Statistics and Applications {ISSN 2454-7395 (online)}
Volume 20, No. 1, 2022 (New series), pp 297-309

Fitting Model for Self-Similar Traffic - Time Dependent Markovian Process and Second Order Statistics

Abhilash Vollala and Malla Reddy Perati
Department of Mathematics, Kakatiya University, Warangal, India.

Received: 18 March 2021; Revised: 28 April 2021; Accepted: 20 May 2021

Abstract

Various self-similar traffic models proposed earlier are asymptotic. In this connection, many Markovian models are proposed, but the performance analysis was possible only when the system is in steady state, and such traffic models are not realistic. In this paper, a procedure is proposed to fit Markov Modulated Poisson Process (MMPP) with time dependent sinusoidal arrival rates that emulates self-similar traffic. This is done by matching the variance of the both during prescribed time scales. Numerical results represent what extent MMPP could reproduce self-similar traffic in specified time scales.

Key words: Self-similar traffic; Variance-time; Markovain Modulated Poisson process (MMPP); Interrupted Poisson process (IPP).

1. Introduction

The pivotal studies at AT&T Bell labs (Leland, *et al.* 1994; Paxson and Flyod, 1995) revealed that IP packet traffic over LAN and WWW internet traffic (Crovella and Bestavros, 1997) are self-similar (fractal like behavior), and this behavior effects efficiency of network nodes such as routers or switches (Misra, *et al.* 2012). Many traffic models such as Chaotic maps, FARIMA, and FBM are proposed to emulate the self-similar behavior. These models are parsimonious, but are asymptotic, hence they are not practically useful (Erramilli, *et al.* 1996; Norros, 1994). Andersen and Nielsen (1998) used Markovian Arrival Process (MAP) in particular Switched Poisson Process (SPP) to model the Long Range Dependent (LRD) characteristics (statistical definition of self-similarity) of traffic over different time scales, and proposed a fitting procedure wherein covariance of second-order self-similar process, and that of resultant MAP (superposition of several SPPs) are equated. Kasahara, *et al.* (2001) proposed a method based on variance of second-order self-similar traffic using Interrupted Poisson Process (IPP). Later, Reddy, *et al.* (2005) extended the work by making modulating parameters of each IPP unequal, and investigated the relation between traffic parameters, time scale, and parameters of fitting. In all the above, resultant MAP is homogeneous and the arrival rates were not functions of time, and queueing behavior of traffic nodes was investigated in the steady state. Steady state outcomes never give actual queueing behavior, since it relies on a prolonged performance of system to nullify the initial conditions (Kelton and Law, 1985). In real time, network traffic is not homogeneous over all time scales, and the modeled system by no means attains steady state. Because of these reasons, necessity of time dependent (transient) analysis is warranted. Abate and Whitt (1988) analyzed *M/M/1* queue in transient with Poisson arrivals using Laplace transforms. Eick, *et al.* (1993) studied the $M_t/G/\infty$ queue with time

Corresponding Author:Abhilash Vollala
Email : ashwesa.abhi@gmail.com

dependent arrival rates, and determined the number of busy servers at time *t*. Different methods were developed by Jennings and Massey (1997) for analysis blocking in circuit switched networks with transient arrival traffic. Massey (2002) proposed canonical queueing models with time-varying rates, and derived necessary mathematical tools for analysis. Qian and Tipper (2004) proposed a framework for adaptively determining optimal channel allocation scheme, and evaluated performance of the scheme under time varying loads. Liu and Whitt (2014) developed an algorithm to find number of servers required to preserve the prerformance in a multi server queue with time varying arrival rates for extension of feed-forward netwoks. Pant and Ghimire (2016) determined the expected queue length and waiting time of the customers at time t using transient arrival rates. In the papers cited above, authors worked with various queueing systems using time based sinusoidal arrival rates to address issues in various domains. These models are based on the fact that many real time arrival phenomenons are almost periodic in nature. In this paper, procedure to fit Markovian Arrival Process (MAP) with sinusoidal arrival rates that emulates self-similar traffic over prescribed time scales is proposed. Variance of number of arrivals due to self-similar traffic, and that of resultant MAP (superposition of IPPs with sinusoidal arrival rates) are equated at certain time points in order to compute the MAP parameters.

The remaining part is arranged in the following way. In section 2, outline of second-order self-similar process and sinusoidal IPP are presented. In section 3, fitting procedure with time dependent sinusoidal arrival rates is given. Numerical results are demonstrated in section 4. Finally, some conclusions are given in section 5.

