The Forecast Model of the Total Population in China Based on Genetic Neural Network

Ning Wang
College of Science ,Zhongyuan University of Technology, Zhengzhou, China
Email: 13663851069@163.com

Xijie Zang
College of Science ,Zhongyuan University of Technology, Zhengzhou, China
Email: jet509@163.com

***Abstract*—In order to avoid shortcomings of the standard BP network algorithm, the forecast model of total population which is based on genetic neural network is constructed . Involving the advantages of GA and BP. The algorithm can simultaneously complete genetic selection within a solution space to find the optimal points. Then the BP algorithm searches the best optimal result from those points by the direction of negative gradient. Thus it not only can avoid the BP algorithm into a local minimum and slow convergence etc, but also can overcome long-search time, slow shortcomings of the GA caused by searching optimal solution in a similar form of exhaustive. Simulation indicates that the algorithm is more accuracy than the standard BP algorithm, faster in calculation and very well in applicability. Genetic network is trained to predict the total population in China in 2015, which is 137102, compared with 136782 in 2014, showing a good rise.**

***Index Terms*— genetic algorithm, neural network, the total population, time series**

I. INTRODUCTION

In the new historical period, we are facing the problem of total population, which is not only the quantity of population, but also the complexity of population control and family planning policy. A special coordination and research organization, and certain financial support is the basic condition for the study. The main research questions include: population, population quality, population distribution and resources environment, and population aging.

At present, the prediction methods in population are more. Zhou predicted annual growth rate of patent applications with a nonlinear chaotic model[1]; Xu used gray system theory to predict the patent trend on the research and application[2]; Rao proposed Wave-type time series which is based on the patent application, and dynamic analysis in authorization data [3].

This genetic neural networks (GA-BP) combined with time series, deals with data of the total population in China, and forecast it. We Propose a prediction method based on GA-BP, exploring the quantitative prediction of total population for further research.

II. THE BASIC PRINCIPLES OF GENETIC NEURAL NETWORK

A. BP neural network and its shortcomings

BP neural network has nonlinear mapping ability of the powerful. So It is used as a kind of artificial neural network model,which is very mature from theory to practice.In theory, simulation has proved that any nonlinear system can be used [4-5]. Figure 1 shows the BP network model with one hidden layer.

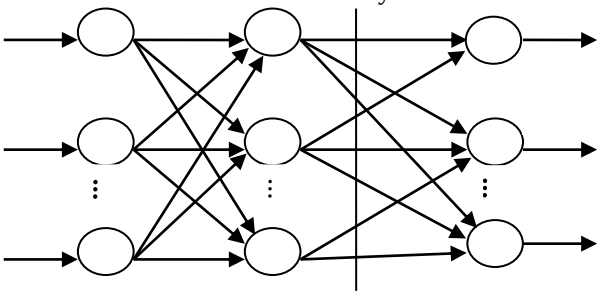

Figure 1. BP network model with a hidden layer

Because the error of BP algorithm reduces by the direction of negative gradient, it is easily landed itself in the local minimum point. when the number of training samples is too large or the relationship between input and output is too complicated, the convergence speed becomes very slow. At the same time, more requirements are needed for network structure and initial value in the algorithm. when these parameters are unreasonable, the network is vulnerable to shocks and even has non-convergence phenomenon [6].

B. Genetic Algorithm, GA-BP

Genetic algorithm simulates the model evolution and natural selection process which is first proposed by Professor J.Holland, the University of Michigan by the United States in 1975. It is the general for a series of parallel random search heuristic algorithm. Holland and Goldberg demonstrated that genetic algorithm can maintain the solution structure of good performance and improve the "adaptive capacity"of the future. Thus it evolved to the optimal solution or sub-optimal solutions. Genetic algorithm has been successful in the applications in a number of areas.

The main features of genetic neural network strategy is searching strategy of group and information exchange

between individuals in group. The search does not depend on gradient information, and there are no special enquirements for solving the problem. The algorithm not only has the global search capability but also enhances local search capabilities. So the genetic algorithm is applied to deal with the complexity of nonlinear problems which is difficult to solve using traditional methods [7].

As a strong macro-genetic algorithm has a good global search capability and chemical properties, genetic algorithm is used to optimized the weights and thresholds of the BP neural network, until the fitness of the solution is no longer significant. That is to say, the solution quality is stable. This combination of parameters obtained by decoding the load is close to the best combination of application. We can use the BP algorithm of artificial neural network for further precise optimization, until the search to the optimal network parameters, and can obtain the optimal parameters accurately. Thus we can reach the global objective faster and efficiently, and can avoid local minimal problems.

C. The specific steps of genetic algorithm[8] :

1. Code: initial population, including the crossover scale, the crossover probability, mutation probability and the weights and threshold initialization; In this initial coding, the use of real coding: set the length of chromosome coding s = m * r + r * n + r + n; and in accordance with [WHI, WHO, BH, BO] to encode the sequence (WHI as the input layer to hidden layer weight vector, WHO for the hidden layer to output layer weight vector BH for the hidden layer threshold vector, BO threshold vector for the output layer). Set the initial population size.

