Proceedings of 2009 Conference on Communication Faculty

Campus Network Multi-ISP Load Balancing Optimization Model Based on BP Neural Networks

XIE Haiyan[1,2], WANG Jian[1], ZHAO Depeng[2], SUN Hui[1]
1. Department of Mathematics, Dalian Maritime University, Dalian, China
2. School of Navigation, Dalian Maritime University, Dalian, China
e-mail: winteriscoming@sina.com, sunhui1376000@126.com

Abstract: In order to solve the problem that user have to make the manual strategy choice when they use the campus network of multi internet service provider (ISP), this paper puts forward a campus networks multi-ISP load balancing optimization model based on BP neural networks. This model can replace the manual strategy choice, automatically evaluate and forecast the performance of the networks according to the current data so as to adaptively choose the current optimal ISP. In this paper, the establishment of the network topology structure was discussed and a detailed description about the data collection and processing were given. The experimental results verified the effectiveness of the model.

Keywords: campus network; multi-ISP; load balancing; BP neural networks

1 Introduction

At present, the campus network mainly used by the Chinese colleges is China Education and Research Network (CERNET), which is the nation's largest public-welfare computer network. CERNET will not only provide a full range of internet services, but also support many large-scale educational information engineering. including Internet Distance Enrollment of Higher learning, modern distance education, digital libraries, education and scientific research projects, etc.

CERNET users belong to the internal network users (or so-called local area network users), quickly response of various educational institution and scientific research institution node could be gotten, and also with the information of education and scientific research. However, due to Chinese specific national conditions, there are many interconnection problems among the operators, The efficiency that one operator's customers visit another operator's is often low. Interconnection between different ISP, that is, cross-ISP link visiting bottleneck problem is the main reason for this phenomenon. For the servers which IP is not the goal for access to CERNET (China Telecom network, China Netcom network), whether the response or downloading speed both can't meet the requirement of the user.

So some of the campus network work with China Telecom, China Netcom, or other network operators, recanalize one or more network export to meet the needs of users. The selection for the export network, campus network center developed a login system, this system is the strategy taken user selection , education network, Telecom network or the Netcom network as the surfing export. Take our campus network of Dalian Maritime University as an example, the schools totally connected to three networks-Education Network (CERNET, Gigabit), Netcom (CNC, 10 trillion) and telecommunications (CHINANET, 10 trillion). For each user who want to surf the external network have to chose the network export when they login the external network.Now we will find a solution which make the system automatically selects a network export and this network export should be responded fastest, download the information fastest instead of the artificial selection .

Many scholars study the campus network ISP problem. In [1], Fu Zhiwei etc. used the network address translation (NAT) technology and the policy-based routing technology on the basis of source address, with an overview of different q/ p, to realize the network service for the campus users. In [2], Zhang Yunchuan etc. devised an alternate architecture based on NAT and standard routing mechanism based on a new topology structure and NAT policy. Request packets originated from campus network can be distinguished from response ones by NAT server at such approach. Reference [3] discussed the designing idea and method of campus multi-export serve system. Reference [4] proposed a king of technology based on policy-based routing selection and NAT, which can solve the network export bottleneck problem effectively. But these references failed to mention how to select ISP automatically. This paper evaluate and forecast the performance of the networks according to the BP neural networks to adaptively choose the current optimal ISP instead of the artificial strategic selection.

This paper has the following organization. Section 2 briefly reviews the basic algorithm of BP network, Section 3 introduces the campus network multi-ISP load balancing optimization model based on BP neural networks, and gives a detailed description of the establishment of the network structure as well as the collection and processing of data. In section 4, we give the experimental results. And the last section briefly concludes this paper.

978-1-935068-06-8 © 2009 SciRes.

2 Introduction of the BP Network Algorithm

2.1 BP Learning Process[5]

2.2.1 Working signal Forward Propagation

The input signals which go through the hidden units spread from the input layer to the output layer. And also produce the output signal at the output terminal.this is defined as the working signal forward propagation. The network weights remain unchanged when the signal forward transfer. The state of neurons in each layer only has influence on the state of neurons in the next layer. If network can not obtain the desired output from the output layer, then transfer to the error signal back propagation.