2. Second-Order Self-Similar and Sinusoidal IPP Processes

Self-similarity is a property, wherein a certain feature of the object is maintained with respect to scaling in space and time. It is statistically defined as follows. Let X be a second order process with variance σ^2, and the time axis is splitted into disjoint sub intervals of unit length. Let $X = \{X_t \ / t = 1, 2, 3, ...\}$ be the points (packet arrivals) in t^{th} interval. Let $X^{(r)} = \{X_t^{(r)}\}$ be a new sequence obtained by averaging the original sequence over non-overlapping blocks of size r . *i.e.*,

$$X_t^{(r)} = \frac{1}{r}\sum_{i=1}^{r} X_{(t-1)r+i} \ , \qquad t = 1, 2, 3, ... \tag{1}$$

The obtained sequence is a second order process, and is called exactly second order self-similar with Hurst Parameter, $H = 1 - \beta/2$, if

$$Var\ (X^{(r)}) = \sigma^2\, r^{-\beta}\ , \ \ \forall r \geq 1 \tag{2}$$

This feature can be emulated by Markov Modulated Poisson Process (MMPP) over the desired time scale (Reddy, *et al.* 2005). For the reason mentioned in the introduction, a special type of two-state IPP is proposed using time dependent arrival rates. The proposed IPP is given as follows:

$$Q = \begin{bmatrix} -d_1 & d_1 \\ d_2 & -d_2 \end{bmatrix}, \quad R(t) = \begin{bmatrix} a + b\sin(t) & 0 \\ 0 & 0 \end{bmatrix}. \tag{3}$$

In the above, Q is transition rate matrix with two distinct parameters and $R(t)$ is arrival rate matrix, which says that, when Markov process is in state 1, arrival process is with sinusoidal arrival rate $\lambda(t) = a + b\sin(t)$, where a, b are the constants, and $b = a\gamma$, $0 < \gamma < 1$ (Eick, *et al.* 1993), and when the Markov process in state 2, there are no arrivals. The number of arrivals in $(0, t]$ of the said IPP is denoted by N_t, and J_t be the state of Markov process at time *t*. The Generating function $P^*(z,t)$ obtained by forward Chapman-Kolmogorov equations with time dependent Markovian process (Fischer and Hellstern, 1993) is given by

$$P^*(z,t) = Exp\left(Qt - (1-z)\int_0^t R(t)dt\right). \tag{4}$$

The mean value of N_t is given by (Heffes and Lucantoni, 1986)

$$E[N_t] = \Pi' \left[\frac{\partial}{\partial z} P^*(z,t)\right]_{z=1} e,$$

Differentiating with respect to *z*, and solving above equation (Neuts, 1979; Coddington and Levinson, 1987), one can obtain as

$$E[N_t] = \Pi' R^{\circ}(t)\, e,$$

$$\Rightarrow E[N_t] = \frac{d_2 \int_0^t \lambda(t)\, dt}{d_1 + d_2},$$

$$\Rightarrow E[N_t] = \frac{d_2\ (at + b(1 - \cos(t)))}{d_1 + d_2}. \tag{5}$$

The variance of N_t is given by (Heffes and Lucantoni, 1986)

$$Var[N_t] = \Pi' \left[\frac{\partial^2}{\partial z^2} P^*(z,t)\right]_{z=1} e + E[N_t] - \{E[N_t]\}^2$$

After differentiating with respect to *z*, and by applying some algebraic manipulation (Neuts, 1979; Coddington and Levinson, 1987), one can obtain as

$$Var[N_t] = E[N_t] + \frac{2}{t}[(E[N_t])^2 - \Pi' R^{\circ}(t)(Q + e\Pi')^{-1} R^{\circ}(t)e]$$
$$+ \frac{2}{t^2}[\Pi' R^{\circ}(t)(e^{Qt} - I) R^{\circ}(t)e],$$

$$\Rightarrow Var[N_t] = \frac{d_2 \int_0^t \lambda(t)\, dt}{d_1 + d_2} + \frac{2 d_1 d_2 \left(\int_0^t \lambda(t)\, dt\right)^2}{t\ (d_1 + d_2)^3} - \frac{2 d_1 d_2 \left(\int_0^t \lambda(t)\, dt\right)^2}{t^2\ \ (d_1 + d_2)^4}[1 - e^{-(d_1 + d_2)t}],$$

$$\Rightarrow Var[N_t] = \frac{d_2\ (at + b(1-\cos(t)))}{d_1 + d_2} + \frac{2d_1 d_2\ (at + b(1-\cos(t)))^2}{t\ (d_1 + d_2)^3} - \frac{2d_1 d_2\ (at + b(1-\cos(t)))^2}{t^2\ (d_1 + d_2)^4}[1 - e^{-(d_1+d_2)t}]. \tag{6}$$

where Π' is steady state vector of Markov chain, e is the vector of appropriate dimension with each entry as 1, and $R^{\circ}(t) = \int_0^t R(t)dt.$

The Index of Dispersion for Counts (IDC) is

$$IDC(t) = \frac{Var[N_t]}{E[N_t]}$$

By using Eqs. (4) and (5), one can have

$$IDC(t) = 1 + \frac{2d_1\left(\int_0^t \lambda(t)\,dt\right)}{t\ (d_1 + d_2)^2} - \frac{2d_1\left(\int_0^t \lambda(t)\,dt\right)}{t^2\ (d_1 + d_2)^3}[1 - e^{-(d_1+d_2)t}]. \tag{7}$$

One can derive the following results:

i. $IDC(t) \to 1,$ as $t \to 0$ (i.e., MMPP move towards to a Poisson process).

ii. $IDC(t) \to 1 + \frac{2d_1 a}{(d_1 + d_2)^2}$, a constant, as $t \to \infty$.

