Time Series Prediction of Wireless Network Traffic Flow Based on Wavelet Analysis and BP Neural Network

Yanchi Li1, Jin Huang2 and Haojie Chen2

1 Faculty of Natural & Mathematical Sciences, King’s College London, London, WC2B 4BG, UK
2 School of Computer, Beijing University of Posts and Telecommunications, Beijing, 100089, China

*Corresponding author email: K1922383@kcl.ac.uk

Abstract. Building an appropriate model is the key issue of network traffic flow prediction [1]. This paper presents a hybrid model based on the combination of wavelet denoising and BP neural network. First, wavelet transform is used to normalize and denoise the collected raw traffic flow. The index for estimating prediction accuracy is Mean Absolute Error (MAE). The original data is supported by a large telecommunications company in China. At present, research has made corresponding progress and the prediction accuracy of this combined model has reached more than 80%.

Keywords: wavelet analysis, BP neural network, network traffic prediction

1. Introduction
A cell is an area covered by a base station in a cellular mobile communication system. In this area, a mobile station can reliably communicate with the base station through a wireless channel. For the traffic-rich cells, the capacity expansion is a necessary condition to support user communication. Otherwise, the cell is likely to have network congestion during the peak period of network traffic. Existing research has recognized the crucial role played by network traffic prediction technology. Sufficiently accurate prediction results are helpful for solving a series of problems such as resource allocation and cell load balancing. In this paper, the telecommunications company provides 842 cells and 211 days of traffic data per cell. It is required to predict the data in the subsequent 27 days through traffic prediction technology. The objectives of this research are as follows:

2. Related technologies
2.1 Limitations of other data preprocessing methods
According to the literature [2], non-stationary sequences tend to become stationary after one or two differentials. The so-called difference is to subtract the value \( T_{i,j} \) of the previous time point by the value \( T_i \) of the latter time point, and the difference can effectively remove the abnormal point in the time series (i.e., the point of sudden rise and sudden drop) after the differential processing This method
is often used in conjunction with the ARMA model. However, the difference method cannot completely remove the outliers.

2.2 Wavelet analysis
In practical applications, the data collected in the actual network channel is often subject to noise interference, showing obvious non-stationary characteristics. Due to the advantages of wavelet analysis in dealing with non-stationary time series, wavelet denoising is used to pre-processing network traffic time series in this paper.[3-4].

2.3 BP Neural Network
BP neural network consists of three parts: input layer, hidden layer and output layer. Its general structure is presented in Figure 2-1.

Figure 2-1. Basic structure of BP Neural Network

3. The prediction model based on the wavelet denoising and (BP) neural network
The model consists of two parts. The first is the preprocessing of data by wavelet denoising, and the second is (BP) Neural network. The function diagram of the model is presented in Figure 3-1.

Figure 3-1. Functional structure of the algorithm
3.1 Detailed design of (BP) Neural Network

3.1.1 Determination of the number of neurons in the input and output layers. 1) Determination of the number of input layer neurons in single-index prediction

In this paper, all of the integers between 1 and 10 are selected for experiments because under such conditions, the network size is moderate. By testing the number of different input layer neurons, it is found that when the number of input layer neurons is 3, the best prediction result can be obtained, and the accuracy rate is the highest.

2) Determination of the number of input layer neurons in multi-index prediction

Unlike single-index prediction, in multi-indicator prediction process, the number of input layer neurons is based on the number of selected indicators instead of cyclic predicting method. In this paper, six out of 44 attributes with the strongest correlation size are selected, so the number of input layer neurons is 6.

3.1.2 Determination of the number of neurons in the hidden layer. Literature [5-6] has conclusively established that the choice of the hidden layer is based on the following three formulas:

\[ S = \log_2 N \# (3 - 1) \]

where \( S \) is the number of hidden layer neurons, \( N \) is the number of samples, and the calculation result of \( S \) is 8.

\[ S = 2N + 1 \# (3 - 2) \]

where \( N \) is the number of neurons in the input layer, and the calculation result of \( S \) is 7.

\[ S = \sqrt{(0.43mn + 0.12mn + 2.54m + 0.77n + 0.35)} + 0.51 \# (3 - 3) \]

where \( m \) and \( n \) are the number of neurons in the input layer and the output layer, respectively, and the calculation results in \( S \) is 11.

3.1.3 Selection of activation function. The neural network adopts the sigmoid activation function, which is the most widely used activation function in (BP) neural network. It is a strictly incremental "S" type function that balances linear and nonlinear behavior well and enables arbitrary nonlinear mapping from input to output. The expression for this function is as follows:

\[ \phi(v) = \frac{1}{1 + e^{-av}} \# (3 - 4) \]

where \( a \) is the slope of the sigmoid function.

4. Results

4.1 Single-index prediction (model combining wavelet denoising and neural network)

In this paper, the original data set is used for forecasting. The data is received from the 842 cells monitored by a broad-scale base station. The time span of each cell is 211. The traffic data of the first 181 days is the neural network training data and the traffic data for the last 27 days is the neural network test data. Using the traffic prediction model to predict network traffic data for 27 days, the sequence of network traffic used for training in a certain cell is shown in equation (4-1), which is obviously \( n=181 \).

\[ \{X|x_1, x_2, x_3, ..., x_n\} \# (4 - 1) \]

In order to test the utility of the model prediction, the average absolute error (MAE) is used to calculate the accuracy. The formula is as follows:

\[ MAE = \frac{1}{m} \sum_{k=1}^{n} \left| \frac{y_k - \hat{y}_k}{y_k} \right| \# (4 - 2) \]

where \( y_k \) represents the test set, i.e. the actual data, \( \hat{y}_k \) represents the value predicted by the algorithm, \( m \) is the step size, for example, here is the day after the first three days of prediction, then \( m \) is equal to
3. MAE corresponds to the accuracy rate. The smaller the MAE, the higher the accuracy.

