http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (12): 2322-2328, 2011
ISSN 1812-5638 / DOI: 10.3923/itj.2011.2322.2328

Research on Fuzzy Self-adaptive Variable-weight Combination Prediction Model for IP Network Traffic

Minsheng Tan, Hui Xu, Lin Zeng and Shiying Xia
School of Computer Science and Technology, University of South China,
Hunan Hengyang, 421001, China

Abstract: In combination prediction of IP network traffic, the single model's mathematical characteristic, prediction accuracy and weight coefficient have significant impact on combination prediction results. As the grey model can depict linearity characteristics of network traffic and the BP neural network model can depict the non-stationary and non-linear characteristics, a Fuzzy Self-Adaptive Variable-Weight Combination Prediction Model (FSVCPM) was composed of them. To improve the prediction accuracy of single model as far as possible, a improved residual grey prediction model was established via indexation processing of residual sequence. By training experiments, neuron number of input layer and hidden layer was identified and corresponding BP neural network was given. By introducing fuzzy decision mechanism and self-adaptive mechanism to calculate fuzzy weight and basic weight, FSVCPM was built and a determination method of variable-weight coefficient was addressed which can make single models to fit effectively. Experimental results validated the correctness and accuracy of the FSVCPM and proved the prediction precision was higher than that of the single model and the Constant-Weight Combination Prediction Model (CCPM).

Key words: IP network traffic, prediction model, constant-weight combination prediction model, variable-weight combination prediction model (VCPM), self-adaptive mechanism

INTRODUCTION

With the rapid development of Internet and its application, network tend to bear more and more application, the features of network behavior become complicated increasingly. Modeling and predicting for IP network traffic is the basic of bandwidth allocation, network traffic engineering, behavioral decision, performance analysis, router and error control, also an important means of network monitoring, abnormity, failure and attack discovery. Via modeling and predicting for IP network traffic, we can find out the condition and trends of network traffic, optimize network more effectively, design routing and load balance for better, work at network congestion control, reduce information loss and delay, make full use of resources and improve service quality (Wu *et al*., 2001; Wang *et al*., 2004; Feldmann *et al*., 2000).

Network traffic prediction model contains regression model, time sequence model, grey model, neural network model, wavelet model and so on. However, as the scale of network becoming larger and larger, these single prediction models have deficiencies generally in prediction accuracy and is hard to predict dynamically. The studies showed that a larger error prediction model and a less error prediction model are combined which can increase prediction performance. Combine plenty of single prediction models correctly, the combination prediction results will be less sensitive to the single prediction model with large error. So combination prediction model can improve prediction accuracy and reliability usually.

An important question in combination prediction is how to work out average weight coefficient and improve prediction accuracy more effectively. For the long term prediction, the CCPM is difficult to obtain satisfactory prediction accuracy, thus VCPM is given.

RELATED RESEARCH

More than ten years ago, it is found the traffic of modern data networks can't be described by the traditional Poisson model (Leland *et al*., 1994, 1993). SEE-IT measure prediction and analysis tools (Yang and Wang, 2001) of F5 NetWork company and a network performance analysis and prediction support system (Sang and Li, 2000) proposed by Wuhan University can complete special network node's performance analysis

Corresponding Author: Minsheng Tan, School of Computer Science and Technology, University of South China, Hunan Hengyang, 421001, China

and traffic measuring and predicting. Sang and Li (2000) analyzed the issues related to multi-scale network traffic prediction based on ARMA and MMP model (Zou and Liu, 2002). Wang *et al.* (2005) introduced network traffic prediction method based on ARMA model, using auto-regressive sliding average model to analyze network traffic whose cycle is 5; Groschwitz and Polyzos (1994) had analyzed minor time granularity for network traffic of auto-regressive characteristics, on the basis of this, present study proposed auto-regressive prediction model and proved that prediction error less than 15% probability is 90% in most cases. Xue *et al.* (2004) proposed a trusted model based on linear prediction and gave the prediction value of cluster traffic based on NSFNET. Also, it concluded that the time sequence based on network traffic could establish model through low level ARIMA model. Xue and Qiu (2005) analyzed network traffic on the basis of actual measure data and built ARIMA model whose prediction error is about 5% in a small prediction step. The time sequence had been analyzed in the research progress of network traffic prediction and the short-related and long-related in time sequence model were described by Shenmin *et al.* (2006), also compared and analyzed their characteristics; Wang and Shan (2002) proposed a new method which used fraction difference, fuzzy prediction and fraction integral to predict network traffic, because the fuzzy auto-regressive model is fine ability to handle non-stationary and nonlinear data, the method can predict actual network traffic more accurately.

