Characteristics of the Token Bucket parameters with Self Similar Network Traffic

DASHDORJ Yamkhin and
CHULUUNBAATAR Nergui
Computer Science and Management
School of Mongolian University
of Science and Technology
Ulaanbaatar, Mongolia zip:13-381
email: dashdorj@mail.mn,
chuka@csms.edu.mn

BATZUL Bold
Dept of Information
System Management of
Institute of Finance and Economics
Ulaanbaatar, Mongolia
email: batzul@ife.edu.mn

SEONGJIN Lee
Dept of Electrical and Computer
Engineering, Hanyang University,
Seoul Korea
email:james@ece.hanyang.ac.kr

***Abstract*—We investigate the influence of self similarity features of the input traffic on the performance of some QoS mechanisms. We model the token bucket of input traffic characteristics on the queue distribution of FBM (Fractional Brown Motion) and find the low bound of token bucket mechanism, at input traffic with the self similarity (Long range dependent).The result is token bucket parameter how to dependent the degree of the self similarity of the traffic.**

Keywords: Token bucket (TB), self-similar network traffic, long range dependent, Fractional brown motion (FBM).

I. Introduction

Over the past several decades network and communication technology has been a significant and growing component of Internet traffic. Integrated broadband networks are expected to support various traffic types such as data, voice, image and video. Traffic generated from these services is substantially different in its statistical characteristics and networks are required to maintain a certain level of throughput during each session for these services. For example, real-time voice communication over computer network requires several Kbits/sec of network bandwidth and further is delay sensitive. To effectively exploit the token bucket of underlying network and to maximize QoS, it is mandatory that we have proper performance model and use this model to used in diffserv environment and shaping a police of burst network traffic to allocate resource for efficient service support. We had analyzed the stochastic characteristics of wireless network traffic in previous work [1] which collect full packet trace from up-and-running IEEE 802.11b network. A number of existing works found that network packet traffic bears long-range dependent property[2], [3], [4]. Discovery of scaling behavior in the measured teletraffic leads to model solutions that can approximate the data characteristics much better than previous techniques. Self-similar processes have been used to successfully model data, which exhibits long-range dependency in a variety of different scientific fields, including [5], geophysics[6], biology[7], telecommunication networks[8], and economics. We develop a the token bucket model for packet traffic of wireless network. We apply Fractional Brownian Motion to model the incoming packet process. Our model accurately models the token bucket and up bound of token bucket of the underlying self-similar network traffic.

II. Related Work

Gilberto Mayor and John Silvester[9] show that an ATM queueing system driven by this long-range dependent source can exhibit very long busy period. They introduce a probabilistic fractional Brownian motion envelope process and develop a method to compute the Token Bucket parameters by using the FBM envelope process. L.Janowski and Z. Papir[10] present a burstiness curve- a tradeoff between Token descriptors-derived for a FBM traffic model using an envelope process. The burstiness curve is defined by traffic parameters (mean, variance, and Hurst parameter) and the considered Token Bucket packet drop probability. Paolo Giacomazzi and et al. [11], are proposed a new approach for the selection of the token bucket parameters on the LBAP (Linear Arrival Process) curve, by taking into account two delay constraints, that is, the delay generated by the regulator itself and the end-to-end network delay and to choose the token bucket parameters that minimize the capacity to be assigned in the network in order to guarantee the delay performance negotiated for the traffic flow. Halabi Hasbullah[12] et al. are to develop a traffic-descriptor, which has only least parameters but well describing the system . A Mathematical model that measures level of burstiness, degree of self-similarity H, and quality of the forwarding link must be developed. The QoS provisioning at the time when resources are to be allocated to the requesting application and at the time when routing decision is to be made to select a forwarding link would then be more accurate and guaranteed[12]. Whereas our study focuses on small time scale statistical characteristics, like estimating the token bucket parameters based on an approximation using FBM process, the estimating the Hurst parameter H and using the Noros equation[13]. In our case, we calculated the theoretical token bucket parameters by using Norros equation[13] and dependency between token rate and bucket size. The paper is organized as follows. In Section 2 we discuss related work

978-1-4244-9037-0/10/$26.00

and Section 3 explain long range dependent properties and Fractional Brown Motion. Section 4 shows measurement, in section 5 the description of the token bucket and analysis for token bucket parameters of token rate and bucket dept with dependent of Hurst parameters and in section 6 for conclusion. I wish you the best of success.

