Effect of the Variable Packet Size on LRD Characteristic of the MMPP Traffic Model

Kang-won Lee* ***Regular Members***, Byung-chun Kwon ** ***Associte Members***

ABSTRACT

The effect of the variable packet size on the LRD characteristic of the MMPP traffic model is investigated. When we generate packet traffic for the performance evaluation of IP packet network, MMPP model can be used to generate packet interarrival time. And a random length of packet size from a certain distribution can be assigned to each packet. However, there is a possibility that the variable packet size might change the LRD characteristic of the original MMPP model. In this study, we investigate this possibility. For this purpose the 'refined traffic' is defined, where packet arrival time is generated according to the MMPP model and a random packet length from a specific distribution is assigned to each generated packet. Hurst parameter of the refined traffic is estimated and compared with the original Hurst parameter, which is the input parameter of the MMPP model. We also investigate the effect of the packet size distribution on the queueing performance of the MMPP traffic model and the relationship between the Hurst parameter and queueing performance.

Key Words : MMPP, Hurst Parameter, Long Range Dependence

Ⅰ. Introduction

Traffic characterization and modeling is a crucial activity towards an efficient dimensioning and resource management of IP networks. A general consensus exists in the fact that Internet traffic is not Poisson at any level of aggregation. Recent traffic studies have shown that Internet traffic may exhibit properties of self-similarity and long range dependence. These characteristics have significant impact on network performance.

Over the last few years a number of attempts were made to develop models for LRD(Long Range Dependence) data traffic, which include multifractal model[17], FBM(Fractional Brownian Motion)[11], chatoc map[13], wavelet decomposition[12], and FRIMA(Fractional ARIMA)[14], etc. Unfortunately, none of the above models that present LRD, self-similarity, and scale invariance allow an analytical solution when they are used as traffic generators feeding queueing systems, even the simplest single server queues. Even their use in simulations is often troublesome [21]. Therefore, there has been a strong demand for a model that matches the LRD characteristic of the measured traffic and yet are analytically tractable and relatively easy to understand.

MMPP(Markov Modulated Poisson Process) model proposed by Anderson[5] appears to be a promising solution for above mentioned problems. The parameters of MMPP are determined so as to match the autocorrelation function. It emulates LRD over a certain range of time scales with finite Markovian models. It was shown that the long-term correlation of traffic beyond a certain threshold does not influence the performance of a system[15],[16]. So, MMPP

* 서울산업대학교 산업정보시스템공학과 (kwlee@snut.ac.kr), ** 대한항공 정보시스템실 (byungckwon@koreanair.com)
논문번호 : KICS2007-05-211, 접수일자 : 2007년 5월 8일, 최종논문접수일자 : 2007년 11월 13일

model where correlation is limited can be successfully employed. The use of MMPP also benefits from the existence of several tools for calculating the queueing behavior. Anderson's MMPP model is used as a traffic model for IEEE 802.16 BWA (Broadband Wireless Access) system[1]. Depending on the matched statistics of the traffic data (such as autocorrelation, variance, and probability distribution), several MMPP models were proposed [6],[7],[20],[21].

Dan et al.[22] show that the burstiness of the IP packet traffic can be influenced by the burstiness of the packet arrival process and also by the packet size distribution. Salvador et al.[23] also argue that accurate modeling of IP traffic requires matching closely not only the packet arrival process but also the packet size distribution. However, MMPP models try to match the LRD characteristic with only the packet arrival process and do not characterize the packet size distribution.

When we generate packet traffic for the performance evaluation of IP packet network, MMPP model can be used to generate packet interarrival time. And a random length of packet size from a certain distribution can be assigned to each packet. However, there is a possibility that the variable packet size might change the LRD characteristic of the original MMPP model. In this study, we investigate this possibility. For this purpose the 'refined traffic' is defined, where packet arrival time is generated according to the MMPP model and a random packet length from a specific distribution is assigned to each generated packet. Hurst parameter of the refined traffic is estimated and compared with the original Hurst parameter, which is the input parameter of the MMPP model. We also investigate the effect of the packet size distribution on the queueing performance of the MMPP traffic model and the relationship between the Hurst parameter and queueing performance.

