The 7th International Symposium on Operations Research and Its Applications (ISORA'08)
Lijiang, China, October 31–Novemver 3, 2008

The monitoring of the network traffic based on queuing theory

Wang Jian-Ping[1, *, †]   Huang Yong[1, ‡]

[1] School of Information Engineer, Henan Institute of Science and Technology, Xinxiang, 453003, china

Abstract Network traffic monitoring is an important way for network performance analysis and monitor. The article explains how to build the basic model of network traffic analysis based on Queuing Theory, using this, we can obtain the network traffic forecasting ways and the stable congestion rate formula, combining the general network traffic monitor parameters, we can realize the estimation and monition process for the network traffic rationally.

Keywords Queuing Theory; M/M/1 queue; traffic monitor; congestion rate; Markov Process

1 Foreword

Queuing Theory, also called random service theory, is a branch of Operation Research in Maths, it's a subject which researches the random regulation of queuing phenomenon, and builds up the maths model by analyzing the date of the network. Through the prediction of the system, we can reveal the regulation about the queuing probability and choose the optimal method for the system.

Adopting Queuing Theory to estimate the network traffic, which becomes the important ways of network performance prediction, analysis and estimation, through this way, we can imitate the true network, it's useful and reliable for organizing, monitoring and defending the network.

2 The maths model of the Queuing theory

In network communication, from sending, transferring to receiving datačňthe proceeding of the data coding, decoding and sending to the higher layer, With all these process ,we can find a simple queuing model. According to the Queuing Theory, this correspond procedure can be abstracted as Queuing theory model [9], like figure 1 showing to us. Supposing this kind of simple data transmitting system satisfies the queue model M/M/1[5].

*Supported by the natural science foundation of Education Department of Henan province grant NO. 2007520017

†He was born in 1981, in BaoJi, ShaanXi province of china, he is a master candidate, his research area is Computer Network and Operation Research.

‡He was born in 1970, in ChangSha, Hunan province of China, he is a lecturer, his research area is Computer Graphics and Computer Software.

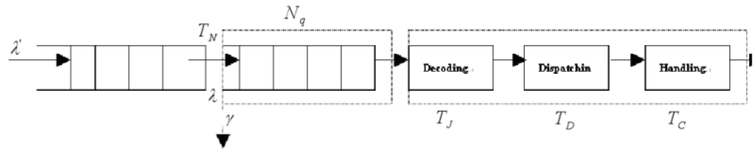

Figure 1: The abstract model of communication process

From the figure above, λ' stands for the sending rate of the sender, T_N is regarded as transportation delay time , λ means the arriving speed of the data packets,N_q is the quantity of data packets stored in the buffer,$\gamma$is the lost rate of the receiver ,and $T_S = T_J + T_D + T_C$ is the service time of data packets in the queue. T_J, T_D and T_C stand for the decoding, dispatching and handling time separately. The sender is abstracted to be a queue, whose sending rate is λ'.

In model M/M/1, the two M represent the sending process of the sender and the receiving process of the receiver separately. They both follow the Markov Process[2], also keep to Poisson Distribution, while the number 1 stands for the channel. Using N(t)=t as the length of the queue at the moment of t, so the probability of the queue whose length is n ,can be shown as the following formula .

$$p_n(t) = p[N(t) = n] \tag{1}$$

According to the formula 1,we can know the equation 2:

$$\begin{cases} p_n'(t) = \lambda p_{n-1}(t) + \mu p_{n+1}(t) - (\lambda + \mu) p_n(t), n = 0,1,2\ldots \\ p_0'(t) = \mu p_1(t) - \lambda p_0(t) \end{cases} \tag{2}$$

In equation 2, supposing λ as the arrival rate while μ as the service rate, when $t \to \infty$, we can get the stable solution of the equation $p_n(t)$.

$$p_n(t) = \rho(1-\rho)n = 0,1,2,\ldots \tag{3}$$

In formula 3, using $\rho = \lambda/\mu \leq 1$ as the probability of the service. N stands for the length of the queue under the balanced condition, It includes the average volume of all the data packets which enter the processing module and store in the buffer.

$$N = \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} p^n (1-\rho) = \frac{\rho}{1-\rho} \tag{4}$$

If N_q shows the average volume of the buffer's data packets, We can conclude the following formula

$$N_q = N - \rho = \frac{\rho}{1-\rho} - \rho = \frac{\rho^2}{1-\rho} \tag{5}$$

If the processing module is regarded as a closed region, the parameter is brought into the formula 5,using the Little Law, we can get the equation(6)

$$\begin{cases} \rho = \lambda T_S \\ \lambda = \lambda^{'} \end{cases} \tag{6}$$

After putting equation (6) into the equation (5) and sorting it out, we can get the formula(7)

$$(\lambda^{'} T_s)^2 + \lambda T_S N_q - N_q = 0 \tag{7}$$

According to the derivative formula (7), we make a conclusion, among three variables, the service time ,the sending rate, and the buffer, if we know any two variables ,it is easy to gain the numerical value of the third one .So these three variables are the key parameters for measuring the performance of the transmission system.