III. LONG-RANGE DEPENDENCE

Long-range dependence is defined in terms of the behavior of the auto covariance $C(\tau)$ of a stationary process as τ increases. For many processes, the auto covariance rapidly decays with τ. For the Poisson increment process with increment L and mean λ, the auto covariance for values of $\tau > L$ is in [1], [4], $C(\tau) = R(\tau) - \lambda^2 = \lambda^2 - \lambda^2$. In general, a short-range dependent process satisfies the condition that its auto covariance decays exponentially:$C(k) \sim a^{|k|} as |k| \to \infty, 0 < a < 1$. The type of data traffic models typically considered in the literature or in the papers employs only short-range dependent processes. Using the equality $\sum_{k=0}^{\infty} x^k = \frac{1}{(1-x)}, |x| < 1$ we can observe that $\sum_K C(K)$ for a short-range dependent process is finite. In contrast, a long-range dependent process has a hyperbolically decaying auto covariance: $C(K) \sim |k|^{-\beta}$ as $|k| \to \infty, 0 < \beta < 1$ where β is the same parameter defined earlier and is related to the Hurst parameter as $H = 1 - \frac{\beta}{2}$. In this case, is $\sum_K C(K) = \infty$. The variances of the aggregated self-similar processes $X^{(m)}, m \geq 1$, decrease more slowly than the reciprocal of the non-overlapping batch size m. This property is given by in [2]: $Var[X^{(m)}] \to cm^{-\beta}$ when $m \to \infty, c$ is a constant and $0 < \beta < 1$. If $\beta = 1$ in this case that processes such as Poisson processes in [2] proved that $H = 1 - \frac{\beta}{2}$ [2]. Variance time function become is [15]: $Var[X_n^{(m)}] = \sigma^2 m^{-2(1-H)}$ where $\sigma^2 = \lambda a \tau^{2H}$ λ and a are incoming rate and variance[15], [2].

IV. DATA STUDY

A. Measurement Setup

Fig 1 shows the network configuration and connection used in this study. To cover wider geographical area, it is more cost, effective to use wireless network than to use wired network technology. Particularly, in a sparsely populated country like Mongolia, wireless network is preferred communication medium. In a wireless network there is one six-sector antenna system where each sector antenna approximately covers 60^0 degree angle and adjacent sector antennas slightly overlaps with each other. 40 wireless clients are connected to 2 Access Point of the provider. Routing of all connections, and also the control and management of throughput are carried out with a router. Each wireless client has throughput ranging from 64 up to 512 kbps. We use packet sniffer to collect the packet [17]. Sniffer is connected to the network so as to record traffic going through Point "1" and simultaneously through Point "2" in Fig 2. Please note that the point of "1" receiving traffic information sharing among wireless customers, and with it the traffic flow of information between customers and Internet. After a point "2" is only the latest of them. All packages are recorded down to the file format tcpdump. More than 12.7 million packages

TABLE I
TRAFFIC DATA SETS DESCRIPTION

Data sets	File description	Protocol layer
eth.dat	Aggregate traffic	2(Ethernet)
eth.src	Upstream traffic	2(Ethernet)
eth.dst	Downstream traffic	2(Ethernet)
tcp.dat	TCP traffic	4(TCP)

were collected in our study. Of these, 70 percent were used to construct the TCP datagram.

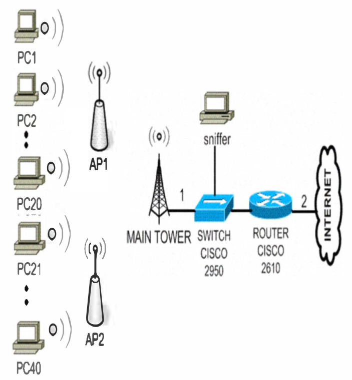

Fig. 1. *The wireless network configuration*

B. The Characteristics of Realizations

Packet was collected from March, 18th, 2005 (Wednesday) at 10:00 to March, 18th at 17:00. The duration is 7 hours; short description is given in table I. Data are presented with two columns in ASCII-format: the first column contains time labels (in sec), and the second column contains the size of the Ethernet-frame in bytes, or the size of a field of data of an IP-packet in case of a TCP-packet.

