is similar to the human brain. Mr. Qian Xuesen thinks that logical thinking, imagination and inspiration are three basic ways of thinking in human's brain. it's possible to get good results in solving practical problems.

The combination of genetic algorithm optimization Genetic algorithm (GA) and neural network can extensively map the ability of neural networks and fast convergence properties of genetic algorithms. They supplement each other--they complete and strengthen each other, which can effectively improve the ability of solving practical problems and adapting local optimization to obtain the optimal solution .It has become one of the hot topics of interdisciplinary academics. BP network maps its function and approximation ability has made extensive use in genetic algorithms to solve common problems of adaptability and optimization in real life and scientific research in the field which has been applied widely. Therefore, from using optimization to improve the traditional BP neural network, genetic algorithm can make good use of BP neural network algorithm which is a powerful function and a ability of local search , and which can play out the global search function of genetic algorithm. Through combining complementing each other, it not only can realize the complementary advantages, but can perfectly solve the problem of all kinds of practical engineering as well. It also has the very good value in research and reference.

II. BP LEARNING ALGORITHM

BP network learning algorithm uses a back propagation learning algorithm which is a teacher's typical learning of error correcting . The learning process is realized by two processes--back propagation error propagation and signal components. BP neural network is a typical type of forward neural network with some advantages such as self-study habits, adaptability and using Yi Shi etc, especially the good nonlinear approximation ability and generalization ability. BP neural network model is composed of input layer neurons, neurons in hidden layer and output layer neurons. The same layer of neurons is not connected while adjacent layer neurons are all interconnected. The traditional BP neural network has a strong recognition to determine the type of causality--certain relationship of input-output. The basic idea is that the forward spread of the incoming input samples from the input layer, each hidden layer is transmitted to the output layer after layer by layer processing. If the actual output with the expected output of the output layer (teacher signal) does not match, the transferred back propagation will be wrong; the back propagation of error is the output error in some form from the hidden layer to the back propagation of input layer , and error apportioned to all unit layers. Each layer unit can obtain an error signal. The error signal that is used as the basis of this correction unit is connected between

the weights and the threshold unit[1,2].Such forward propagation and back propagation is carried out again and again by leading to constantly adjust the network parameters and being carried out to reduce the output error of the network until it reaches an acceptable level.

BP algorithm consists of four processes: input mode; Feed forward neural network model is a neural network model of one-way flow of information containing M neurons; N-input single-layer; single-layer feed-forward neural network model in front of the structure as shown.

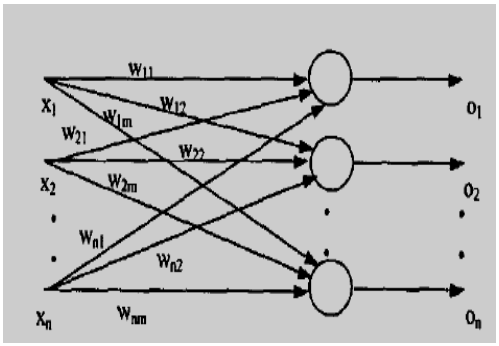

Figure 1. Feed forward neural network model structure.

The study sample P is $(x_1, x_2 ....., x_n)^T$, the desired output for y_p . When the desired output and the actual output of the network are different, the presence of the output error ep is defined as the following:

$$e_p = \frac{1}{2}(r_p - y_p)^2 \tag{1}$$

The above error is defined DEPLOYED to the hidden layer, there is

$$e_p = \frac{1}{2}(r_p - y_p)^2 = \frac{1}{2}\left[y_p - f\left(\sum_{j=1}^{m} v_j o_{jp} \right) \right]^2 \tag{2}$$

It can be further expanded to the input layer,

$$e_p = \frac{1}{2}\left[y_p - f\left(\sum_{j=1}^{m} v_j o_{jp} \right) \right] = \frac{1}{2}\left\{ y_p - f\left[\sum_{j=1}^{m} v_j f\left(\sum_{j=1}^{m} w_{ij} x_{jp} \right) \right] \right\}^2 \tag{3}$$

It's not difficult to see from the above formula, several output error e_p network is actually a function of $w_{ij} v_j$ and $\theta_j \left(i = 1,2,...\pi,\ j = 1,2,...m \right)$. The network's training purpose is to reduce the output error to adjust these parameters through the network. According to the gradient descent principle, the objective function contains free variables. The objective function's value is the fastest decline in the direction of the negative gradient direction of the objective function of variables. Thus, we only amend them along e_p negative gradient direction of free parameters which can reduce the error function. Specific process is:

$$\Delta w_{ij} = -\eta \bullet \frac{\partial e_p}{\partial w_{ij}},\quad \Delta w_{ij} = -\eta \bullet \frac{\partial e_p}{\partial v_j},\quad \Delta \theta_j = -\eta \bullet \frac{\partial e_p}{\partial \theta_j} \tag{4}$$