Although, network traffic has characteristics related complex, the output has short-related characteristics in identical scale through wavelet transformation. Riedi *et al.* (1999) proposed an indirect prediction method of network traffic based on wavelet transformation and the experimental results explained that this method could better predict network traffic of LAN and WAN. A multi-scale model is proposed to combine long-related characteristics network traffic data by Cong and Han (2003). Tang (2006) uses multi-fractal model of network traffic based on wavelet transform to analyze network traffic self-similar characteristics. Liu *et al.* (2000), it compares merit with demerit in the grey prediction method and supports vector machine method. A new prediction model-grey support vector machine prediction method has been presented by combining with the former two methods. This method not only has accumulation generation of grey prediction method but also weakens the influence from random disturbance factors in original sequence, enhance data regularity and avoid theoretical defects in grey prediction method and model. Liu *et al.* (2000) uses BP neural network to predict network traffic through one step five steps and ten steps prediction and uses experiment to test prediction effect. Since, neural network can train samples constantly and correct own parameters timely, prediction effect is better.

From the above, studies on IP network traffic modeling and predicting brought close attention from scholars around the world. They have done a lot of researches on it, but the current models are the single prediction models with widespread defects in adapting and prediction accuracy. In addition, the current combination model is constant weight one and can't conduct dynamic prediction.

IMPROVED RESIDUAL GREY PREDICTION MODEL

Residual grey prediction model is on the basis of GM (1, 1), residual sequence is from the difference between true value and prediction value. That is $E = \{\varepsilon(1), \varepsilon(2), ..., \varepsilon(n)\}$ and $\chi(\acute{\iota}) = \chi(\acute{\iota}) - \chi(\acute{\iota})$. Rebuild GM (1, 1) model for residual sequence E can obtain prediction value of residual sequence:

$$\hat{\varepsilon}^{(1)}(k+1) = (\varepsilon^{(0)}(1) - \frac{b''}{a''})e^{-ak} \tag{1}$$

Then, network traffic prediction value:

$$\hat{\chi}(k+1) = (1 - e^{\hat{a}})(\chi^{(0)}(1) - \frac{\hat{b}}{\hat{a}})e^{-\hat{a}k} + (1 - e^{\hat{a}'})(\varepsilon^{(0)}(1) - \frac{\hat{b}'}{\hat{a}'})e^{-\hat{a}'k} \tag{2}$$

When residual sequence of grey model is non-negative number, the residual grey model could be built. However, the residual sequence E produced by the GM (1, 1) modeling of network traffic needn't to meet the condition of non-negative sequence, so this study improves the grey model.

First, build GM (1, 1) model for IP network traffic sequence, second obtain residual sequence through difference between true value and prediction value. The third, deal with the residual sequence by absolute value and indexation to get two new sequences E_1 and E_2. By building GM (1, 1) model after the operating of BX data generation on E_1, using the model to predict and operating prediction value of inverse BX data generation, the prediction value of numerical part γ (i+1) on residual sequence E is got. For the indexation E_2, using the same operation can get the prediction value of E_2, then determine the value of current symbolic part h (i+1) by prediction value according to advance symbol criterion rule. Finally, the value of the symbolic part and numerical part on prediction value of residual sequence can be determined by E_1 and E_2.

$$\hat{e}(i+1) = h(i+1)\hat{\gamma}(i+1) \tag{3}$$

Then prediction value of network traffic is:

$$\hat{x}(k+1) = (1 - e^{\hat{a}})(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}})e^{-\hat{a}(k)} + h(k+1)\hat{\gamma}(k+1) \tag{4}$$

BP NEURAL NETWORK MODEL

Determination of neuron number in the input layer and hidden layer: In order to establish neural network, neuron number in the input layer and hidden layer must to be determined firstly. But up to now, there is not a theory to guide how to choose the optimal neuron number but to determine it often based on experience or experiment. Moreover, the learning performance and generalization ability of the network is to be tested by experimental comparison.