V. CHARACTERISTICS OF TOKEN BUCKET

A. Buffer Overflow Probability Approximation Analysis

We analyze the tail probability of a given queuing system with finite queue. We assume that incoming traffic follows FBM (Fractional Brownian Motion). FBM is one of the most widely known model for self-similar process[13]. Let us briefly explain fractional Brownian motion. FBM process, $A(t)$, is defined as in Equation (1).

$$A(t) = mt + \sqrt{ma} Z_{(H)}(t), \ t \in \Re \qquad (1)$$

where m and a denotes average rate and variance of arrival process. $Z_H(t)$ is Gaussian process with zero mean and variance of $Var[Z_{(H)}(t)] = |t|^{2H}$. H denotes Hurst parameter and satisfies $H \in [0.5...1]$. FBM traffic is modeled by the three parameters $FBM(m, a, H)$. Norres et. Al. established a relationship between queue length and overflow probability especially when incoming traffic bears long range dependent property Equation (2)[13].

$$\ln(Pr[Q > L]) = -\frac{1}{2ma(1-H)^2} \frac{C(1-\rho)(1-H)}{H}^{2H} * L^{2(1-H)} \qquad (2)$$

where the buffer size L, service rate C and the traffic parameters m,a and H for the bounder value[18].The analysis of a single queue construction with FBM at the input was presented for the first time in [13], where it was shown that the queue length distribution can be approximated by Weibull distribution. The approximation tail probability is follows:

$$Pr(Q > L) \approx \exp(-\gamma * L^{2(1-H)}) \tag{3}$$

where $L \rightarrow \infty, \rho = \frac{m}{c}$ and $\gamma = \frac{1}{2ma(1-H)^2}(\frac{C(1-\rho)(1-H)}{H})^{2H}$ is the offered load. If the observed traffic, that is, the traffic extracted traffic data with 1 sec can be approximated $H = 0.82$, $m = 2290kbps$,and $a = 280,9kpbs$. Thus, tail probability can be calculated with Equation(3). In previous studies, we analyzed dependency for parameters of variance and Hurst parameter[1].

B. The Token Bucket Model with Self Similar Network Traffic

We are explained basic model for the token bucket and how to dependent the some parameters from self similar network traffic. This token bucket model is used in diffserv environment. Token bucket represents the Policing function of traffic condition block of diffserv. A token bucket flow is defined by (r, b) r denotes the rate at which tokens are accumulated and b is the depth of the token pool (in bits). If the bucket is full, the incoming tokens will be thrown away. If a packet arrives to the token bucket then token number is verified. If there are enough tokens then the packet is sent immediately and token number is reduced by a packet bit length. Otherwise, the packet either waits for tokens or is dropped from service. The token pool is refilled at a constant rate, if the bucket is full, tokens are lost. The token bucket uses two parameters to control the connection. The first parameter is the bucket size b [bits] and the second one a token accumulating rate r [bps]. Dropping incoming packets undergoes two restrictions. The bucket size has to be greater than the maximum length of packets in transmission in order to prevent a systematic dropping of the longest packets. On the other hand, dropping packets can occur only in a so called congestion interval that is a contiguous time period with the token bucket being not fully replenished. For the congestion intervals a maximum accumulated amount of data sent through the token bucket is limited in Equation (4) to[9]:

$$L(t) = rt + b \tag{4}$$

where i is a congestion interval duration.
Pair (r, b) of token bucket parameters for the same drop probability create so called a burstiness curve $b = b(r)$ [10], [12].
Let us denote the overflow probability ε with maximum time interval as Equation (5):

$$\varepsilon = Pr(mt + \sqrt{am}Z_H(t) > rt + b) \tag{5}$$

where m-mean rate a-variance coefficient, r is token rate and b token pool size. We defined the input of FBM process in previous work [19]. Thus, FBM traffic has exactly second-order self-similarity and long-range dependence.

$$Pr\{Z_H(t) < \frac{(r-m)t+b}{\sqrt{am}}\} = \Phi\{\frac{(r-m)t+b}{t^H\sqrt{am}}\} \tag{6}$$

where $\Phi(\bullet)$ is the complementary distribution function of the standard Gaussian distribution.
The maximum congestion interval find from (6) equation until a token size reaches its peak, that t.

$$t_{max} = \frac{b * H}{(r-m)(1-H)} \tag{7}$$

where m-mean rate, r-token rate, b-token pool and H is hurst parameter $(0.5 < H < 1)$. We are doing some mathematical rearrangement and get the result which token pool size is dependent token rate and Hurst parameter. We get the following result and to plot the Fig. 2.

$$b = [\sqrt{-2am * \ln(\varepsilon)}(1-H)^{1-H}H^H(r-m)^H]^{\frac{1}{1-H}} \tag{8}$$

where m-mean input rate, a-variance coefficient, ε-token pool overflow probability, r-token rate, H-parameter and b token pool.
We derived the mathematical formula for token pool with dependent token rate and H-parameter. If hurst parameter is closer to $H = 0.8$, token pool size is slowly down at increasing token rate. If we changed the variance coefficient $a = 561.6kbps$ with H-parameter, $H = 0.8$ at increasing token rate is very slowly down, shown in Fig. 3. All these result are shown in Fig. 2, 3 which is dependent bucket size for token rate. If input rate is equal to token rate $(m = r)$, bucket size is zero from Equation (8). The bucket size is more dependent H-parameter and variance coefficient in Fig. 2, 3.