2.2.2 Error signal Back Propagation

The difference value between actual output and desired output in the networks, namely the error signal .And the error signal started with the output terminal backward spreads layer by layer. This is defined as the error signal back propagation. The conjunction weights of the neural network are continuously modified by error feedback to reduce the errors between the actual output and desired output in the process of the error signal back propagation.

2.2 BP Learning Algorithm Steps

Step 1: Set up variables and parameters;

Step 2: Initialization, assignment each of the $W_{MI}(0)$, $W_{IJ}(0)$, $W_{JP}(0)$ a smaller random nonzero value ;

Step 3: Input the random samples X_k, $n=0$;

Step 4: For the input samples X_k, calculate input signal u and output signal v of the neurons of the each layer in the BP network;

Step 5 : Calculate the error $E(n)$ from the desired output d_k and the actual output $Y_k(n)$ that is obtained from the preceding step .And judge whether the error meet the requirement , if that meet the requirement ,then transfer to step 8 ,otherwise ,turn back to step 6;

Step 6: Judge that whether $n+1$ is bigger than the highest iteration times ,if that is true ,then transfer to step 8,otherwise ,just back calculate the local gradient of the neurons of the each layer ;

Step 7: Calculate the correction value of the weights, and correct the weights, $n=n+1$, then transfer to step 4;

Step 8: Judge that whether all the training samples have been learned, end up all if that is true, otherwise, turn back to step 3

3 The Campus Network Multi-ISP Load Balancing Optimization Model Based on BP Neural Networks

3.1 The Structure Design of Neural Network Model

According [6], five indexes which are used to measure the network performance is respectively: availability, response time, network utilization, network throughput, network bandwidth capacity. So we can get the neurons of the input layer in the BP network model is respectively: ① Bandwidth utilization of the sending data of the Chinanet port; ②Bandwidth utilization of the receiving data of the Chinanet port; ③Response time of the Chinanet port; ④ Bandwidth utilization of the sending data of the CNC port; ⑤ Bandwidth utilization of the receiving data of the CNC port; ⑥Response time of the CNC port.

The neurons of the output layer is respectively:① The score of the Chinanet port ② The score of the CNC port. That score based on the time between the user send out the request data and receive the data completely, the smaller the score is, the better the performance of the network is.

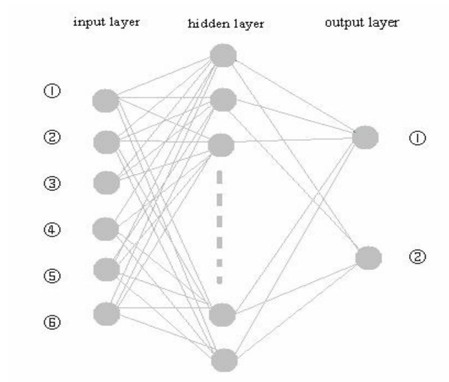

Figure 1. Neural network model

Robert Hedit-Nielson put forward an argument in 1989: As any continuous function in the finite, closed interval can be approximated with a hidden layer of BP network, so any function of n-dimensional space can be mapped to the m-dimension space by a three-layer BP network .A hidden layer is often designed with this kind of principle in mind: decreasing the scale of network as far as possible, to reduce the complexity of the network. It has proved by multiple tests that the numbers of hidden layers designed by using empirical formula $n_1 = 2n+1$, namely the neurons of hidden layer are 13, get the fastest convergence speed of BP neural network. The network topological structure of the model is shown as Fig.1.

3.2 Methods of Collecting Data

There are many factors affecting the properties of Neural Networks which can be summarized into three main factors: the topological structure of neural networks, the strength and weakness of learning algorithm and data processing. After determined the topological structure and algorithm of the networks, we need to collect the relevant data, and the process of collecting data will be introduced as follow.

Two servers are used in this experiment, namely, one host named Curl_data and other host named Snmp_data. Curl_data is responsible to obtain the response time to the target website by their respective network providers, the time from the start of user request to the user gets response resources and downloading speed. Snmp_data host is responsible to obtain the situation of network load of CNC and CHINANET export (that is the data quantity of provider who pass in and out the network).In this experiment target sites include http://www.baidu.com and http://www.google.com.

First of all, synchronize the time between the two servers, modify/etc/crontab files, then the system can operate the two scripts timing automatically. According to the monitor data of current network, we can know that both the overload time and idle time of the network appear in the period of the selected three_day sampling.

