Internet time-delay prediction based on Wavelet transformation and ARIMA

Liang Zhao[a], Ruobing Liang[b], Jili Zhang[b] Wenyan Wu[c]

[a]Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

zliang@dlut.edu.cn

[b]Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116024, China

[c]Faculty of Art and Creative Technologies, Staffordshire University Stafford, Staffordshire, ST18 0AD, UK

Abstract. Based on the analysis of the fundamental characteristics of the end-to-end network time delay, this paper presents a new network delay prediction method based on wavelet- autoregressive integrated moving average model (W-ARIMA). The wavelet filter is applied to eliminate the noise signal from the time delay data and then the ARIMA model is used to predict the time delay sequence. An experimental example is given to show that the performance of the proposed delay predicting method, the results show that the W-ARIMA proposed in this paper has much higher prediction accuracy compared with the method based on the direct autoregressive integrated moving average (ARIMA) model.

Keywords: Internet time-delay, ARIMA, wavelet transformation, prediction

1 INTRODUCTION

In recent years, with the maturing of Internet technology, the Internet network control system has widely applied to the industrial field control, fault detection and remote monitoring [1-3]. However, the network delay greatly increases the complexity of the network control system analysis and design, which is exiting inevitably and randomly [4]. Hence, accurate modeling and prediction of Internet network delay is the basis of the analysis and design of networked control system.

In 1978, Wong J W [5] used the Poisson arrival assumption to take on the analysis of traditional network delay queue, which considered that the data arrival interval was unrelated. However, the simulation showed that it had a great difference between the predicted value and measured network delay by using queuing theory. Qiong L, Ming Y, and Jiao L, proposed the use of linear time series forecasting model based on AR (Auto-Regressive), ARMA (Auto-Regressive Moving Average), ARIMA (Auto-Regressive Integrated Moving Average) forecast model respectively in 2001, 2003, 2006 [6-8], forecasting analysis for network delay. It is obtained that the low order AR or ARIMA model can accurately model the network time delay, and provide accurate prediction accuracy through simulation results.

In addition, there are many scholars use intelligent methods such as neural networks for prediction of network delay [9-11]. The neural network has the ability of nonlinear identification, which can therefore be used to predict network time delay, but it is easy to fall into local extremum due to the reasons of “over training” or “lack of training” in the traditional neural network. In 2007, Cun Li [12] proposed a prediction method of two level network delay based on the SVM (support vector machine), which is suitable for the prediction of network delay with strong nonlinearity. However, this method still has the problem of long computing time, more seriously, with the increase of the specimen and its accuracy will be decreased. In [13], Xiao T offers a prediction delay modeling by using the grey theory, which is very similar with autoregressive model, but the prediction precision of grey method is poor if there exits random jitter in network delay time series.

As we know, the autoregressive moving average model (ARMA) has a high prediction accuracy and advantages of simple modeling, fast computational speed in the short term, compared with other prediction methods. However, it will bring lots of computation and increase the complexity if we use the ARMA model directly, as for the time-varying and nonlinear characteristic performance of network delay. It is difficult to guarantee the real-time capability of network control system. To avoid this problem, we adopt differential treatment and wavelet transform to eliminate the non-stationary of the network delay sequence, and extract the change trend of network delay sequence. It could reduce the order of the model and improve the accuracy of prediction of network delay by this method.

This paper first analyzes the composition of network delay, and then collects the real end-to-end time delay data by Ping method from a client to the Sina site. Secondly, proposes the network delay prediction method based on Wavelet-Autoregressive Integrated Moving Average model (W-ARIMA) to judge the network quality.

2 TIME DELAY DATA COLLECTING

2.1 The composition of network delay

Generally speaking, end-to-end network delay is mainly divided into four parts, namely: the processing delay, transmission delay, propagation delay and queuing delay.

(1) Transmission delay: refers to the time taken from the entrance of the first bit of data packet into the communication link to entrance of the last bit of the data packet into the communication link.