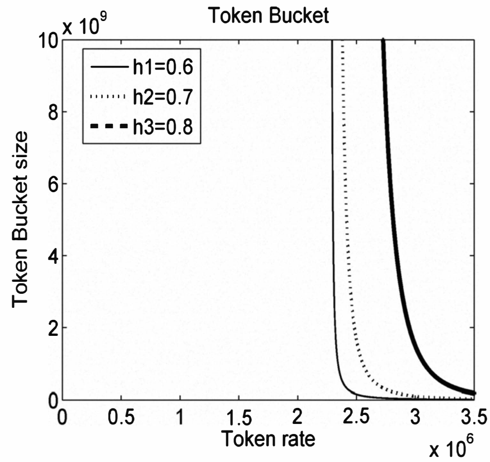

Fig. 2. *Token bucket of self similar network traffic with variance a=280.8kbps*

C. A Low Bound for Token Pool

At first we define that an overflow occurs whenever the arrival packet exceeds the token bucket size. In this case, we

978-1-4244-9037-0/10/$26.00

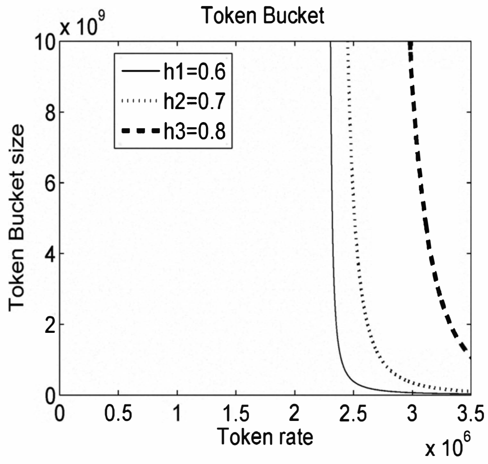

Fig. 3. *Token bucket of self similar network traffic with variance a=561.6kbps*

find the low bound for token bucket as following express:

$$\varepsilon = Pr\{m + \sqrt{am}Z_H(1) > b + r\} \tag{9}$$

We can express the Equation 9 as

$$Pr\{Z_H(1) < \frac{b+r-m}{\sqrt{am}}\} = \Phi(\frac{b+r-m}{\sqrt{am}}) \tag{10}$$

So we obtained the low bound of token bucket size for large buffers.

$$b = m - r + \sqrt{-2am\ln(\varepsilon)} \tag{11}$$

In Equation (11) we can find that the low bound does not dependent on the H parameter. For known traffic parameters (m, a, H) and the considered drop excess probability ε, we plotted the Fig. 2,3 proves that serving a traffic with a higher hurst parameter requests much slow token refilling rate r and deferent hurst parameter is required deferent token bucket. Then we derived the express for a low bound token bucket size in Fig. 4. It is no required the H parameter.

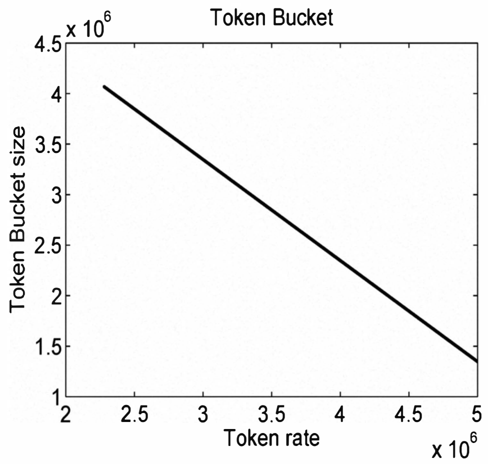

Fig. 4. *A Low bound of Token Bucket*

VI. Conclusion

In this study, we perform comprehensive analysis on 802.11b network traffic. We capture packet traces from the existing wireless LAN environment. Whereas our study focuses on small time scale, like estimated the tail probability based on an approximation using FBM tail probability, estimated a token bucket parameters and a low bound of token bucket parameters. Our case, we calculated the theoretical token bucket by using equation(8). How dependent between token bucket and token rate. We found that among the models tested, token bucket model with self similar traffic more accurately synthesizes the long memory characteristics of the underlying traffic. The results of our work can be used in many areas. They include a token bucket police planning in various network related hardware, and etc.