Pay attention to the choice of sample number to research the problems of time sequence. Neuron number in the input layer is corresponding to the length of retrospective time. Prediction performance of neural network relates to the input neuron number and its output value is related to its previous value. According to the choosing method for neuron number in the hidden layer, the experiment selected the original value 5 as the neuron number in the hidden layer according to the method of choosing the neuron number in the hidden layer, separately setting to 4, 5, 6 as the number of neurons in the input layer by basic BP algorithm, to test the network training and generalization ability, the results are shown in Table 1.

From the Table 1, when neuron number in the input layer is different, the neural network shows different prediction ability towards the same time sequence. The neuron number in the input layer should differ from each other when predict different time sequences. When neuron number ranges from 4 to 6 in the input layer, the value of Equality Coefficient (EC) is over 0.9 which indicates that the generalization ability of neural network is preferable. When neuron number in the input layer is 5, the test after training Mean Absolute Error (MAE), Mean Relative Error (MSE) and Sum of Squares Errors (SSE) has each got its minimum value, while EC gets the maximum.

Neuron number in the hidden layer is associated with complexity of problems to be solved and the amount of input and output, which have great influence on the

Table 1: Performance comparison of input layer neurons

NNIL	TT	SSE	MAE	MSE	EC
4	14375	0.0061	7.2316	0.16742	0.93473
5	10508	0.0013	2.0143	0.10238	0.94626
6	12973	0.0046	5.7876	0.29147	0.93596

NNIL: Neuron number in input layer, TT: Training times, NNHL: Neuron number in hidden layer

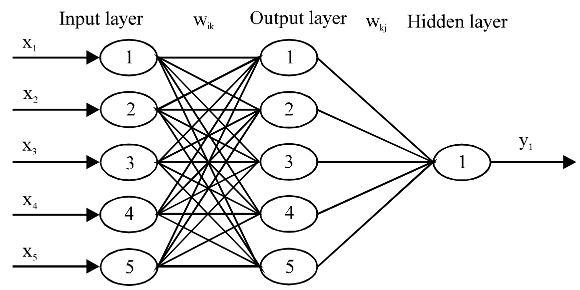

Fig. 1: 5×5×1 of the BP neural network

performance of network. So far there is no perfect theoretical basis, so it is difficult to determine the neuron number in the hidden layer only to rely on the experience. If the neuron number in the hidden layer is too little, neurons network will be hard to deal with the complex question and modeling will be not perfect. Otherwise, it not only makes structure complex, computation large and iterative learning time long on neural network but also maybe provide excess degrees of freedom to adapt noise which will make the network training excessive and lead to errors larger and so on. According to the experiment of backtracking time, combine the empirical formula of neuron number in hidden layer as reference, the network with different neuron number in hidden layer conducted train and test of generalization ability when the neuron number in input layer is separately set among 4~6. From the experiment about comparing the different combinations with neuron number in hidden layer and input layer, both the neuron number in input layer and hidden layer are 5, the BP network's generalization ability is best.

Establishment of BP neural network: The 5×5×1 structure of BP neural network according to the BP neural network training result and shown in Fig. 1, the input layer and hidden layer transfer function is Sigmoid type and the output layer is linear transfer function. Due to the rang of Sigmoid function is (0, 1), set the maximum learning number with 20,000 times , learning rate is 0.01, and take the sum of squared errors 0.1 as learning objectives, set up network connection weights of the initial value of random Numbers in (-1, 1).

FSVCPM

The object of single model which constitutes the combination prediction model are more, such as grey prediction model, neutral network model, ARMA and Wavelet model. For constituting a combination prediction model, the two or more models can be selected. In this study, combination prediction model is structured by improveiing residual grey prediction model and the BP neural network prediction model, because the grey prediction model can describe better in the linear features of network traffic while the BP neural network is better able to depict the non-stationary and nonlinear characteristics.

Determination of fuzzy strategy and fuzzy weights: Many scholars have been exploring how to combine the information with different models so as to improve the target of prediction accuracy. Thus, weight coefficient is the key to achieving the goal. This study gives a method of determining the variable-weight coefficient which makes many single models combined preferably to achieve the desired effect.

