Prediction of Terrorist Attacks in China based on BP Improved Algorithm and GTD

Le Hong
School of Chongqing University, Chongqing 40044, China.
bgbgjbhl0906@163.com

Abstract. Due to the different national conditions, the driving forces and factors of national terrorist attacks vary. Therefore, this paper takes GTD China sample data as the research object to study and predict. The prediction process is as follows: On the basis of the BP network-based model for predicting the most dangerous areas, combined with the GTD sample data, the best number of nodes in the implicit layer of the prediction model is automatically selected by combining the empirical formula with the MATLAB program. Three improved BP algorithms are used to train the network model. The results show that the training error of Levenburg Marquardt algorithm is minimal and the convergence speed is fastest. Through the training and simulation of the model, it is proved that the model has high precision and can meet the requirement of practical application.

Keywords: BP network; GTD sample; number of terrorist attacks; prediction.

1. Introduction

Accurate prediction of the number of terrorist attacks in different regions is of great practical significance to prevent terrorist attacks, reduce casualties and financial damage, and fundamentally solve the harmfulness of terrorist attacks. Since the occurrence of terrorist attack is a nonlinear system affected by many factors, we can use BP neural network algorithm to predict it. However, as the number of hidden layer nodes in BP neural network is selected by empirical formula, the prediction accuracy is sometimes high and sometimes low, so this paper improves the method to determine the number of hidden layer nodes in BP neural network and the advantages and disadvantages of various BP algorithms are analyzed and compared.

2. BP Neural Network Algorithm

BP neural network is a multi-layer feedforward neural network based on error back-propagation algorithm (Back Propagation). It usually consists of input layer, hidden layer and output layer. The learning process is composed of two parts: the forward propagation of information and the back propagation of error. The specific algorithm process is as follows:

Suppose the input of the network is X, the output is Y, the expected output is T, the input neuron has m inputs, the hidden layer neuron has s nodes, the transfer function is f_1, the output neuron has n outputs, and the corresponding transfer function is f_2. The connection weight from the input layer to the hidden layer is W_{ji}(i= 1,2... s ;j= 1,2... m). The connection weight from the hidden layer to the output layer is k_i(i= 1,2... n). The threshold values of hidden layer and output layer are a_1 and a_2 respectively.

2.1 Forward Information Transmission Process:

a.Output of the i_{th} neuron in the hidden layer is:

$$y_{1i} = f_1(\sum_{i=1}^{s} w_{1j} + a_{ij})$$

b. The output of the k_{th} neuron in the output layer is:

$$y_{2k} = f_2(\sum_{i=1}^{n} w_{2ki} y_{1i} + a_{2k})$$

c.The error function is:

$$E = \frac{1}{2}\sum_{k=1}^{n}(t_k - y_{2k})^2$$

2.2 Reverse Propagation Process

a. The weight change from the ith input to the KTH output is:

$$\Delta w_{2ki} = -\eta \frac{\partial E}{\partial w_{2ki}} = \eta(t_k - y_{2k}) f'_2 \, y_{1i} = \eta \delta_{ki} y_{1i}$$

In the same way:

$$a_{2ki} = -\eta \frac{\partial E}{\partial a_{2ki}} = \eta \left(t_k - y_{2k}\right) f'_2 = \eta \delta_{ki}$$

b. The weight change from the JTH input to the JTH output is:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{1ij}} = \eta \sum_{k=1}^{n}(t_k - y_{2k}) f'_2 \, w_{2ki} f'_1 \, x_j = \eta \delta_{ij} x_j$$

In the same way:

$$\Delta a_{1ij} = \eta \delta_{ij}$$

Among them:

$$\delta_{ij} = e f'_1 ; e_i = \sum_{k=1}^{n} \delta_{ki} w_{2ki}$$

3. Establishment of Prediction Model of Terrorist Attacks Number

A large amount of data shows that the factors influencing the number of terrorist attacks in the region not only include the fact that whether there are ethnic conflicts and other disputes in the region, but also include economic level and political power, and there is a non-linear relationship between the frequency of terrorist attacks in cities and various influencing factors. According to GTD data analysis, the number of terrorist attacks has a strong correlation in space and time. Therefore, factors influencing the number of terrorist attacks include the number of attacks in the past 12 years, whether ethnic conflicts or other disputes are recognized, and the ranking of economic level in China in the recent times.