Following the introduction, we discuss the LRD of the Internet traffic and the effect of the LRD on network performance in section 2. The MMPP traffic model is also briefly explained. In section 3, the MMPP traffic model is constructed, which can be used to generate packet for this study. And the refined traffic is generated considering packet size distribution. Time series data are defined to estimate Hurst parameter for the refined traffic. In section 4, Hurst parameter estimation method is briefly discussed and estimation results are given. Queueing performances of the MMPP traffic model are also discussed. Final conclusion and future works are mentioned in section 5.

Ⅱ. Background

2.1 Self-Similarity and Long-Range Dependence

Long-range dependence has been found in various kinds of network traffic, such as in local and wide area networks [4],[8],[9],[10]. In order to model such kind of traffic, self-similar processes are introduced instead of traditional Poisson-based models.

The self-similarity and LRD definitions are given as follows. Let $X=(X_t : t=0,1,2,...)$ be a covariance stationary stochastic process with mean μ, variance σ^2, and autocorrelation function $r(k), k\geq 0$. Assume $r(k)$ is of the form

$$r(k) \sim k^{-\beta},\ as\, k\rightarrow\infty \quad (1)$$

where $0<\beta<1$.

For each $m=1,2,3,...,$ let $X^{(m)}=(X_t^{(m)}: t=1,2,3,...)$ denote the new covariance stationary time series obtained by averaging the original series X over non-overlapping blocks of size m, i.e.,

$$X_t^{(m)}=(X_{tm-m+1}+\cdots+X_{tm})/m,\ t\geq 1 \quad (2)$$

The process X is called (exactly) second-order self-similar if for all $m=1,2,3,...,var(X^{(m)})=\sigma^2 m^{-\beta}$ and

$$r^{(m)}(k) \sim r(k),\ k \geq 0 \tag{3}$$

The process X is called (asymptotically) second-order self-similar if for k large enough,

$$r^{(m)}(k) \rightarrow r(k),\ \ as\ \ m \rightarrow \infty \tag{4}$$

The key property of this class of self-similar process is that the covariance does not change under block aggregation and time scal changes. The Hurst parameter is defined as H = 1- β/2 . Note that here 0.5 < H <1, since 0 < β < 1. A self-similar process with 0.5 < H < 1 is long range dependent. Processes with LRD are characterized by an autocorrelation function that decays hyperbolically (as compared to the exponential decay exhibited by traditional traffic models). Hyperbolic decay is much slower than exponential decay, and since β < 1, the sum of the autocorrelation values of such series approaches infinity. The Hurst parameter is thus a key indicator of LRD behavior. The higher the value of H is, the stronger the long range dependence.

There are several studies, which show the effect of LRD on queueing and network performance[2],[3]. They showed that the packet loss and queueing delay behavior in simulation using real traffic data were very different from those of traditional network models.

2.2 MMPP Traffic Model

MMPP is a doubly stochastic Poisson process. Consider L MMPP's. In the case of m-state MMPP$_i$ $(i=1, 2, \ldots, L)$ its arrival rate is determined by the state of a continuous-time Markov chain with infinitesimal generator Q_i and Poisson arrival rates $\lambda_{ij}(j=1, 2, \ldots, m)$. That is, arrival rate is λ_{ij} when the Markov chain of the i^{th} MMPP is in state j. Matrix Λ_i which describes Poisson arrival rates is called the arrival rate matrix. In the two-state case, called SPP (Switched Poisson Process), Q_i and Λ_i are given by

$$Q_i = \begin{bmatrix} -C_{i1} & C_{i1} \\ C_{i2} & -C_{i2} \end{bmatrix},\ \Lambda_i = \begin{bmatrix} \lambda_{i1} & 0 \\ 0 & \lambda_{i2} \end{bmatrix}$$

Superposition of MMPPs with two-state (SPP) is suggested as a very versatile tool for the modeling of variable packet traffic with LRD[5]. The volume of traffic modeled by each of the individual two-state sources can be associated with the volume of the traffic showing variability on a given particular time scale. This should be reflected in the choice of time constants for each individual SPP, i.e., the choice of modulating parameters C_{i1} and C_{i2} . For example, consider a model consisting of a three SPP's. If we assume $\lambda_{i2}=0$, this model can be considered as a superposition of 3 IPP's (Interrupted Poisson Process).