In network traffic prediction, assuming there are N (N>2) kinds of prediction methods and the fuzzy control of each method includes two aspects:

At the moment of i, the relative error of the j-method is expressed by e_j (i):

$$e_j(i) = \frac{y(i) - f_j(i)}{y(i)} \tag{5}$$

C_j (i) is the specific value between the different true value when the time is in i moment and the arithmetic mean value and the arithmetic mean value before m stage before m stage, as follow:

$$c_j(i) = \frac{y(i) \sum\limits_{J=i-m-1}^{i} y(i)}{\frac{1}{m} \sum\limits_{j=i-m-1}^{i} y(i)} \tag{6}$$

the fuzzification process of relative error as follows: e_j (i) is set to a sequence which is between (-1, 1), then the sequence is divided into several parts, each section corresponds to unrelated set, therefore, a unrelated set of X is formed. This study uses A to mark the semantic variables of relative error, so this discrete is divided into five semantic values, as follow in Table 2:

In fact, the C_j (i) value may not exist in (-1, 1), however in (-M, M) (M is an integer which is determined by some special factors) but can convert it into (-1, 1) with formulate (7):

Table 2: The semantic value of prediction relative error in discrete set

UN	X_1	X_2	X_3	X_4	X_5
	[-1,-0.6]	(-0.6,0)	0	(0,0.6)	[0.6,1]
SV	A_1	A_2	A_3	A_4	A_5
	NB	NS	ZE	PS	PB

UN: universe; SV: semantic value

Table 3: Fuzzy control rules

Fuzzy	Weight	NB	NS	ZE	PS	PB
Error	NB	PB	PB	PB	PS	ZE
	NS	PB	PS	PS	ZE	ZE
	ZE	PB	PS	ZE	NS	NB
	PS	ZE	ZE	NS	NS	NB
	PB	ZE	NS	NB	NB	NB

$$e_j(i) = \frac{c_j(i)}{M} \tag{7}$$

At the time i of the j-method, E is the deviation size between prediction value and actual value. C is the deviation degree between prediction value and actual value. K_{ij} represents the fuzzy weight. Based on the above control rules: if E = e_j(i) and C = c'_j(i) and then K_{ij} = k'_j (i). A fuzzy strategy is designed by adopting two input (e_j(i), c'_j(i)) and one output K_{ij}. This study gets a fuzzy control rule table (Table 3) based on the tracking desired trajectory of semantic space. It is necessary to adjust the rule, since the new table couldn't achieve better control effect.

Therefore ,according to the prescribed formula, the K_{ij} is transformed into a precise weight K'_j(i) ,then gets a fuzzy weight value of j-method through by using the formulate (8) at the i+1 moment.

$$k_j(i+1) = \frac{k'_j(i)}{\sum\limits_{j=1}^{n} k'_j(i)} \tag{8}$$

Determination of Self-adaptive mechanism and basic weights: Using the relative error e_j(i) of the j-method at the moment of i, the actual values of the former m stages y (i-m), y (I-m+1),..., y (i-1) at the moment of i and the D-value C_j (i) between the true values in the moment i and the arithmetic mean value of the former m stages, this stduy could predict y(i+1), the grey weight r (e_j(i)) and s (c_j(i)) can be measured by Eq. 9 and 10.

$$r(e_j(i)) = \frac{\mu + \lambda\theta}{|e_j(i)| + \lambda\theta} \tag{9}$$

$$s(c'_j(i)) = \frac{\mu + \lambda\theta}{|c'_j(i)| + \lambda\theta} \tag{10}$$

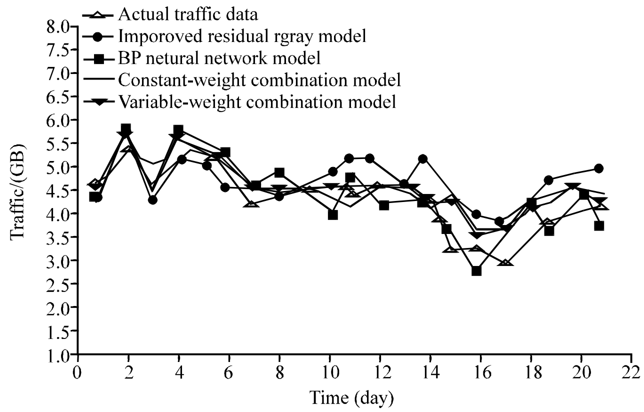

Fig. 2: VCPM prediction results for IP network traffic (3 weeks)