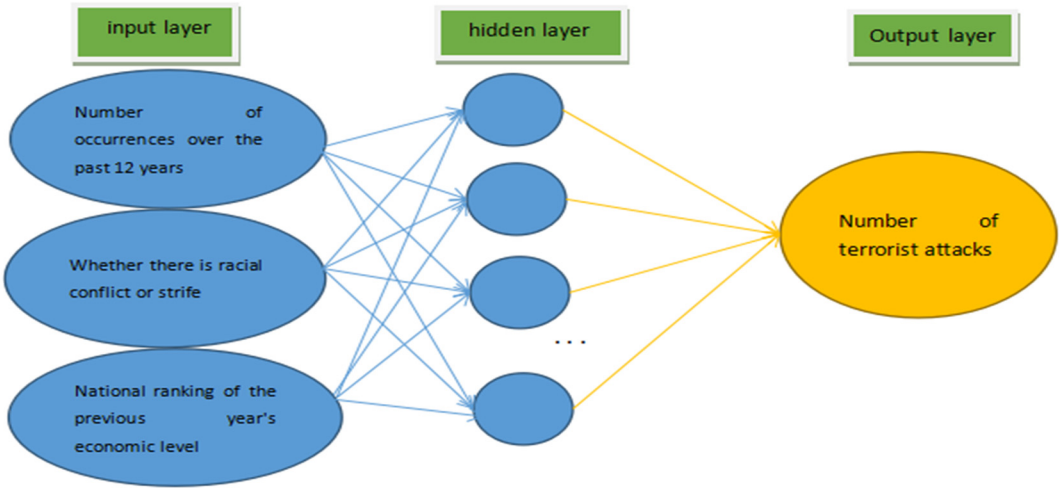

Fig. 1 BP artificial neural network for predicting the number of terrorist attacks

3.1 Selection of Samples

The source of the sample should be true and reliable and the selected sample should be representative. Therefore, the sample data of text is from GTD (table1).

Table 1. training and prediction model sample set2 Network structure design

	Number (2004-2016)	The existence of contradiction	Economic Ranking (2016)	Number (2017)
GuangXi	2	0(N)	19	0
JiangSu	1	0(N)	2	0
YunNan	4	0(Y)	22	0
GuangDong	4	0(Y)	1	1
Tibet	6	1(Y)	31	0
ShanDong	3	0(N)	3	0
ShangHai	3	0(N)	10	1
XinJiang	36	1(Y)	26	2

3.2 Network Structure Design

a. Determination of node points in the input and output layers

From the previous BP network diagram, the number of input nodes is 3, and the output is only one, I. E. the number of terrorist attacks Y.

b. Determination of hidden layer nodes

The determination of the number of nodes in the hidden layer is the core of the network structure design. If the number of nodes is too small, the convergence speed of the network is fast, but the accuracy is low. If the number of nodes is too many, the training time of the network will be long and the error will not necessarily reach the best. So far, there is no theoretical guidance to determine the number of hidden layer nodes. In most cases, empirical formula is used.

$$m = \sqrt{n + l} + a$$

In the formula, m is the number of hidden layer nodes, n is the number of input nodes, l is the number of output nodes, and a is constant, generally between 1 and 10. Based on this, this paper attempts to design a network with variable number of hidden layer nodes by combining empirical formula with computer software MATLAB program. Through error comparison, the optimal number of nodes is determined. The results of MATLAB operation are shown in Table 2.

Table 2. Network training errors

Layer Nodes	4	5	6	7	8	9	10	11	12	13
Error	0.0641	0.0918	0.0875	0.1080	0.0968	0.882	0.0631	0.0708	0.0831	0.0968

It can be seen from Table 2 that when the number of hidden layer nodes is 10, the training error reacher the lowest-0.0631, so the number of hidden layer nodes is determined to be 10, so the topological structure of the network model is (3-10-1).

3.3 Selection of BP Algorithm

For the performance of the common BP algorithm (momentum BP algorithm, additional momentum adaptive learning rate BP algorithm, Levenburg-marquardt algorithm), this paper uses the traingdm function based on momentum BP algorithm in combination with the sample data of the table 1. The levenberg-Marquardt theory and trainlm function, and the trainlm function of BP algorithm with momentum adaptive learning rate, are used to train the regional terrorist attack prediction model based on BP neural network. The training is based on the network prediction model. The most suitable BP calculation and training function is selected. The maximum number of iterations is 2000 steps, and the convergent target is 0.001, and the training results are compared in Table 3.

Table 3. Performance comparison of three training functions

Training function	iterations	mse
traingom	20000	0.0087401
triaingdm	5	0.000478142
traingdx	201	0.000989668

3.4 Training and Prediction of Network Mode

After determining the network structure (including the number of network layers, the number of neurons in each layer) and the network training function, the prediction model of the number of terrorist attacks based on BP neural network is established. The ediction model can be trained by using the sample data of the first six groups in Table 1. The maximum number of training times is set as epochs=800, target error goal=0.0000001, calling function train to train the BP network prediction model of the number of terrorist attacks. The result is shown in Figure 2.