In this formula, the negative sign is on behalf of the gradient descent; $\eta \in (0,1)$ is the learning rate which control to modify the value of each step. From the above equation to equation (4), it is not difficult to obtain samples of $\Delta\omega_\eta, v_j$ and θ_j from all learning. According to the above procedure, we can adjust the weights and thresholds of the network [3,4,5]. Once we obtain the average error ,those fixed connection weights and the threshold in all samples can definite as follows:

$$E = \frac{1}{P} \bullet \sum_{P=1}^{M} \mathrm{e}_p \tag{5}$$

P is the total number of learning samples there.It's Easy to see that the average error E measures all learning network. At this point, an iterative process is completed. BP neural network is one of the neural network recognition pattern which is widely used. Figure 2 shows the schematic structure of BP neural network.

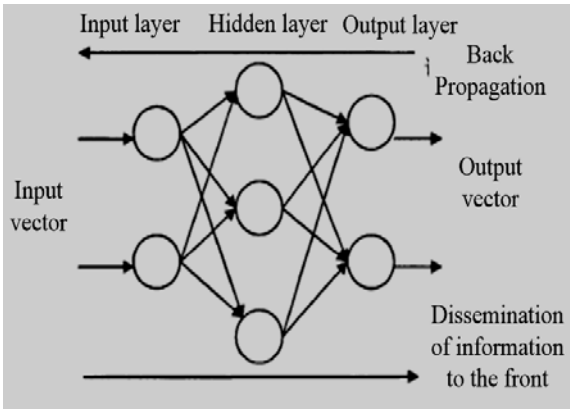

Figure 2. Schematic structure of BP neural network

III. THE BASIC PRINCIPLES OF GENETIC ALGORITHMS

Genetic Algorithms (GcncfieAlgorithm, GA) is an evolution which based on natural genetic population algorithm to explore an efficient mechanism. It is first proposed in 1975 by American scholar Holland and Genetic algorithm find qualified multiple solutions at the same time. It's easy to converge to the global optimal and overcome the shortcoming of BP neural network that is easy to fall into local convergence. Besides, the genetic algorithm is a kind of heuristic search with higher search efficiency, which can quickly determine the approximate scope of optimal solution[6] and have obvious advantages.

It abandoned the traditional search methods of simulating the natural process of biological evolution, and use the approach of artificial evolution to search the target space randomly . It may be the solution to the problem of a group of individuals or chromosomes and encoded into a string of symbols for simulating each individual form of Darwinian genetic selection and the natural elimination process of biological evolution and it based on the population genetics of repeated operation (genetic, crossover and mutation). According to a predetermined objective fitness function which was evaluated for each individual, rules of survival of the fittest and survival of the fittest evolutionary, it continues to be better groups while the global parallel search find a method to search optimization groups with the best individual and to meet the requirements of the optimal solution. Genetic algorithms are a class of randomized algorithms, but it is not a simple random. It can effectively utilize some ways of processing information to improve the quality of the solution string, which is similar to natural evolution, genetic algorithm act on chromosome genes and chromosomes will look a good way to solve problems. When solving the problem, genetic algorithms will be ignored naturally. It needs only a pair of each chromosome generated by the algorithm to evaluate whether chromosome have more chance to satisfy the applicability of reproducing. Using GA to research was asked and determine the best hypothesis by searching the space of candidate hypotheses. In GA, the "best hypothesis" is defined as fitness (Fitness). For the best hypothesis, the fit measure is the number of pre-defined in the current problem. For example, if the task is to learn the input and output in the training examples by giving an unknown approximation of function to this function, the fitness can be defined as the assumption of accuracy on the training data[7,8,9]. If the task is to learn chess strategy, it can be defined as an individual fitness and win rate of individual chess for the other in the current population. We can Select from the current population comprising a number of assumptions in the next generation. These assumptions are selected by using a probabilistic approach while in the selection probability it can be assumed that ht is calculated by the following formula:

$$\Pr(ht) = \frac{Fitness(ht)}{\sum_{j=1}^{p} Fitness(ht)} \tag{6}$$

After this operation generates a new cross-member, a new generation of group already contains the required number of members. Next, we can randomly selected from these members and the members of a certain percentage of Shim random variation. Thus, the GA algorithm performs a stochastic parallel columnar search according to the fitness function found in good assumption. Hybridization of two random individuals are selected as parents of individuals according to judgment of the probability of cross-hybridization such as cross. The intersecting position is random when generating two new individuals which is shown in Figure 3.