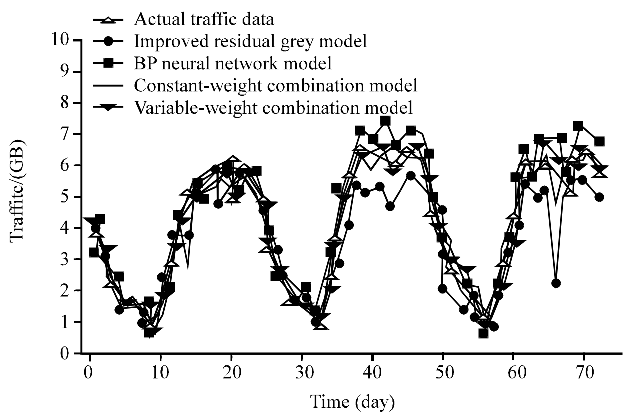

Fig. 3: VCPM prediction results (3 days)

Table 4: VCPM error analysis (3 weeks)

	ME	RE	VAR
IRGM	0.6840	9.16	0.853
BPNN	0.5362	7.25	0.721
CCPM	0.4278	5.83	0.577
VCPM	0.3943	5.46	0.541

ME: Mean error, RE: Relative error, VAR: Variance, IRGM: Improved residual grey model, BPNN: BP Neural network

$\mu = \min_j \min_i |y(i)\text{-}f_j(i)|$, $\lambda = \max_j \max_i |y(i)\text{-}f_j(i)|$, = g [fj (i-m), f_j (i-m+1),..., f_j (i-1)), θ is the grey correlation distinguish coefficient, which often is taken by 0.5.

The basic weight l_j of the j- method at the moment of (i+1) is:

$$l_j(i+1) = \omega_i \gamma(e_j(i)) + (1-\omega_i)s(c'_j(i)) \tag{11}$$

where, ω_i is an adaptive control coefficient, $0<\omega_i<1$ and it is decided by Eq. 12:

$$\omega_i = 1-(\frac{i-1}{i})^N \tag{12}$$

N is a positive and N = 0.5 in general.

Table 5: VCPM error analysis (3 days)

	ME	RE	VAR
IRGM	0.5640	7.57	0.753
BPNN	0.4854	6.82	0.586
CCPM	0.4178	5.76	0.534
VMCM	0.3719	5.24	0.518

Establishment of FSVCPM: Through the argument above, this study gets fuzzy adaptive weight of the j-method at the moment of i+1:

$$K_j(i+1) = \frac{l_j(i+1)k_j(i+1)}{\sum_{j=1}^{n} l_j(i+1)k_j(i+1)} \tag{13}$$

Equation 13 describes the weight that this method is in influence of prediction effect which makes the weight of each method more reasonable and improves the prediction accuracy greatly. The network traffic prediction value f (i+1) can be got after getting all the fuzzy adaptive k_j (i+1) of every kind of prediction method by using the determination method of fuzzy variable-weight at the moment of i+1.

$$f(i+1)=\sum_{j=1}^{n}k_j(i+1)f_j(i+1) \tag{14}$$

EXPERIMENTAL RESULTS AND ANALYSIS OF FSVCPM

As mentioned above, experimental data was acquired from a certain IP MAN operator from November 1, 2008 to January 31, 2009, real-time traffic information was collected once every 5 minutes and there were 26496 network traffic records in total. This stduy selected 20 different moments and did a set of experiments each moment, then 20 sets of experiments had been done, at last the experimental error analysis table was obtained from the statistics data of the 20 sets of experiments.

First, take the data of network inflow traffic as training samples at 12:30:00 every day in 5 weeks from December 7, 2008 to January 10, 2009, then predict the network inflow traffic at the time every day in the next 3 weeks (i.e., January 11, 2009 to January 31, 2009). The results are shown in Fig. 2. At last select another moment to do 20 sets of experiments, the error analysis are shown in Table 4. Next, take the network inflow traffic as training samples at every integral point (such as 8:00am, 9:00am, 10:00am, etc) in 5 days from January 1, 2009 to January 5, 2009, then predict the network traffic at the time of every hour in the next 3 days from January 6 to January 8. The results are shown in Fig. 3. At last Select another moment to do 20 sets of experiments, the error analysis are shown in Table 5.

The results of experiments above showed that in the prediction of big time granularity, VCPM is the best in prediction precision and prediction performance and the ability of adaptation is the greatest, CCPM is the second and Single model is poorer than Combination Model in prediction performance.