1) First IPP:
$\lambda_1 = (\lambda_{11}) = 6.0,\ C_{11} = C_{12} = 10^{-2}$

2) Second IPP:
$\lambda_2 = (\lambda_{21}) = 6.0,\ C_{21} = C_{22} = 10^{-4}$

3) Third IPP:
$\lambda_3 = (\lambda_{31}) = 6.0,\ C_{31} = C_{32} = 10^{-6}$

In Fig. 1, autocorrelation functions of three IPP's are shown as a function of time scale. From this, we can see that only IPP 3 contributes significantly to the correlation for $k=10^6$, while

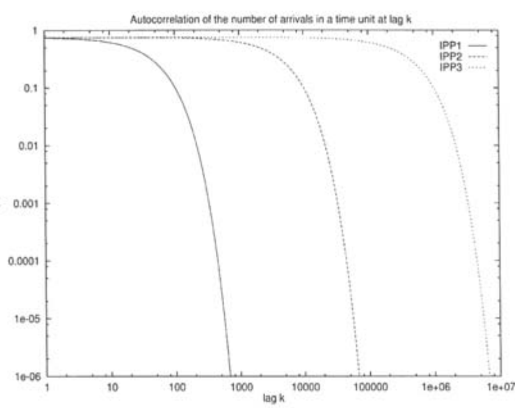

Fig. 1. Autocorrelation Functions of IPP's

Table 1. Required Input Parameters

Parameter	Meaning
λ	Packet arrival rate of process
d	Number of IPP's
ρ	Lag 1 correlation
k	Time scale
H	Hurst parameter

both IPP 2 and IPP 3 contribute significantly to the correlation for $k=10^{4}$, and finally for $k=10^{2}$ all three IPP's contribute significantly to the correlation.

Depending on the statistics (such as autocorrelation and variance) of the traffic data matched using the MMPP model there can be several methods to construct the MMPP traffic model[6],[7],[20],[21]. The required input parameters are shown in Table 1.

Ⅲ. Generation of the Refined Traffic and Time Series

3.1 Construction of the MMPP Model

In this study Anderson's model is used to generate MMPP traffic, where traffic is modeled by the superposition of several IPP's. The parameters of IPP's are determined so as to match the autocorrelation function. For this study, packet arrival rate, number of IPP's, time scale, and lag 1 correlation are assumed as 3/unit time, 4, 10, and 0.5, respectively. For the Hurst parameter we use three different values: 0.6, 0.75, and 0.85. Following the flow diagram in Anderson's work[5], three MMPP models are constructed as shown in Table 2.

Table 2. MMPP Models

Input parameters	source	λ_i^{IPP}	c_{1i}	c_{2i}
d = 4 k = 10^5 λ = 3.0 H = 0.60 ρ = 0.50	IPP_1	8.253	$5.714*10^{-1}$	$2.286*10^{-1}$
	IPP_2	1.799	$1.231*10^{-2}$	$4.924*10^{-3}$
	IPP_3	0.381	$2.652*10^{-4}$	$1.061*10^{-4}$
	IPP_4	0.088	$5.714*10^{-6}$	$2.286*10^{-6}$
d = 4 k = 10^5 λ = 3.0 H = 0.75 ρ = 0.50	IPP_1	6.579	$5.715*10^{-1}$	$2.285*10^{-1}$
	IPP_2	2.562	$1.231*10^{-2}$	$4.923*10^{-3}$
	IPP_3	0.914	$2.653*10^{-4}$	$1.061*10^{-4}$
	IPP_4	0.448	$5.715*10^{-6}$	$2.285*10^{-6}$
d = 4 k = 10^5 λ = 3.0 H = 0.85 ρ = 0.50	IPP_1	4.748	$5.715*10^{-1}$	$2.285*10^{-1}$
	IPP_2	3.383	$1.231*10^{-2}$	$4.923*10^{-3}$
	IPP_3	0.983	$2.653*10^{-4}$	$1.061*10^{-4}$
	IPP_4	1.837	$5.715*10^{-6}$	$2.285*10^{-6}$

To generate packet traffic from above MMPP models, ARENA simulation is used.

3.2 Packet Size Distribution and Refined Traffic Generation

The burstiness of the IP packet traffic can be influenced by the burstiness of the packet arrival process and also by the packet size distribution[22]. Therefore, accurate modeling of IP traffic requires matching closely not only the packet arrival process but also the packet size distribution. The MMPP models, which only address the packet arrival process, can not characterize the packet size distribution.