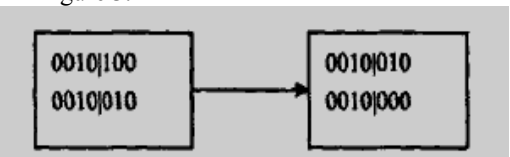

Figure 3. Crossover operation ("I" for the hybrid position).

IV. GENETIC ALGORITHM OPTIMIZATION OF NEURAL NETWORK ANALYSIS

In summary, the back-propagation algorithm (BP) solve the hidden layer weights to fix the problem, but it is a nonlinear function of the extreme demand, which is likely to fall into local minimum gradient method. Therefore it cannot guarantee to converge to the global minimum points. BP optimization algorithm has precise characteristics, and genetic algorithm has a strong macro-search capability and good global optimization. This paper considers to overcome the deficiencies of BP network, the first genetic algorithm and BP neural network training by using genetic algorithms before the combination of genetic algorithm optimization after narrow the search so as to re-use BP network to accurate solution and achieve the purpose of global optimization fast and efficiently. With the scheme, we can effectively overcome the disadvantage that the back - propagation algorithm is easy to fall into local minimum in the solving process and convergence's speed is relatively slow which can play advantages that search the global optimal solution with a high convergence precision and convergence's speed in the parallel genetic algorithm to find out the global optimal solution quickly and the effect is very obvious.

This article describes the GBA as a global natural selection and genetic optimization algorithm which is based on mathematical algorithms evolutionary mechanism system and abstracting from the mechanism of natural selection , crossover and mutation genetic operators and calculate encoded string of these three parameter. Through paralleling to the different areas of the parameter space, the search in its turnover in generations find the optimal global solution

direction. The disadvantage is that the GA converges to a certain extent, + through crossover and mutation operator generates significantly reduce the probability of higher adaptability of the individual so that the convergence rate is restricted. You can not accurately determine the location of the optimal solution--the local search space without the ability to fine-tune and so on. Research shows that the current number of conventional genetic algorithm is not necessary for the best method of solving a problem. And the unique genetic algorithm and knowledge of the problem integrated into hybrid genetic algorithm may produce an excellent way to solve the performance, which also continues to improve the search performance of genetic algorithm and provides a new way of thinking and a lot of practice has proved Genetic algorithm an effective tool for global search "4. The chapter describes how to use genetic algorithms to train the network which takes full advantage of the global search and strong local search ability and strong features of BP neural network algorithm.

Combining artificial neural networks with genetic algorithms are considered as a promising way to reproduce the behavior of smart um usually in some global optimization algorithms. It can solve the neural network training in local minima problems. In order to solve the slow convergence of neural network training that the problem is easy to fall into local minimum, the current study focused on two aspects: First, there are premature convergence problems in initial weights of neural network optimization and genetic algorithm itself , so these method is still not guaranteed that after-dimensional network training will not fall into the local minimum area; Secondly, the use of neural network evolution (Evolving Neural Networks, referred ENN), completely replace BP learning by genetic algorithms and use gradient descent method to avoid defects. However, due to genetic algorithm itself feature of weaker local search makes evolutionary neural network requires a wide range of initial weights' area resulting in increased complexity which makes the trained neural network generalization become not ideal.

V. INTEGRATION OF NEURAL NETWORKS AND GENETIC ALGORITHMS

Neural network ensemble is a finite number of neural network with a problem. The integrated output and the input sample can be decided by it so as to constitute an integrated neural network output under the sample . (See Figure 4) Perrone and Breiman also proved this theory. From the mid-1990s , a lot of researchers began to study this field and the theoretical study of neural network ensemble received great attention. The theory and application of the results continue to emerge and make the neural network ensemble become an international machine learning and neural computing profession a very active focus of research which attract more and more attention. Because the performance of the method is very obvious, it become the focus of neural computing and machine learning research, and has been successfully applied in various fields in recent years.