CONCLUSION

IP Network Traffic shows the obvious fractal, long-rang correction, self-similarity and characteristics of sudden, thus traditional IP network traffic models (For example: Poisson model, Markov model, auto-regression model, ARMA model) couldn't depict these traffic characteristics effectively. The combination prediction method has aroused more and more attention. If many single models are combined correctly. This combination can make combination prediction results less sensitive to the single prediction model which has large deviation. The single model's mathematical characteristic, prediction accuracy and weight coefficient have important influence on combination prediction results. A grey prediction model was established based on residual via the improvement indexation processing of grey model's residual sequence. By training experiments, the neuron number in input layer and hidden layer is identified and corresponding prediction model of BP neural network is built. Then FSVCPM is brought out on the basis of the residual improvement grey prediction model and the BP neural network model. Experimental results show the VCPM is the best in prediction precision and prediction performance and the greatest in the adaptive ability followed by CCPM and Single model is poorer than Combination Model in prediction performance. In addition, the VCPM could make dynamical prediction since its weight is fuzzy self-adaptive.

ACKNOWLEDGEMENT

Present study was supported by Project 10JJ9025 of the Hunan Provincial Natural Science Foundation of China, Project 2009GK3036 of the Hunan Provincial Science and Technology Plan, Project 10C1185 and 09C1162 of the Hunan Provincial Education Office Science Research of China.

REFERENCES

Cong, S. and L. Han, 2003. Multifractal model of network traffic based on discrete wavelet transform. J. Commun., 24: 3-8.

Feldmann, A., A. Greenberg, C. Lund, N. Reingold and J. Rexford, 2000. Netscop: Traffic engineering for IP network. IEEE Network Magaz., 14: 11-19.

Groschwitz, N.K. and G.C. Polyzos, 1994. A time series model of long-term NSFNET backbone traffic. Proc. IEEE Int. Conf. Commun., 3: 1400-1404.

Leland, W.S., M. Taqqu, W. Willinger and D. Wilson, 1993. On the self-similar nature of Ethernet traffic. Comput. Commun. Rev., 23: 183-193.

Leland, W.S., M. Taqqu, W. Willinger and D. Wilson, 1994. On the self-similar nature of Ethernet traffic. IEEE/ACM Trans. Networking, 2: 1-15.

Liu, S., T. Guo and Y. Dang, 2000. The gray system theory and its application. Science and Technology Press, Beijin.

Riedi, R.H., M.S. Crouse, V.J. Ribeiro and R.G. Baraniuk, 1999. A multifractal wavelet model with application to network traffic. IEEE Trans. Inform. Theory, 45: 992-1018.

Sang, A. and S.Q. Li, 2000. A predictability analysis of network traffic. Proc. INFOCOM, 1: 342-351.

Shenmin, H. X. Weisheng, W. Zhongjie and Y. Youling, 2006. Network traffic modeling and prediction based on fractional differencing and fuzzy-AR. Comput. Eng. Appli., 42: 104-106.

Tang, W., 2006. A new prediction model based on grey support vector machine. J. Syst. Eng., 4: 411-413.

Wang, J., X. Xiao and W. Gao, 2005. Small time granularity network traffic autoregression prediction analysis. Comput. Eng. Appl., 26: 129-132.

Wang, S., D. Xuan, R. Bettati and W. Zhao, 2004. Providing absolute differentiated services for real-time applications in static-priority scheduling networks. Networking, IEEE/ACM Trans., 12: 326-339.

Wang, X. and X. Shan, 2002. A wavelet-based method to predict Internet traffic. Proceedings of International Conference on Communications, Circuits and Systems and West Sino Expositions, 29 June-1 July, IEEE, pp: 690-694.

Wu, G., E.K.P. Chong and R. Givan, 2001. Congestion control via online sampling. Proceedings of 20th Annual Joint Conference of the IEEE Computer and Communications Societies, April 22-26, Anchorage, AK, USA., pp: 1271-1280.

Xue, K., Z. Li and L. Liu, 2004. The network traffic prediction based on ARIMA model. Microelectron. Comput., 4: 182-186.

Xue, L. and Z. Qiu, 2005. The application and research of time series analysis in network traffic prediction. Proceedings of the 10th National Conference on Youth Communication, (NCYC'05), Beijing, China, pp: 234-236.

Yang, J. and F. Wang, 2001. Intelligent network analysis and forecast support system. J. Wu Han Univ., 2: 75-78.

Zou, B. and Q. Liu, 2002. The network traffic prediction based on ARMA model. J. Comput. Res. Dev., 39: 1646-1652.