(2) Propagation delay: refers to the time taken from sending the first bit of data packet by client to receipt of the bit by server. The transmission delay is mainly decided by the propagation speed and the physical distance between the client and server, and it is considered to be the same in most cases, because the propagation speed is very fast (2/3 the speed of light).

(3) Queuing delay: refers to the packet in the router's buffer, waiting time before transmission or processing, which is mainly affected by router performance and network load. In general, the value of queuing delay changes randomly.

(4) Processing delay: refers to the time between packet arrives at the input end of a certain node and arrives the output end of the node, which is mainly affected by the computing power of that processing node. The processing delay is random on the whole, as for the speed of each probing packet processed in the router is not exactly the same, and it can be divided into a definite part and a random part.

2.2 Network delay collection

In this paper, we select the typical method ‘Ping’ to obtain the end-to-end delay data, the client computer is in an office of Dalian University of Technology, and the Server selects Sina server (IP:121.194.0.239), with 10 intermediate routers between the client and server.

We do the Ping experiment from midnight, 12, Aug. 2014 to midnight, 13, Aug. 2014, for 24 hours in order to observe the network delay characteristics in a long term. The probe packet size is set 32B, the Ping period is 1 minutes, and we can collect 1,440 sets of data. Figure 1 shows the changing curve of measured Ping network delay between client and the server. As can be seen, network delay is distributed randomly within a certain range, and presents a periodicity over time. The value of delay is higher within daytime, when network load is larger, and delay fluctuation is greater; on the other hand, it is lower during late night and early morning, and delay fluctuation is relatively smooth.

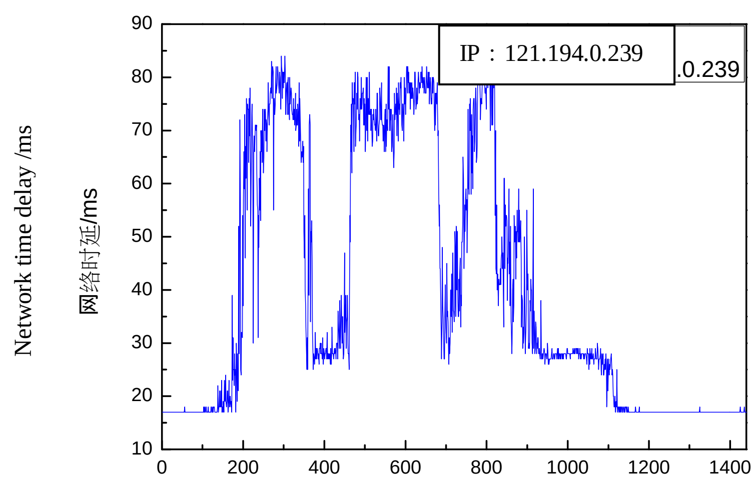


Figure 1 Network time delay sampling

3 WAVELET DECOMPOSITION FOR TIME DELAY SEQUENCE

3.1 The principle of wavelet transform

In 1987, Mallat presented an analysis of time domain, frequency domain and has the function of wavelet method [14], it can clearly reveal the various changes in the cycle of hidden in the signal, fully reflect the trend of changes in different time scales in the system. Wavelet makes a complex non-stationary signal more obvious after wavelet transform, also can make a qualitative estimate of the future development trend of the system. The Mallat architecture is put forward on the basis of a fast wavelet transform algorithm in the multi resolution analysis, namely the famous Mallat Pyramid shaped algorithm. This algorithm can do signal time domain and frequency domain analysis for the original signal by constructing a two channel filter coefficients, in the case of unknown of the scale function $\phi(t)$ and wavelet function $\psi(t)$.