Acknowledgment

The authors would like to thank Professor T. Batbayar, Vice President for International Cooperation of Mongolian University of Science and Technology and Won Youjip of department of Electronics of Computer Engineering Hanyang University, Seoul, Korea

References

[1] Dashdorj Yamkhin and Won Youjip, *Modeling and Analysis of Wireless LAN Traffic.* Journal of Information Science and Engineering Nov 2009. Vol.25 No.6: p.1783-1801

[2] Willinger .W and K. Park, *Self-Similar Network Traffic and Performance Evaluation* 2000: A Wiley-Interscience Publication.

[3] Xianhai Tan, Yuanhui Huang and Weidong Jin *Modeling and Performance Analysis of Self-Similar Traffic Based on FBM.* 2007 IFIP International Conference, 18-21 Sept 2007: Liaoning of China, p. 543-548

[4] Rao Yunhua, Xu Zhongyang and Lei Xu, *Performance Analysis of Queue System with Self-Similar traffic input.* in ICCT 2003, International Conference on 9-11 April 2003 Seoul, Korea. p. 1437 - 1439

[5] Beran. J, *Statistics for Long-Memory Processes.* Vol. 39(1). 1994: Jeffrey Glosup Technometrics. p.105-106.

[6] Frisch. U, *Turbulence the Legacy of A.N. Kolmogorov.* Cambridge, 1995.

[7] Collins. J.J and C.J de Luca, *Upright, correlated random walks: A statistical-biomechanics approach to the human postural control system.* Chaos: An Interdisciplinary Journal of Nonlinear Science, March 1995. Vol.5 No.1: p. 54-63.

[8] Taqqu, M.S, V. Teverovsky and W. Willinger, *Estimators for Long-Range Dependence: An Empirical Study.* Fractals, 1995. Vol.3: p. 785-798.

[9] Gilberto Mayor and John Silvester, *Time Scale Analysis of an ATM Queueing System With Long-Range Dependent Traffic.* in Proceedings of the INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. 1997: Washington, DC, USA. p. 205-212.

[10] L. Janowski and Z. Papir, *Analysis of a Burstiness Curve for FBM Traffic and Token Bucket Shapping Algorithm.* in ConTEL'05 8-th International Conference on Telecomunications. June 15-17: Zagreb, Croatia. p. 439-442.

[11] Paolo Giacomazzi, Luigi Musumeci, Gabriella Saddemi and Giacomo Verticale,*Optimal Selection of Token Bucket Parameters for the Admission of Aggregate Flows in IP networks,* in GLOBECOM '06. IEEE 16 April 2006 San Francisco, CA p. 1 - 6

[12] Halabi Hasbullah, Mahamod Ismail and Kasmiran Jumari, *PARSIMONIOUS TRAFFIC-DESCRIPTOR FOR QOS ROUTING DECISIONS IN BLUETOOTH NETWORK.* Journal of Theoretical and Applied Information Technology, Sept 2009. vol.5 no.2: p. 119-128.

[13] Norros. I, *On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks.* IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATION, August 1995. VOL.6: p. 12-36.

[14] Stallings, W, *High-Speed Networks and Internets: Performance and Quality of Service.* Second Edition 2002: New Jersey: Prentice Hall Upper Saddle River.
[15] Willinger. W, M.S Taqqu , and A. Erramilli, *A bibliographical guide to self-similar traffic and performance modeling for modern high speed networks.* In Theory and Applications; Oxford University Press, 1996. Vol. 4: p. 339-366.
[16] Mandelbrot .B and J.R. Wallis, *Computer experiments with fractional Gaussian noises-Part 1: Averages and variances.* Water Resources Research, 1969. 5: p. 228-267.
[17] Matthews .J, *Computer Networking: Internet Protocols in Action.* 2005: WILEY Johon Wiley and Sons, Inc.
[18] Teymori .S and W. Zhuang, *Queue analysis and Multiplexing of Heavy-tailed Traffic in Wireless Packet Data Networks.* Mobile Networks and Applications, Springer, 2007. Vol. 12, No. 1: p. 31-41.
[19] Dashdorj Yamkhin and Won Youjip, *Modeling Heavy-tailed Behavior of 802.11b Wireless LAN traffic.* Journal of Digital Contents Society, Jun 2009. Vol. 10 No. 2: p. 357-365.

978-1-4244-9037-0/10/$26.00