2.Fitness value calculation: calculation of evaluation function for each individual, and the sort, press the select the network type of individual probability values:

$$P(i) = f/ \text{sum}[f(i)] \quad (1)$$

Fitness value for individual i which can be used to measure the sum of squared errors E:

$$F(i) = 1/E(i) \quad E(i) = \text{sum}[\text{sum}(d-y)^2] \quad (2)$$

Where i = 1 ..., n for the number of chromosomes; o = 1, ..., q as the output layer nodes; k = 1, ..., m the number of samples for the study; for teachers signal, the actual output for the network. By the formula (1) and then calculate the probability of selection by roulette wheel selection principle of the implementation of chromosome evolution operations.

3. The crossover probability of individual and cross-operation, to generate new individuals and there is no direct cross-replication of individual operations;

4. The use of new individuals generated by mutation probability;

5. New individuals into the population p, and calculate the new individual's evaluation function;

6. To determine whether the end of the algorithm, if found satisfactory the individual, the end of, or transferred to 3 computing the next round of genetic manipulation.

7. To optimize the initial value of the GA as the initial genetic weight, and BP algorithm to train the network until the specified accuracy.

8.Algorithm end. If the required performance indicators, will be the final decoding groups in the best individual you can get the optimized coefficient of the network connection weights.

III. THE PREDICTION MODEL BASED ON GENETIC NEURAL NETWORK

A. Data preprocessing

In this paper, the data of the total population in china is given from 1980 to 2014 in Table I.

TABLE I.
1980-2014 THE TOTAL POPULATION IN CHINA

year	The total population	year	The total population
2014	136782	1996	122389
2013	136072	1995	121121
2012	135404	1994	119850
2011	134735	1993	118517
2010	134091	1992	117171
2009	133450	1991	115823
2008	132802	1990	114333
2007	132129	1989	112704
2006	131448	1988	111026
2005	130756	1987	109300
2004	129988	1986	107507
2003	129227	1985	105851
2002	128453	1984	104357
2001	127627	1983	103008
2000	126743	1982	101654
1999	125786	1981	100072
1998	124761	1980	98705
1997	123626		

According to the total population and habits of the five-year plan, we take the amount of total population from 1980 to 1984, as input of the first sample, and take the data of it in 1985 as corresponding output. Then we take those from 1981 to 1985 as input of the second sample, and take the data of it in 1986 as the output of the second sample. In turn, 30samples were obtained.

Standardization of data for processing, the data specification between -1 to 1, matlab in premnmx () function can be realized. The data is divided into two parts. The first part of a total of 21 sets of data used for training. The second part of the 9 groups for authentication, the use of predicted value and actual value, calculated MAPE, the mean absolute percentage error [1].

B. parameter Design of genetic networks

The body of the secondary title. Part of the genetic algorithm, using real-coded, easy to get chromosome length s = 137; crossover probability is 0.3, select the probability is 0.8, mutation probability 0.05, population Algebra 100, population size 50;

BP algorithm is part of the input data, output data each 5 to 5-12-5 structure as a predictor of network topology, the input layer and hidden layer activation function between the "tansig", between the hidden layer and output layer The activation function for the "purelin", training function as "trainlm"; the largest number of 2000 training, convergence and accuracy of 1.0e-28 learning rate 0.3.

C. Genetic network compared with the standard BP network

Standard BP neural network converges slowly [9-10], or even difficult to converge, as shown in Figure 2

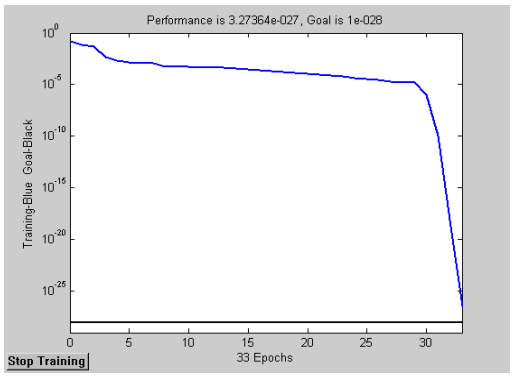

Figure 2. training process of standard neural network

Better genetic neural network approximation, Figures 3 and 4 respectively, the fitness function decline curve and error curve

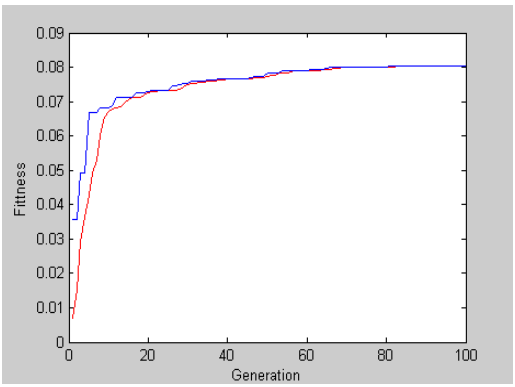

Figure 3. the curve of fitness function GA running for 100 generations

Mean absolute percentage error (MAPE) compared

Standard BP algorithm: MAPE = 20.65%, It shows that the prediction error of granting greater generalization ability is poor, the model used to predict the effect of authorized quantities less than ideal.