(1) Decomposition, Mallat wavelet decomposition algorithm is described as Equation (1)

$$\begin{cases} c_{j+1}(k) = \sum_{n=-\infty}^{\infty} c_j(n) h_0(n-2k) = c_j(k) * \overline{h_0}(2k) \\ d_{j+1}(k) = \sum_{n=-\infty}^{\infty} c_j(n) h_1(n-2k) = c_j(k) * \overline{h_1}(2k) \end{cases} \tag{1}$$

Where, $c_j(k)$ and $d_j(k)$ are the discrete approximation coefficients of signal multi-resolution analysis, $h_0(k)$ and $h_1(k)$ are two filters for difference equation.

(2) Reconstruction, the Mallat wavelet reconstruction algorithm is shown as Equation (2).

$$c_j(k) = \sum_{n=-\infty}^{\infty} c_{j+1}(k) h_0(k-2n) + \sum_{n=-\infty}^{\infty} d_{j+1}(k) h_1(k-2n) \tag{2}$$

Where, $c_{j+1}(k)$ and $d_{j+1}(k)$ can be obtained from Equation (1).

3.2 Time delay sequence wavelet decomposition

The signal wavelet decomposition is actually the process of implementation of the discrete wavelet transform using the filter, as shown in Figure 2. The S is representation of the original time series, through two complementary filters to generate A, D two signals, wherein A represents the signal approximation (Approximation), low frequency part of the signal, D reflects the signal details value (Detail), which is of high frequency part of the signal. Generally speaking, the low frequency part of the signal is the most important, and the high frequency part of an "additive" role.

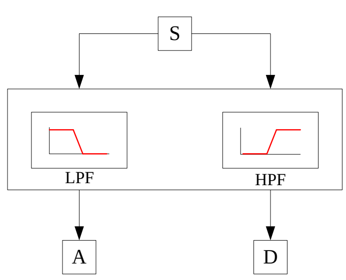


Figure 2 Filtering process of two-channel Wavelet

Thus, the discrete wavelet transform can be expressed as the composition of a tree consists of by a low-pass filter and a high pass filter. This tree is called the wavelet decomposition tree, the original signal is decomposed by a number of such filters, and the decomposition layers are depended on the need of data to be analyzed and the requirement of user. In order to select the appropriate decomposition layers, first of all, starting from the overall delay sequence, here, we use the data collected to design wavelet decomposition.

Figure 3 shows the network delay wavelet decomposition for network time delay of Sina. From Figure 3, it can be seen that, when choosing the 2 layer wavelet decomposition, the delay time series are not smooth enough, and jitter phenomenon still more obvious; when choosing the 4 layers wavelet decomposition, but filter out much of the delay characteristics, and can't correctly reflect the distribution characteristics of the time delay; when choosing the 3 layers, the decomposed signal is kept the original time series period characteristics, but also reflects the most of the changing details. Therefore, we finally set the decomposition structure is 3 layers.