In this study we want to see the effect of packet size distribution on LRD characteristic of the MMPP traffic. For this purpose the MMPP traffic model is refined, where packet arrival time is generated according to the MMPP traffic model and a random packet size from a certain distribution is assigned to each packet. Following three distributions are assumed. For convenience'sake, mean packet length is assumed as 1.

- constant
- exponential
- pareto (α =2, β =0.5)

Now the refined traffic is generated using the MMPP model of Table 2 and one of the packet

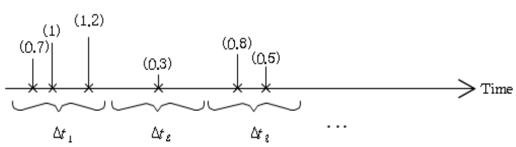

Fig. 2. Refined Packet Traffic

size distributions mentioned above. To investigate the effect of variable packet size on LRD characteristic of the MMPP models, following two Hurst parameters are compared.

1. Hurst parameter H, which is used as a input parameter for MMPP model
2. Estimated Hurst parameter H^* of the refined traffic.

3.3 Time Series Generation

The refined traffic trace is characterized by two variables: packet arrival time and packet length. From this trace, time series with only one variable must be generated in order to estimate Hurst parameter for this variable.

During Δt time period we have random number of packet arrivals. Within non-overlapping time intervals of size Δt we sum the size for the packets arriving in each interval Δt_i, and obtain the time series {Traffic X } = $\{X_i, i=1,2,3,...\}$. Following Fig. 2 shows the refined traffic obtained from the MMPP model by assigning random packet length. The arrival times are marked as X and the values in parenthesis represent the packet sizes.

From the above Figure 2 time series are determined as $X_1 = 2.9 = (0.7+1+1.2)$, $X_2 = 0.3$, $X_3 = 1.3 = (0.8+0.5)$, ⋯.

Ⅳ. Hurst Parameter Estimation

To obtain the time series from the refined traffic we used three different values of Δt : 0.15, 0.3, and 0.6. Assuming that mean packet arrival rate is 3, the mean numbers of packet arrival during this interval are 0.45, 0.9, and 1.8, respectively.

4.1 Hurst Parameter Estimation

In this study the aggregated variance method[19] is used to estimate Hurst parameter H^*. Time series data are divided in blocks of length m and compute the sample average and variance.

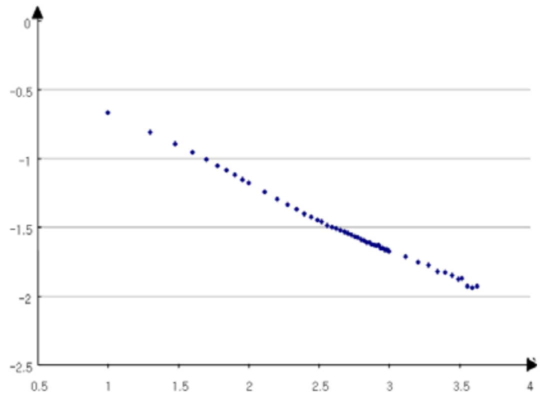

Fig. 3. Log(m) versus Log(Sample Variance)

And we construct the graph of Log(m) versus Log(sample variance). From the slope of this graph Hurst parameter is estimated. For sufficiently large values of m the slope estimates 2H-2. The graph slope can be obtained using the least square method.

Following Figure 3 shows the graph of Log (m) versus Log(sample variance) for the exemplary refined traffic. To generate the refined traffic, the MMPP model of H = 0.75 (Table 2) and exponential packet size distribution with mean 1 are assumed. Δt = 0.3 is used to generate time series.

Using the least square method the graph slope is determined as -0.49, and the Hurst parameter is estimated as 0.755.