3 Queuing theory and the network traffic monitor

3.1 The method of monitoring the network traffic

Traffic monitoring is the basic way for scheduling, concerting and operating the net[7]. Network usually uses Peak Information Rate(PIR),Peak Burst Length(PBL),Commitment Information Rate(CIR),Commitment Burst Length(CBL),Extra Information Rate(EIR)and Extra Burst Length(EBL),the six parameters as the network traffic monitoring parameters.

Using the six parameters above, we can realize two kinds of Bandwidth Measurement Algorithms, they are PIR/CIR Algorithms and CIR/EIR Algorithms[8], PIR/CIR uses four parameters, they are PIR ,PBL,CIR and CBL.CIR/EIR also uses four parameters, they are CIR,CBL, EIR and EBL. In these algorithms, each of the PIR/PBL,CIR/CBL, EIR/EBL informs a Token Bucket partly[6] ,from this we know that every algorithms contains two token buckets, so they are called dual-token buckets algorithms.

3.2 Forecasting the network traffic using Queuing Theory

The network traffic is very common[4], What's worse, when the traffic becomes under extreme situation, it will lead to the net's paralysis[1]. There are a great deal of research about monitoring the congestion at present ,besides, the documents which make use of Queuing Theory to research the traffic rate appear more and more. for forecasting the traffic rate, we often test the data disposal function of the router used in the network.

Supposing a router's arrival rate of data flow in groups is λ, and the average time which the routers use to dispose each group is $\frac{1}{\mu}$, the buffer of the routers is C, if a certain group arrives, the waiting length of the queue in groups has already reached, so the group has to be lost. When the arriving time of group timeouts, the group has to resend, Supposing the group's average waiting time is $\frac{1}{\mu}$, we identify $p_i(t)$ to be the arrival probability of the queue length for the routers group at the moment of t, supposing the queue length is i:

$$p(t) = (p_0(t), p_1(t), \ldots, p_{n-1}(t)), i = 0, 1, \ldots, C+1 \tag{8}$$

Then the queuing system of the router's date groups satisfies simple Markov Process[10], according to Markov Process, we can make the diversion strength of matrix as follow:

$$Q=\left\{\begin{array}{cccccccc} -\lambda & \lambda & & & & & & \\ \mu & -(\lambda+\mu+\nu) & \lambda+\nu & & & & & \\ & \mu & -(\lambda+\mu+2\nu) & \lambda+2\nu & \ldots & & & \\ & & & \vdots & & & & \\ & & & & & \mu & -(\lambda+\mu+C\nu) & \lambda+C\nu \\ & & & & & & \mu & -\mu \end{array}\right\} \tag{9}$$

3.3 Network Congestion Rate

Network congestion rate is changing all the time [3], the instantaneous congestion rate and the stable congestion rate are often used to analysis the network traffic in network monitor. The instantaneous rate $A_C(t)$ is the congestion rate at the moment of t .The $A_C(t)$ can be obtained by solving the system length of the queue's probability distributing, which called $p_{n-1}(t)$.According to some properties of Markov Process, we know that$p_i(t)(i=0,\ldots,C+1)$satisfies the following differential equation systems:

$$\left\{\begin{array}{l} p_0^{'}(t)-\lambda p_0(t)+\mu p_1(t) \\ p_{\lambda}^{'}(t)=[\lambda+(k-1)\nu]p_{k-1}(t)-[\lambda+\mu+k\nu]p_k(t)+\mu p_{k+1(t),}k=1,\ldots,C \\ P_{C+1}^{'}(t)=[\lambda+C\nu]p_C(t)-\mu p_{C+1}(t) \\ p_0(0)=1,p_2(0)=p_3(0)=\ldots=p_{C+1(0)=0} \end{array}\right. \tag{10}$$