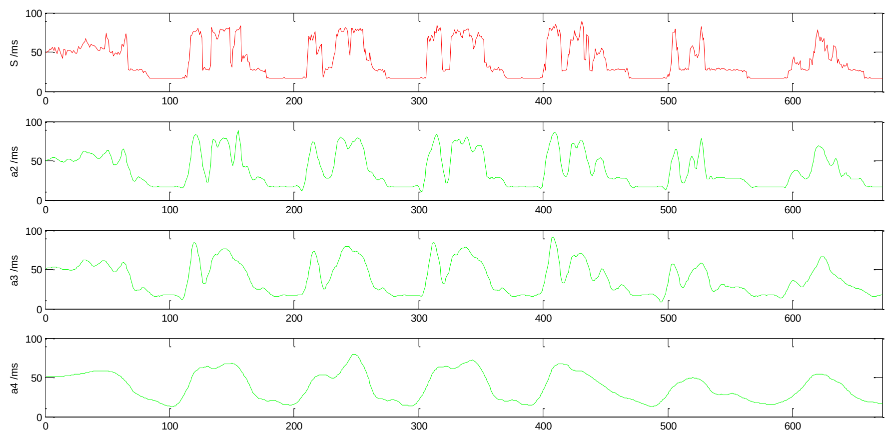


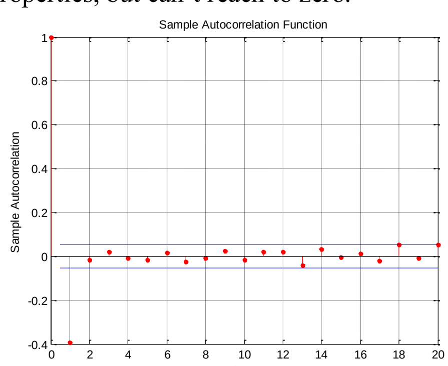
Figure 3 Wavelet decomposition for network delay sequence

4 PREDICTION OF NETWORK TIME DELAY BASED ON W-ARIMA MODEL

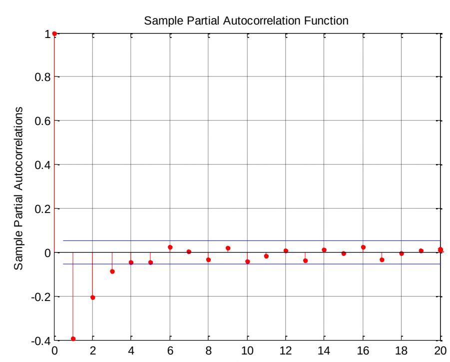
4.1 Network delay modeling

Process modeling of network delay is actually the time sequence of the delay composition process of ARMA model identification. Before the modeling of network delay sequence, we should first determine the stationary characteristics of the time delay sequence series, due to the AR model, MA model or ARMA model are only for stationary time series modeling. If the delay time series are non-stationary series, then needs to do differential treatment make it into a stationary sequence, and the ARMA model becomes ARIMA (Autoregressive Integrated Moving Average Model) model after differential treatment.

The model identification process often judged by the "tailing" and "truncation" characteristics of self-correlation coefficient and the partial correlation coefficient of the stationary time series. The so-called "truncation ", refers to the existence of 'q', the covariance function satisfies $Cov(X_t, X_{t+k}) = 0$ when k>q; on the other hand, the so-called "tailing", refers to the function presents negative exponential decay properties, but can't reach to zero.



(a)



(b)

Figure 4 Autocorrelation and partial autocorrelation coefficients of network delay sequence of Sina after first order difference process

Figure 4 (a) shows the corresponding autocorrelation and partial autocorrelation coefficient of the first 20 items for time delay of Sina, it can be clearly seen from the figure, the autocorrelation coefficient presents the slow linear decrease trend, so the delay sequence is non-stationary time series. While, Figure 4 (b) describes the corresponding autocorrelation and partial autocorrelation coefficient of the first 20 items after the first order difference process. It's not hard to see that, the autocorrelation coefficient of delay sequence satisfies the 'truncation' characteristics in the 1 step, which means that we can use the MA (1) model to simulate the delay sequence. On the other hand, the partial correlation coefficient delay sequence satisfies the 'tailing' characteristics, which implies that we can select the AR (2) model to simulate the delay sequence. Finally, we also could select the ARMA (3,0) model according to the principle of Box-Jenkins modeling, which has less amount of calculation.

Followed by the Akaike Information Criterion (AIC) compares the 3 models [15], as shown in Table1. The final parameters are selected as: p=3, q=0, d=1, namely ARIMA (3,1,0) model, the parameters of the least square estimation are $\hat{\varphi}_1 = -0.4934, \hat{\varphi}_2 = -0.2464, \hat{\varphi}_3 = 0.0876$.