The platform used to denoise is the Wavemenu toolbox in MATLAB. The Wavelet Packet 1-D toolkit is used and DB3 wavelet basis function is used for three-layer decomposition of the original sequence. The signal sequence after denoising tends to be stable, and the tip and the mutation are reduced. The analysis of the autocorrelation function and the partial correlation function of the processed sequence shows that the processed traffic sequence has basically conformed to the stationary sequence[7-8].

The comparison between the predicted value and the actual value of the 0th cell after the wavelet denoising is shown in figure 4-1. MAE is 0.0994.

![Comparison of training set results and comparison of test set results of no.0 cell after denoising](image)

**Figure 4-1** Comparison of training set results and comparison of test set results of no.0 cell after denoising

A summary analysis of 842 communities after denoising and predicted by single-index BPNN is shown in Table 4-1, where the first line represents the MAE value, and the second line represents the number of corresponding cells.

| MAE Value (%) | Corresponding Number (%) |
|---------------|--------------------------|
| 0 - 20%       | 17.5%                    |
| 20% - 30%     | 35.9%                    |
| 30% - 40%     | 48%                      |
| 40% - 100%    | 79%                      |
| >100%         | 21%                      |

### Table 4-1. Summary of MAE value distribution of all cells using single-index prediction together with wavelet denoising

4.2 Multiple indicator prediction (model combining wavelet denoising and neural network)

According to the experimental results of 4.2, we can see that the wavelet noise reduction method has reduced the overall prediction accuracy, which indicates that the existing model must be improved. Since the fluctuation range of the data is still large after denoising, in order to achieve higher accuracy when simulation is performed and eliminate the Gibbs effect during wavelet denoising, the traffic flow is then normalized, and the data is denormalized after BPNN is used [9]. The normalization function used is shown below.

\[
y = \frac{(y_{\text{max}} - y_{\text{min}}) \cdot (x - x_{\text{min}})}{(x_{\text{max}} - x_{\text{min}}) + y_{\text{min}}} \cdot (4 - 3)
\]

where \(y_{\text{max}}\) defaults to 1 and \(y_{\text{min}}\) defaults to -1. \(x\) is the original data item to be processed. \(x_{\text{max}}\) and \(x_{\text{min}}\) are the maximum and minimum values in the original data respectively.

The prediction result of the 0th cell test set is shown below, and the MAE is 0.116.
Figure 4-2 The prediction result of the 0th cell test set

A summary analysis of 842 communities after denoising and predicted by multiple-index BPNN is shown in Table 4-2, where the first line represents the MAE value, and the second line represents the number of corresponding cells.

Table 4-2. Summary of MAE value distribution of all cells using multiple indicator prediction

| ≤ 20% | ≤ 30% | ≤ 40% | ≤ 100% | >100% |
|-------|-------|-------|--------|-------|
| 33.7% | 62.9% | 65.2% | 94.7%  | 5.3%  |

4.4 Comparison between neural network model and other models

In this paper, three models are compared to evaluate the prediction result. Comparison of the three models is presented in Table 4-3.

Table 4-3. Comparison of different models

| BP Neural Network (Single indicator) | BP Neural Network (Multiple indicators) | ARMA model | ELM |
|-------------------------------------|----------------------------------------|------------|-----|
| 0.3489                              | 0.2919                                 | 0.4227     | 0.3063 |

The table shows that the performance of BPNN is better than the other two models.

5. Conclusion

Network traffic flow prediction had long been a largely under-investigated domain in the past two decades and there is no definitive model which can get the best prediction results. The combinational model proposed in this paper is not optimal to some extent. However, despite its disadvantages, BPNN has higher execution efficiency compared with those widespread used models such as ARMA. This paper provides an overview of different prediction models and simultaneously outlines a new rational prediction model with high execution efficiency and good accuracy, helping technicians quickly determine which cells have large traffic fluctuations in the next month thus need to be expanded.

Reference

[1] Sun Guang. Network Traffic Prediction Based on the Wavelet Analysis and Hopfield Neural Network. International Journal of Future Computer and Communication, Vol. 2, No. 2, April 2013

[2] Zeng Wei. Differential evolution algorithm for BP neural network optimization of energy
demand prediction [D], [Master's thesis]. Huazhong University of Science and Technology: 2015-04-24

[3] Sven A M, Harsha, Sirisena. The Influence of long-range dependence on traffic prediction [A]. IEEE International Conference on Communications [C]. Helsinki: IEEE Press, 2001. 1000-1005

[4] Micro-Electronics 141--Analysis of ECG Signal Denoising Algorithm Based on Wavelet Transform[C]. College Students' Papers Comparative Matching Library. 3140434030: 2018-06-15

[5] Chang Xue: Campus Network Bandwidth Traffic Prediction Based on BP Neural Network Theory [D], [Master's thesis]. Harbin Engineering University: 2008-04-01

[6] ARMA-based traffic prediction and overload detection of network. Zou Bo Xian and LIU Qiang (Information Network Laboratory, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080) 2002 12

[7] Guang-Bin Huang, Qin-Yu Zhu, Chee-Kheong Siew. Extreme learning machine: Theory and applications. Volume 70, Issues 1–3, December 2006, Pages 489-501

[8] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, Deep learning. Nature Publishing Group 2015/5

[9] REN Xiaoxia; SONG Haiming; Shanxi. Short-term Freeway Traffic Flow Prediction Based on Wavelet Neural Network. Technology and Business University; Economic and Technical Research Institute, State Grid Xinjiang Electric Power Corporation; 2017 6