Genetic algorithms: MAPE = 8.53%, that with respect to the above model, the model prediction error is small, generalization ability, the simulation results are reliable, this model can be used to predict the amount authorized. Please do not revise any of the current designations.

Genetic network is trained to predict the the total population in China in 2015, which is 137102, compared with 136782 in 2014, showing a good rise.

IV. CONCLUSION AND OUTLOOK

Using the common artificial neural network learning, there are inherent shortcomings such as local minimal, slow convergence etc. In order to overcome the shortcomings,we proposed genetic neural network prediction model which integrated the advantages of GA

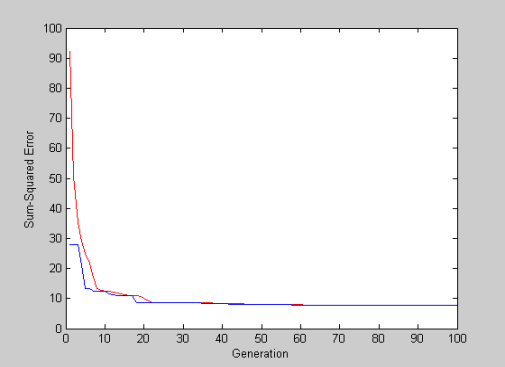

Figure 4. the decline curve of error GA run for 100 generations

and BP. The algorithm can simultaneously complete genetic selection within a solution space to find the optimal points. Then the BP algorithm searches the best optimal result from those points by the direction of negative gradient. Thus it not only can avoid the BP algorithm into a local minimum and slow convergence etc, but also can overcome long-search time, slow shortcomings of the GA caused by searching optimal solution in a similar form of exhaustive.

Verified through actual simulation, the proposed method of artificial neural network has better applicability than the average prediction accuracy. Its calculation has higher speed and it has ideal prediction. Another addition to genetic algorithms, simulated annealing algorithm can also choose the wavelet algorithm, particle swarm optimization, ant colony algorithm and principal component analysis, optimization of BP neural network.

ACKNOWLEDGMENT

This article is supported by The National Natural Science Fund(11401605), The National Natural Science Fund(11501279), and the scientific and technological project of Henan Provincefund (13210221014).

REFERENCES

[1] R.Zhou, J.Yu, “The chaotic forecast of annual growth rate of patent applications based on non-linear ravioli,” *Statistics and Decision*, vol.5, pp. 31–32, 2007.

[2] J.Xu, Y.Wang, “Prediction of the patent on the research and application based on grey system,” *Computer Applications* vol.22(4), pp.4–7, 2006.

[3] M.Rao, Y.Lin and G.Li, “Patent application and grant the amount of time series analysis,” *Operations Research and Management*, Vol.16(6), pp.157-161, 2007 .

[4] J.Guo, “Based on BP Neural Network Prediction of Chinese demand for iron ore,” *Mineral Research* , Vol .11, pp.41-44, 2008 .

[5] J.Hu, X.Tang, “BP algorithm of artificial neural network and its application,” *Information Technology*. Vol .4, pp. 1-4, 2004 .

[6] Y.Liu, L.Zhang, “BP network and RBF neural network implementation and performance comparison,” *Electronic Measurement Technology Rresearch and Design*, Vol .4, pp. 77-80, 2007 .

[7] P.Doganis, A.Alexandridis, P.Patrinose, et al, “Time series forecasting for short shelf-life food products based on artificial neural network and evolutionary computing”,

Journal Of Food Engineering, Vol .75 (2), pp.196-204, 2006.

[8] Sung-Kwun Oh, W.Pedrycz, "Multi-layer self-organizing polynomial neural networks and their development with the use of genetic algorithms," *Journal Of The Franklin Institute*, Vol.176(5), pp.475-489, 2006.

[9] K.Kim, I.Han, "Genetic angorithms approach to feature discretization in artificial neural networks for the prediction of stock price index," *Expert System with Application*, Vol .19 (2), pp.125-132,2000.

[10] C.Jin, "Based on genetic neural network prediction of cancer mortality rates," *Systems Engineering Theory& Practice*, pp.141-144,2000.

Ning wang: the author is born in Henan province in China in 1976. The author was graduated from Mathmatics department of NanKai University and received Bachelor degree in 1998. !n 2007, the author was graduated from FuDan University and received Master degree in financial engineering.