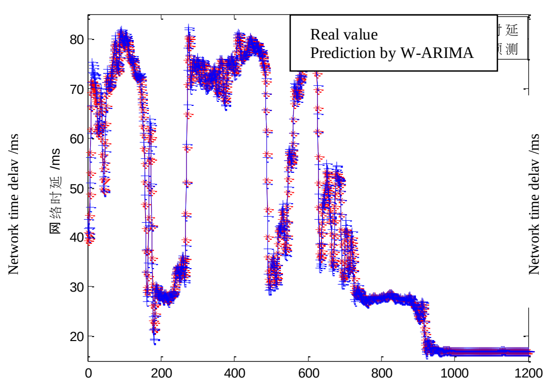
Tab. 1 AIC parameters of model

Model	Residual sum of squares	AIC
AR (3)	3857.22	6.102
AR (2)	38603.43	6.1081
MA (1)	38773.71	6.2204
ARMA (2,1)	38768.12	6.1108

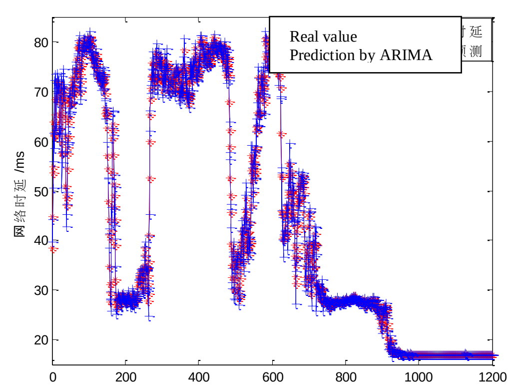
4.2 Prediction of network delay

After modeling, we use Matlab program to realize the prediction of end-to-end time delay. The input parameter of the real-time wavelet filter is 128, and starts to wavelet filter process in a rolling manner from the 129th input delay data. Firstly, the time delay data needs to do a first-order differential procedure, and then sent into the ARMA model forecast.

Figure 5 is the output results prediction. Figure 5 (a) and (b) are the overall views of prediction. Figure 5 (a) shows the ARMA forecasting result after the wavelet filtering, and Figure 5 (b) shows the directly ARMA forecasting result without wavelet filtering procedure. In order to observe more intuitive, we intercept the 100 points to amplify, as shown in figure (c) and (d). By comparing the local prediction effect, it can be found that the whole trend of wavelet denoising is more smooth compared with directly using of the ARIMA model. Besides that, the W-ARIMA has higher prediction precision.

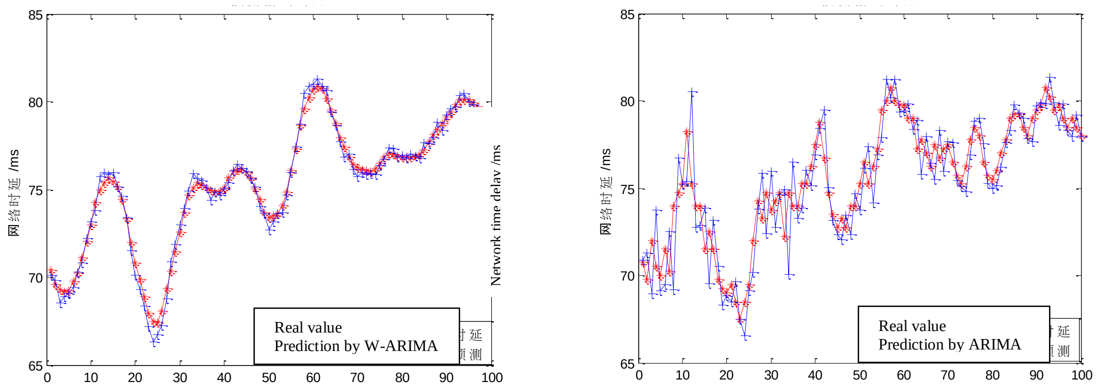


(a)



(b)

Network time delav /ms



(c) (d)

Figure 5 Prediction results of time delay

5 CONCLUSION

This paper mainly studies the end-to-end network delay composition, delay measurement and delay prediction. The network delay sequence data are collected by using the Ping method, and analyzes the distribution characteristics of it. A new W-ARIMA method is proposed to predict the network delay. Wavelet filtering transform could eliminate the non-stationary of the network delay sequence, and grasp the delay variation trend. The predicted results show that wavelet prediction accuracy is better than direct prediction. This method offers a new idea for the study of the Internet network delay prediction.