Table 3. Hurst Parameter Estimation (H*)

Hurst parameter of the MMPP model	Time interval	Constant	Exponential	Pareto
H=0.6	0.15 0.3 0.6	0.603 0.620 0.619	0.613 0.618 0.613	0.610 0.615 0.620
H=0.75	0.15 0.3 0.6	0.750 0.750 0.753	0.758 0.755 0.753	0.753 0.758 0.756
H=0.85	0.15 0.3 0.6	0.851 0.852 0.853	0.855 0.850 0.854	0.856 0.851 0.855

From above Table 3, we can find following facts:

- The estimated Hurst parameter value of the MMPP traffic is almost same as the Hurst parameter value H, which is the input parameter of the MMPP model. This fact shows that MMPP faithfully matches the LRD characteristic of the measured traffic.
- The estimated Hurst parameter values, H^*, remain almost same irrespective of the packet size distributions and time interval Δt, which shows that the LRD characteristic of the MMPP traffic model is not affected by the packet size distributions.

4.3 Queueing Performance

In this section, the effect of the packet size distribution on the queueing performance of the MMPP traffic model is investigated. Also, the relationship between the Hurst parameter and queueing performance is addressed. For this purpose we consider the single-server model, infinite queue size, and FIFO discipline. The MMPP models of Table 2 are used as the offered traffic. The server loads are assumed as 0.4, 0.5, and 0.6. Following Table 4 summarize the simulation results.

Table 4. Simulation Results of the Queueing Performane (* denotes server load)

Hurst parameter of the MMPP model	Packet size Distribution	Waiting time in queue			Number of packet in queue		
		0.4 *	0.5 *	0.6 *	0.4 *	0.5 *	0.6 *
H=0.6	Constant	0.590	1.393	3.014	1.775	4.188	9.084
	Exponential	0.718	1.567	3.292	2.158	4.713	9.924
	Pareto	0.917	1.903	3.634	2.759	5.720	10.954
H=0.75	Constant	0.417	1.226	4.602	1.257	3.631	13.955
	Exponential	0.551	1.466	5.044	1.663	4.342	15.296
	Pareto	0.725	1.914	5.932	2.186	5.667	17.990
H=0.85	Constant	0.598	5.265	52.225	1.887	16.898	176.38
	Exponential	0.804	5.663	53.161	2.539	18.178	179.54
	Pareto	1.112	6.719	54.320	3.509	21.565	183.46

From above Table 4 we can find following facts:

- The queueing performances of the MMPP model are influenced by the packet size distribution.
- It is anticipated that the traffic with higher Hurst parameter shows more LRD traffic burstiness, and thus leads to much larger queue size. This is true for the relatively high server load, 0.6. However, the MMPP traffic model of H = 0.6 gives slightly larger queue size and waiting time than the MMPP traffic model of H =0.75 when the server load is 0.4 or 0.5. This shows that the Hurst parameter is not by itself an accurate predictor of the queueing performance for a given LRD traffic trace. (This result is also observed by Gerla[18])

V. Conclusion and Future Works

The MMPP traffic model is refined by assigning a random packet size to each generated MMPP packet. To estimate Hurst parameter time series are constructed by summing the sizes for the packets arriving during time interval Δt.

It is shown that Hurst parameter of the refined traffic is almost same as that of the original MMPP traffic. That is, the LRD characteristic of the MMPP traffic remains same under the variable packet size following a certain distribution.

It is also shown that the queueing performances of the MMPP traffic model are influenced by the packet size distributions. We show a case, where the queue size and delay of the traffic with low Hurst parameter are larger than those of the traffic with relatively high Hurst parameter and conclude that the Hurst parameter is not by itself an accurate predictor of the queueing performance for a given LRD traffic trace. In this study we have seen that the Hurst parameters of the traffic traces with different packet size distributions are same. However, that does not mean their queueing performances are same. Further study is required to analytically show the relationships between the Hurst parameter and queueing performance.

Referrnces

[1] IEEE 802.16 Broadband Wireless Access System, IEEE 802.16.3c-01/30r1, March 2001.

[2] A. Erramilli, O. Narayan, W. Willinger,

"Experimental queueing analysis with long-range dependent packet traffic", Bellcore.

[3] K. Park, G. Kim, M. Crovella, "On the effect of traffic self-similarity on network performance", *SPIE*, 1997.

[4] M. Crovella, A. Bestavros, "Explaining world wide web traffic self-similarity", Boston University, October 1995.

[5] A. Anderson, B. Nielsen, "A Markovian approach for modeling packet traffic with long-range dependence", *IEEE JSAC*, vol.16, no.5, pp.719-732, June 1998.