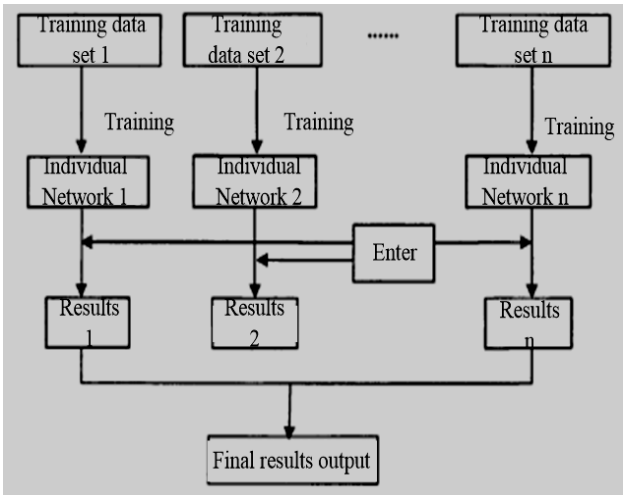

Figure 4. Schematic diagram of neural network ensemble.

Chen Gangue, from the University of Hong Kong, who proposed a correction method basing on neural network to determine the number of relationships and experimental results between the calculations of functional density theory recently. The number of such relationships are then used to eliminate the absorption spectra of organic molecules theoretical bias. Hutchison patients and others with six kinds of calculation methods assess the 60 organic molecules absorbing energy spectrum.

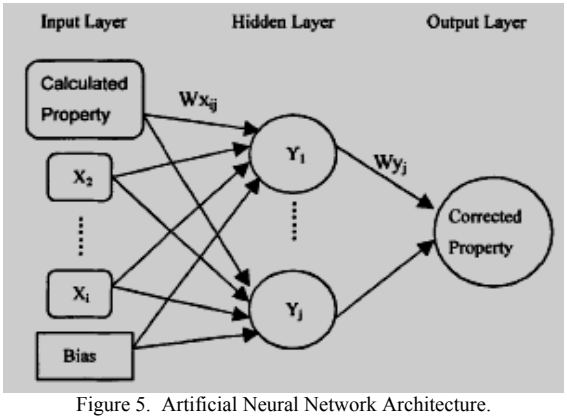

Figure 5. Artificial Neural Network Architecture.

These six kinds of methods are ZINDO / CIS, ZINDO / RPA, HF / CIS, HF / RPA, TDDFT / RPA and TDDFT. The conclusions show that TDDFT / CIS and TDDFT / RPA methods combined with linear regression method to get more accurate results. Chen Gangue, who used TDDFT / B3LYP to calculate the absorption of these 60 molecules artificial neural networks and multiple linear regression to correct the calculated the TDDFT / B3LYP results obtained from corrected neural network which is better than multivariate linear regression calibration results. In the process of neural network computing, the output of the neural network mainly

depends on the network connection weights and weights of these networks in the process of the form of the iterative training .The overall distribution of the weights of neural network connection contains all the knowledge of neural network system. Thus different network connection weights will lead to a different network output. However, the traditional neural network algorithm can't achieve such effect .

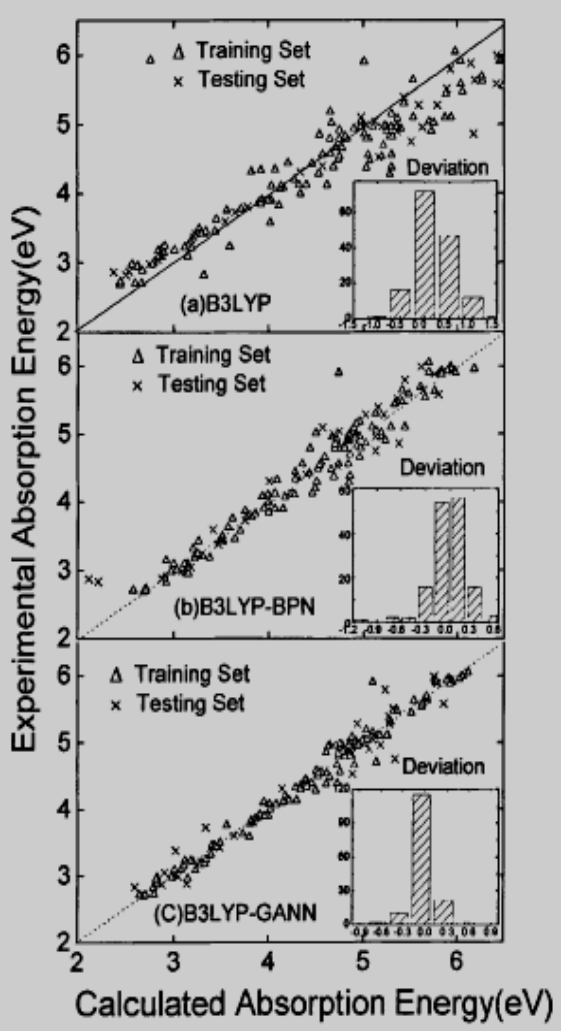

Figure 6. Experimental and theoretical values of 150 organic molecules: (a.c) DFT, BPN and GANN corrected calculated value :(unit: eV).