By solving this differential equation systems, we get the instantaneous congestion rate $A_0(t)$ is

$$A_0(t)=P_1(t)=\frac{\lambda}{\mu+\lambda}(1-e^{-(\mu+\lambda)t}) \tag{11}$$

The instantaneous congestion rate can't be used to measure the stable operating condition of the system, so we must obtain the stable congestion rate of the system. The so-called stable congestion rate means it will not change with the time changing, when the system works in a stable operating condition. The definition of the stable congestion rate is

$$A_C(t)=\lim_{t\to\infty}A_C(t) \tag{12}$$

Supposing $\rho=\lim_{t\to\infty}p(t)$ as the distributing of the stable length of the queue, C as the buffer of the router, the stable congestion rate can be obtained in two ways: firstly, we obtain the instantaneous congestion rate ,then make its limit out . According to its definition, it can be obtained with the distributing of the length of the queue. Secondly, according to the Markov Process, we know that the distributing of the stable length of queue can be get through evaluating equation (13)

$$\left\{\begin{array}{l} pQ=0 \\ \sum_{i=n}^{C+1}P_i=1 \end{array}\right. \tag{13}$$

In the equation $p = (p_0, p_1, \ldots, p_{C+1})$, When C=0, the equation (13)'s evaluation is

$$A_0 = P_1 + \frac{\lambda}{\mu + \lambda} \tag{14}$$

When C=1, the equation (13)'s evaluation is

$$A_1 = P_2 = \frac{\lambda(\lambda + \nu)}{\lambda(\lambda + \nu) + \mu(\mu + \lambda)} \tag{15}$$

When C =2, the equation (13)'s evaluation is

$$A_3 = P_4 = 1 - \frac{\mu}{(\lambda + \mu + 2\nu)A_2 - (\lambda + \nu)(1 - A_1)A_0 + \mu} \tag{16}$$

on the analogy of this , we conclude that ,the stable congestion rate is

$$A_C = P_{C+1} = 1 - \frac{\mu}{(\lambda + \mu + C\nu)A_{C-1} - (\lambda + (C-1)\nu(1 - A_{c-1})A_{C-1} + \mu} \tag{17}$$

4 Conclusion

The article explains how to build the queue model based on the Queuing Theory, how to realize calculation and analyzing the network traffic through M/M/1 Queuing Theory, and then gets the forecast way and the formula of the stable congestion rate of the network traffic. The network traffic monitoring model which is tested , experimented and analyzed by the actually system ,shows that using the Queuing Theory will optimize the network traffic, it's convenient and simple for calculating and monitoring the network traffic properly.

References

[1] Wang Ting, Wang Yu. Survey on a Queue Theory Based Handover Scheme for UAVS Communication Network. Chinese Journal of Sensors and Actuators, 2007, 04

[2] Li Da-Qi, Shen Jun-Yi. Queuing Theory Supervising K-Means Clustering Algorithm and ITS Application in Optimized Design of TTC Network .Journal of Astronautics, 2007, 03

[3] Han Jing, Guo Fang, Shi Jin-Hua. Research on the traffic monitoring of the distributed network based on human immune algorithm. Microcomputer Information, 2007, 18

[4] Wang Pei-Fa, Zhang Shi-wei, Li Jun. The Application and Achievement of SVG in Network Netflow Monitor Field. Microelectronics & Computer, 2005, 04

[5] Ren Xiang-Cai, Xiong Qi-Bang. An Application of Mobile Agent for IP Network Traffic Management .Computer Engineering, 2002, 11

[6] Davison B D. A Web Caching Primer. IEEE Internet Computing, 2001, 5(4)

[7] Guo Yang, Gong Wei-Bo, Don Towsley. Timestepped Hybrid Simulation (TSHS) for Large Scale Networks. In: Proceedings of IEEE Infocom 2000. Israel: IEEE, 2000

[8] Jong Suk Ahn, Peter B Danzig. Packet Network Simulation: Speed up and Accuracy Versus Timing Granularity. IEEE/ACM Trans on Networking, 1996, 4(5)

[9] Vern Paxson, Sally Floyd. Why We Don' t Know How To Simulate the Internet. In: Proceedings of the 1997 Winter Simulation Conference. USA: ACM, 1997

[10] Gunther N. The Practical Performance Analyst. New York: McGraw-Hill, 1998

[11] ITU-T G.729A. Coding of Speech at 8kbps using Conjugate-Structure Algebraic Code-excited Linear-prediction Coding. USA: ITU-T, 1998