Curl_data host runs the script once per hour, each time the data will be collected three times, and these data include the bonding time, total time and downloading speed. Then execute the preliminary treatment for the data, which means to get their respective average value. This host is just to test the value of the microenvironment , namely the performance test of several commonly used target website .

Snmp_data host also runs the scripts once pre hour, nearly 10 seconds for each time. CPU time and network time can be neglect when the data is accurate to microsecond to calculate, and also the response connected to the switch can be neglected .After the end of the process to collect data, It still needs to do some simple data-processing, that is to figure out the flow amount which pass in and out of the network during that time .

There are totally 87 collecting data through above method, including 67 training data and 20 testing data.

3.3 Data Processing Method

Data sets often contain noises, errors and so on in the process of sampling, which often tend to cover up the main information, therefore data processing becomes very important.

Firstly it needs to examine the distribution of data when sufficient data is collected, and if necessary, transform the data to improve the network learning. According to [7], the data is most effective to the network learning when the variable shows normal distribution, so we can check in advance whether the variables present normal distribution. Commonly used method of calculation is to work out the skew coefficient and coefficient of prominent degree of data distribution. The calculation method is as follows:

$$S_c = \frac{[(x_i - \overline{x_i}) / SD]^3}{n} \tag{1}$$

$$K_c = \frac{[(x_i - \overline{x_i}) / SD]^4}{n} - 3 \tag{2}$$

where $\overline{x}$ is average value, SD is standard deviation, n is the number of data.

$\overline{x}$ and SD is calculated as follows:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x_i})^2}{n-1}} \tag{4}$$

If S_c is between -0.5 and 0.5, K_c is between -1 and 1, the data distribution could be taken as being close to a normal distribution. All the collected data processed by the method listed above meet the normal distribution.

However, we still need to normalize the data, which does not affect the distribution of primitive data, so as to get good network performance.

As for the six neurons input, firstly it needs to figure out the utilization ratio of bandwidth of each exit, that is

$$x = \Delta\text{ifOutOctets}/\left(\Delta\text{time} \times 10000000\right) \tag{5}$$

where ΔifOutOctets is the increment of flow amount, Δtime is the increment of time.Then use the standard normalization method to process data:

$$x^* = (x - agv(x)) / SD \tag{6}$$

SD is standard deviation as (4) shows, and $agv(x)$ is mean value of all the data, x^* is the value of normalization.

As for the two neurons output, use the method as follows to process the data. Firstly, it needs to sum up all the data to figure out the value $agv(x)$; then get the value of normalization according to (7),

$$x^* = x - agv(x) \tag{7}$$

finally, limit the data between 0.01 and 0.09.

It is noticeable that if the learning sample set used in training the neural network is normalized data, the new testing sample used to examine the neural network after studying should also use the sample set normalized, and of course the maximum value, the minimum value, the average value, the standard deviation and so on should use the related data in the learning sample set.

978-1-935068-06-8 © 2009 SciRes.

4 Experimental Result Analyses

Input the training sample set whose data have been processed to the BP Network which is shown as Fig.1 to

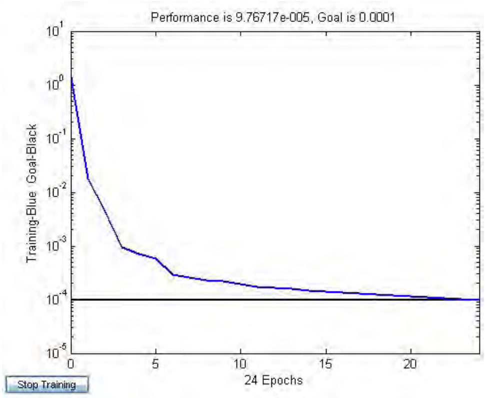

Figure 2. Error curve

train, and then make stimulation with the testing sample sets. Modify the weight by used of MATLAB neural network toolbox, and the selection of activation function of each layer requires a great number of training. It is perfect to make a choice with hyperbolic tangent function and a linear function as the activation function of hidden layer and output layer respectively. And as long as average iteration times are less than 30, it will be able to meet the requirement that error is 0.0001. Fig.2 is one error curve obtained through some training.