We adjust the number of hidden layer neurons changing l to 10 to determine the optimal neural network structure. We have found that the hidden layer neurons contain 5 steps to produce the best output. Therefore, we have adopted a neural network structure 10.5.1.

As what can be seen from Table 1, after the training , BPN rms error correction is dropped from the original 0.48 0.20 eV; For the test set, the error was down from 0.41 to 0.28 eV. For GANN correction method of training and test sets, errors were reduced to 0.14 and 0.24 eV. As what can be seen from the analysis, GANN correction methods improve training and test sets of density functional theory calculations. But there are still some data in the test set, and their relative error is higher, the main cause of this situation is that this type of data is rare. So in the process of training, the neural network of this type occur in the feature of training set so that the data cannot be effectively extracted. If you can provide more and better experimental data, the prediction accuracy GANN method can be further improved.

TABLE 1 . TDDFT / B3LYP / 6-31G (D), BPN AND GANN ERROR CORRECTED TABLE (UNIT: EV)

	TDDFT/B3LYP/6-31G(d)	BPN	GANN
Training set	0.48	0.20	0.14
Testing set	0.41	0.28	0.24
Overall	0.47	0.22	0.16

VI. CONCLUSIONS

This thesis proposes to improve genetic algorithm with BP neural network method test its performance through experiments. Genetic algorithms and neural networks are the key technologies of computational intelligence techniques in 21st century. Genetic algorithms are adaptive algorithm and does not have the other global optimization and implicit parallelism. It reflects a strong ability of solving problems. Neural networks are interconnected massively with parallel processor. The optimization problem can be solved by highly interconnected neurons and enhance the mutual complementarity between them.

To obtain more robust representation and the ability to solve practical problems, this paper attempts to make a combination of BP neural networks and genetic algorithms to solve engineering optimization problems to achieve - given effect. It shows that the method is feasible and effective. For data on BP neural network forecasting method and genetic algorithms, there are still many deficiencies which require further study and improvement. This article is mainly predicted by BP artificial neural network to achieve the number of nodes in the hidden layer model by adopting spreadsheet approach. How to effectively determine the science hidden layer nodes, require further study. In the calculation model, genetic algorithm play a very important role. However, this article doesn't have further discussion on genetic algorithm itself . Genetic algorithms should be a certain room for improvement so that the final forecast accuracy is greatly improved which has to research in the future yet. Because of their limited knowledge of the level of data collection and hasty limited time, there may be a lot of problems in this article. The text of the opinions and conclusions are to be further tested and refined and hope experts to make a correction.

ACKNOWLEDGMENTS

Guangdong Distance Open Education Research Fund Project (YJ1333); science and technology innovation project of Guangdong Province (2013KJCX0168).

REFERENCES

[1] Lu RQ,"Knowledge Engineering&Knowledge Science of the Century,"Journal of Tsinghua University, pp. 53-58, 2011.

[2] Hu L H, Wang X J, Wong L H, et a1. "Combined first Principles calculation and neural Network correction approach for heat offormation," Journal of Chem Phys,vol. 26, pp. 635-639,2013.

[3] Wang XJ, HuL H, Wong L H, eta 1. "A combined first-principles calculation and Neural Networks correctiona pproach for evaluating Gibbs energy offormation,"Journal of Moi Simul, vol. 13,pp. 144-149, 2011.

[4] Alexandridis A, Patfinos B Sarimveis H, et a1. "A two-stage evolutionary algorithm for variable selection in the development of RBF neural network models,"Journal of Chemometrics and Intelligent LaboratorySystems, vol. 33, pp. 75-80,2011.

[5] Janson D J, Frenzel J F. "Training product unitn eural networks with genetic algorithms,"Journal of IEEE Expea, pp. 474-477,2010.

[6] Xu XL. "Genetic algorithm optimization of rotating machinery neural network forecasting model ,"Journal of Mechanical Engineering, vol. 24, pp. 266-269,2003.

[7] Irikura K K, Frufip D J. "Computational Thermochemistry : Prediction and Estimation of Molecular Thermodynamics,"Journal of Washington, DC : American Chemical Society,vol. 14, pp. 247-251,2010.

[8] Holland J H. "Adaptation in Natural and Artificial Systems,"Journal of Ann Arbor : The University of Michigan Press, vol. 27, pp. 347-351,2011.

[9] Cramer C J. "Essentials of Computational Chemistry : Theories and Models,"Journal of West Sussex : Wiley. vol. 11, pp. 27-31,2012.