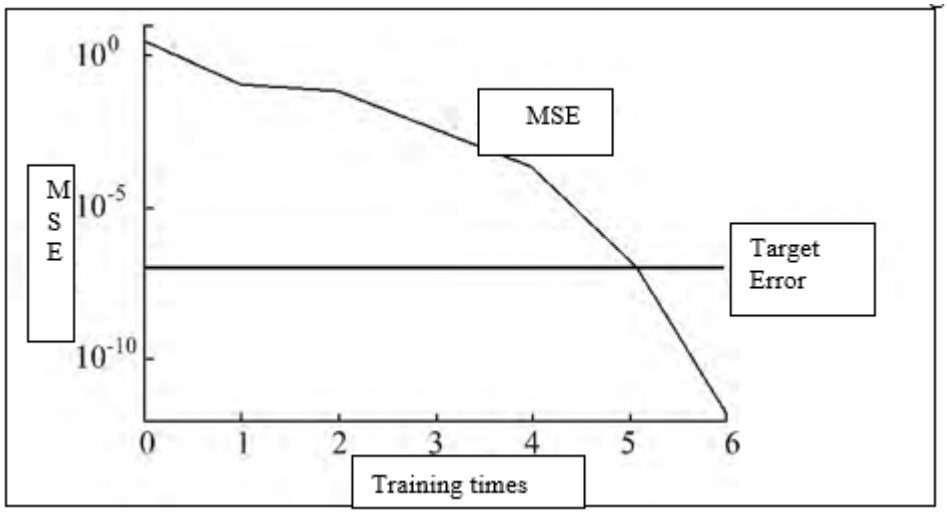

Fig. 2 Network training results

It can be seen that the prediction model of the number of regional terrorist attacks based on BP network has minimal mean square error after training.

The seventh and eighth groups of sample data are used as test samples to test the effect of the number of terrorist attacks in the established area. In order to analyze the test results intuitively, we use the workplace window in the Matlab toolbox to find out the simulation results after the test sample prediction model is used, and then compare the simulation results y with the target variable t _test = [1,2] of the test sample. The results are shown in table 4. As can be seen from the table, the predicted number of regional terrorist attacks based on BP neural network is [0.97751 ,2.02285]. So, the error is 2.249% and 1.1425%, respectively.

Table 4. Comparison of actual and predicted terrorist attacks in 2017

Sample serial number	Actual value	Predicted value	Absolute error	Relative error (%)
7	1	0.97751	0.02249	2.249%
8	2	2.02285	0.02285	1.1425%

It can be concluded from Table 4 that the model has high precision and can meet the demand of practical application. Therefore, we can predict the number of terrorist attacks in some provinces with frequent terrorist attacks by using this model, which is shown in Fig 3. Therefore, Xinjiang Province will produce 2 (2.0221) terrorist attacks in 2018, which is the largest number of terrorist attacks in the country. Therefore, we can draw the conclusion that Xinjiang Province is the most dangerous region in China.

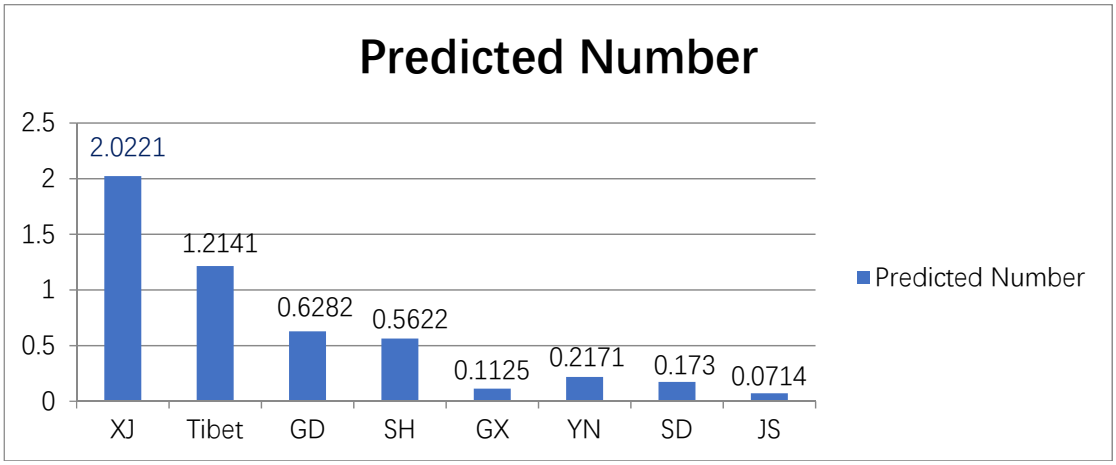

Fig. 3 Predicted number in 2018

4. Conclusion

The levenberg-marquardt algorithm with the fastest convergence speed and the smallest mean square error was selected by combining the empirical formula with the MATLAB program of computer software and comparing the training results of the prediction model with three improved BP algorithms. The BP neural network model is used to predict the number of regional terrorist attacks. The prediction results show that the model has a high accuracy and can meet the needs of practical application. It is also predicted that the region with the largest number of terrorist attacks in China in 2018 is XinJiang province.

References

[1]. Di Grande L,Neria Y ,Brackbill RM ,et al.Long-term posttraum stress symptoms among 3271 civilian survivors of the september 11,2001,terrorist attacks on the world trade center[J].Am J Epidemil,2010,173(3): 271-281.

[2]. LECUN Y, BENGIO Y, HINTON G E. Deep Learning[J]. Nature, 2015, 521; 436-444. DOI:10. 1038/nature 14539.

[3]. Ying Jianping, Lin Yixun, et al. Power Electronics Technology Foundation [M]. Mechanical Industry Press, 2003.

[4]. Li Zeyuan. Power Electronics Technology in the 21st Century [J]. Telecommunications Technology, 1999, 11.