Table 1. Contrast table between practical value and predictive value

serial number	practical value	predictive value	Relative error(percent)
1	0.2490	0.2284	9.02
	0.2656	0.2615	1.52
2	0.7622	0.7545	1.02
	0.1772	0.1738	1.89
3	0.2490	0.2449	1.63
	0.2656	0.2779	4.64
4	0.3464	0.3238	6.51
	0.5229	0.5193	0.68
5	0.3320	0.3627	8.48
	0.4906	0.4793	2.31
6	0.1032	0.1118	8.37
	0.2533	0.2779	9.72
7	0.1638	0.1522	7.59
	0.1019	0.0982	3.68

According to a number of fast convergent situation, It can be seen that the selected input neurons, output neurons, the numbers of hidden layers and factors ,the sample set and so on for the test is reasonable.

The collected 67 groups of training data are used to calculate the optimal weight of the network, and then use 20 sets of data which are for test to predict simulation by the adoption of "sim" function of MATLAB toolbox, Table 1 lists a part of the practical value and predictive value and relative error .AS can be seen from the results of the table, the minimum value of relative error is 0.68%, the maximum value of relative error is 9.72%, and the average value of relative error is 4.79%.

The establishment of the BP network model is designed to find the best ISP by automatically assessment for network performance. According to the principle, smaller score (smaller output value) better network performance, may obtain the contrast table between actual best ISP and predicted optimal ISP as is shown in Table 2.

Table 2. Contrast table between actual best ISP and predicted optimal ISP

serial number	actual best ISP	predicted optimal ISP
1	ISP_1	ISP_1
2	ISP_2	ISP_2
3	ISP_1	ISP_1
4	ISP_1	ISP_1
5	ISP_1	ISP_1
6	ISP_1	ISP_1
7	ISP_2	ISP_2

It can be included from Table 2 that although error exists between the actual output and the desired output, which makes a great difference, will not affect the optimal choice of ISP. Among the 20 groups of data, two of them predicted failure, that is, the success rate of prediction was 90%. While the actual network evaluations for these two groups are: 0.3584, 0.3615 and 0.0160, 0.0168. By the view of the value of the practical network performance evaluation, both in the network environment above, the different choice of ISP make no difference. Therefore, it can be said that this model will completely implemented the adaptive selection of ISP.

5 Conclusion

This paper studies balancing optimization model of campus networks of multi-ISP load based on BP neural networks, and gives a detailed description of the establishment of the network structure as well as the data collection and processing, and we can draw the conclusion from the simulation of data in MATLAB that the model is effective.

Generally speaking, changes in tactics are made by a manual written routing table in the existing solution. The model proposed in this paper can conduct performance testing of each ISP, substitute into current network index, predict the performance that in accordance with the network status of each ISP to specific IP, so that the system could finally decide adaptively which ISP to be used, and it is a big progress.

The shortcomings of the model include great amount of data to be collected previously, long time to prepare and the changing top network.

Owning to the high redundancy of top network, even though a router has been optimized, it has little effect on total network, and because of the changing network, the setting of update cycle also needs to be taken into consideration.

6 Acknowledgment

Here, I thank my tutor and my schoolmates very much for their supports.

References

[1] Fu Zhiwei, Liu Yue, and Tan Ru, The Design and Implementation of the Access of Campus Network to the Multi-ISP [J], Journal of Liaoning Normal University (Natural Science Edition), 2003, 26(4), P382-384 (Ch).

[2] Zhang Yunchuan, Wang Yaoqing, and Qian Jianghong, Study and Implementation on Large Scale Campus Network Multi-ISPs Connection[J], Microcomputer Development, 2005, 15(4), P11-13 (Ch).

[3] Jiang Hua, Qiu Weidi, Analysis and Designing of Campus Multi-Export Server System [J], Journal of Guangxi University (Natural Science Edition), 2006, 31, P292-295 (Ch).

[4] Huang Min, Zhang Weidong, Network Design and Realization Based on Policy-Based Routing [J], Computer Applications, 2002, 22(5), P72-74 (Ch).

[5] GAO Jun. The artificial neural network principles and simulation examples [M]. Beijing: China Machine Press, 2007.

[6] Tang Kai, Netperf and Network Performance Test [J], IBM DeveloperWorks, 2004(Ch).

[7] HAGAN Martin T., DEMUTH Howard B., BEALE Mark H., Neural network design [M]. Beijing: China Machine Press, 2005.

978-1-935068-06-8 © 2009 SciRes.