Acknowledgments

The authors acknowledge the financial support of the Seventh Framework Programme of the European Union (EU FP7) Marie Curie Actions – SmartWater (PIRSES-GA-2012-318985).

References

Zhang J, Guo P, Lyu M, et al. (2014). Fault Detection for Wireless Network Control Systems with Stochastic Uncertainties and Time Delays. *Abstract and Applied Analysis*. Article ID 250314, 9 pages. Doi: 10.1155/2014/250314.

Xu F, Cheng Y, Ren H, et al. (2014). Research on Adaptive Neural Network Control System Based on Nonlinear U-Model with Time-Varying Delay. *Mathematical Problems in Engineering*. Article ID 420713, 7 pages. Doi:10.1155/2014/420713.

Liu L, Liu Y, Chen C L P. (2013). Adaptive Neural Network Control for a DC Motor System with Dead-Zone. *Nonlinear Dynamics*, 72 (1) 141-147. Doi: 10.1007/s11071-012-0698-2.

Tian Z, Gao X, Li K. (2014). A hybrid time-delay prediction method for networked control system. *International Journal of Automation and Computing.* 11 (1): 19-24. Doi: 10.1007/s11633-014-0761-1

Wong, J W. (1978). Queueing Network Modeling of Computer Communication Networks. *ACM Computing Surveys.* 10 (3): 343-351. Doi: 10.1145/356733.356740

Jiao L, Zhang D, Houjie B. (2006). Differential AR algorithm for packet delay prediction. *Progress in Natural Science*. 16 (4): 437-440. Doi: 10.1080/10020070612330016

Ming Y, Li X R, Huimin C, et al. (2004). Predicting Internet end-to-end delay: an overview. System Theory. *Proceedings of the Thirty-Sixth Southeastern Symposium on.* Doi: 10.1109/SSST.2004.1295650

Li Q, Mills D L. (2001). Jitter-based delay-boundary prediction of wide-area networks. *Networking, IEEE/ACM Transactions on.* 9 (5): 578-590. Doi: 10.1109/90.958327

Rahmani B, Markazi A, Mozayani N. (2008). Real time prediction of time delays in a networked control system. *3rd International Symposium on Communications, Control and Signal Processing. ISCCSP*. Doi: 10.1109/ISCCSP.2008.4537416

Sadeghzadeh N, Afshar A, Menhaj M B. (2008). An MLP neural network for time delay prediction in networked control systems. *Control and Decision Conference. CCDC.* Doi: 10.1109/CCDC.2008.4598345

Li H, Wang H, Gui C. (2006). Internet Time-delay Prediction Based on Autoregressive and Neural Network Model. *International Conference on Communications, Circuits and Systems Proceedings.* Doi: 10.1109/ICCCAS.2006.285014

Li C, Xiao J, Zhang Y. (2007). Time delay estimation and adaptive control of networked control systems. *Journal of Southwest Jiaotong Universit*y. (02): 175-180. Doi: 10.1109/ISCIT.2006.339832

Xiaomin T, Yang S, Minrui F. (2011). Internet Time-Delay prediction based on unbiased grey model. *9th World Congress on Intelligent Control and Automation (WCICA).* Doi: 10.1109/WCICA.2011.5970638

Mallat, S G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 11 (7): 674-693. Doi: 10.1109/34.192463

Kavousi-Fard A, Kavousi-Fard F. (2014). A new hybrid correction method for short-term load forecasting based on ARIMA, SVR and CSA. *Journal of Experimental & Theoretical*, 25 (4): 559-574. DOI:10.1080/0952813X.2013.782351