[6] T. Yoshihara, S. Kasahara, Y. Takahasi, "Practical time-scale fitting of self-similar traffic with Markov-modulated Poisson process", *Telecommunication Systems,* vol.17, no.1-2, pp.185-211, 2001.

[7] P. Salvador, A. Pacheco, R. Vaiadas, "Multiscale fitting procedure using Markov modulated Poisson process", *Telecommunication Systems,* vol.23, no.1-2, pp.123-148, June 2003.

[8] W. Leland, M. Taqqu, W. Willinger, D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)", *IEEE/ACM Transaction on Networking,* vol.2, no.1, pp.1-15, Feb 1994.

[9] V.Paxon, S.Floyd, "Wide-area traffic: the failure of Poisson modeling", *IEEE/ACM Transactions on networking,* vol.3, no.3, pp.226-244, 1995.

[10] W. Willinger, M.S. Taqqu, R. Sherman, D.V. Wilson, "Self-similarity through high variability: statistical analysis of ethernet LAN traffic at the source level", *IEEE/ACM Transaction on Networking,* vol.5, no.1, pp.79-86, 1997.

[11] I. Norros, "On the use of fractional Brownian motion in the theory of connectionless networks", *IEEE JSAC*, vol. 13, pp.953-962, 1995.

[12] S. Ma, C. Ji, "Modeling heterogeneous network traffic in wavelet domain", *IEEE/ ACM Transaction on Networking,* vol.9, no.5, pp.634-649, 2001.

[13] R.J. Mondragon, J.M. Pitts, D.K. Arrowsmith, "Chaotic intermittency-sawtooth map model of aggregated self-similar traffic stream", *Electronic Letters,* vol.36, no.2, pp.184-186, 2000.

[14] M. Krunz, A. Makowski, "A source model for VBR video traffic based on M/G/infinity input processes", *IEEE Infocom '98,* SF, CA, USA, pp.1441-1448, April 1998.

[15] M. Grossglauser, J.C. Bolot, "On the relevance of long-range dependence in network traffic", *IEEE/ACM Transactions on Networking,* vol.7, no.5, pp.629-640, Oct 1999.

[16] A. Nogueria, R. Valadas, "Analyzing the relevance time scale in a network of queues", *SPIE*, August 2001.

[17] H. Riedi, D.V. Vehel, "A multifractal wavelet model with application to data traffic," *IEEE Transaction on Info. Theory*, vol.45, no.4, pp.992-1018, 1999.

[18] R. Ritke, X. Hong, M. Gerla, "Contradictory relationship between Hurst atameter and queueing performance', *Computer Science Department,* UCLA, 2004.

[19] M. Trovero, "Long-range dependence: a light tale for practitioner", Oct 2003.

[20] A. Nogueira, P. Salvador, R. Valadas, "Fitting self-similar traffic by a superposition of MMPPs modeling the distribution at multiple time scales", *IEICE Trans. on Communications,* vol.E27-B, no.3, pp.678-688, March 2004.

[21] L. Muscariello, M. Mellia, M. Meo , A. M. Marsan, R.M. Cigno, "Markov models of internet traffic and a new hierarchical MMPP model", *Computer Communication*, vol.28, pp.1835-1851, 2005.

[22] G. Dan, V. Fodor, G. Karlsson, "Packet size distribution: an aside", QoS-IP, pp.75-87, 2005.

[23] P. Salvador, A. Pacheco, R. Vaiadas, "Modeling IP traffic: joint characterization of packet arrivals and packet sizes using BMAPs", *Computer Networks,* vol.44, pp.335-352, 2004.

이 강 원 (Kang-won Lee) ??회원

1980 서울 대학교 공과대학 공학사(산업공학)
1982 서울 대학교 대학원 공학석사(산업공학)
1985 Kansas State Univ. U.S.A 공학박사(산업공학)
1989 한국전자통신연구원 TDX 개발단
2007 서울산업대학교 산업정보시스템공학과
<관심분야> 정보통신, 품질 및 신뢰성, O.R.

권 병 천 (Byung-chun Kwon) ??회원

2006 서울 산업대학교 산업정보시스템공학사(산업공학)
2007 서울 산업대학교 산업대학원 정보산업공학석사(산업공학)
<관심분야> O.R., ERP