3. Fitting Procedure

Variance based fitting procedure is to obtain the MMPP parameters using time dependent arrival rates. From earlier, (Reddy, *et al.* 2005; Kasahara, *et al.* 2001) it is known that modelling of self-similar traffic involves superposition of number of two-state MMPPs (in particular IPPs), and the fundamental requirements for fitting process are

i) $[r_{\min}, r_{\max}]$: Minimum, maximum limits of the time scale range.

ii) $\lambda_w(t)$: Arrival rate of whole process at time t.

iii) n : Number of superposed two-state IPPs.

iv) H : Hurst parameter.

v) σ^2 : Variance.

The j^{th} IPP of the process is given by

$$Q_j = \begin{bmatrix} -d_{1j} & d_{1j} \\ d_{2j} & -d_{2j} \end{bmatrix}, \quad R_j(t) = \begin{bmatrix} a + b_j \sin(t) & 0 \\ 0 & 0 \end{bmatrix}, \quad 1 \le j \le n \tag{8}$$

where a, b_j are the constants and $b_j = a\gamma_j$, $0 < \gamma_j < 1$. In this context, a is assumed that, it is equal to the whole arrival rate. The superposition of n IPPs, and a Poisson process is stated as

$$Q = Q_1 \oplus Q_2 \oplus \cdots \oplus Q_n,$$

$$R(t) = R_1(t) \oplus R_2(t) \oplus R_3(t) \cdots \oplus R_n(t) \oplus \lambda_p(t). \tag{9}$$

Here, time dependent arrival rate of classical Poisson Process is denoted by $\lambda_p(t)$, and $\oplus$ indicates the kronecker's sum, where the resultant of above sum is also an MMPP. The arrival rate of whole process at time t is given by

$$\lambda_w(t) = \lambda_p(t) + \sum_{j=1}^{n} \frac{d_{2j}}{d_{1j} + d_{2j}} \lambda_{1j}(t). \tag{10}$$

Let Y_t be the number of arrivals in whole Markovian process, and the arrivals from j^{th} IPP and Poisson process are denoted by $N_{<j,t>}$, $N_{<p,t>}$ respectively, during t^{th} time slot. The count of arrivals after the averaging process are $Y_t^{(r)}$, $N_{<j,t>}^{(r)}$ and $N_{<p,t>}^{(r)}$ respectively.

Therefore, $Y_t^{(r)} = \sum_{i=1}^{n} N_{<j,t>}^{(r)} + N_{<p,t>}^{(r)}$. (11)

The variance of j^{th} IPP and Poisson process are given as

$$Var\,[N_{<j,t>}^{(r)}] = \frac{d_{2j}\left(\int_o^r \lambda_{1j}(t)\,dt\right)}{r^2(d_{1j}+d_{2j})} + \frac{2d_{1j}d_{2j}\left(\int_o^r \lambda_{1j}(t)\,dt\right)^2}{r^3(d_{1j}+d_{2j})^3} - \frac{2d_{1j}d_{2j}\left(\int_o^r \lambda_{1j}(t)\,dt\right)^2}{r^4(d_{1j}+d_{2j})^4}[1 - e^{-r(d_{1j}+d_{2j})}], \tag{12}$$

$$Var\left[N_{<p,t>}^{(r)}\right] = \frac{\int_0^r \lambda_p(t)\,dt}{r^2}. \tag{13}$$

From Eqs. (11), (12), (13), and using the fact that variance of a resultant process is preserved by the superposition of distinct sub-processes, the following relation is obtained.

$$Var\left[Y_t^{(r)}\right] = \frac{\lambda_w(t)}{r} + \sum_{i=1}^{n} \xi_j \left(\int_0^r \lambda_{1j}(t)\,dt\right)^2, \tag{14}$$

where $\xi_j = \dfrac{2d_{1j}d_{2j}}{r^3(d_{1j}+d_{2j})^3} - \dfrac{2d_{1j}d_{2j}}{r^4(d_{1j}+d_{2j})^4}[1 - e^{-r(d_{1j}+d_{2j})}]$. (15)

Using (1) and (14), at n different points r_j, $j = 1, 2, 3, .., n$ variance of both processes are equated. The time scale over which self-similarity of traffic exhibits is taken as $[r_{\min}, r_{\max}]$ (*i.e.*, $r_{\min} \le r \le r_{\max}$), then r_j is given by

$$r_j = r_{\min}\alpha^{j-1}, \quad j = 1,2,\ldots,n. \tag{16}$$

where $\alpha = \left[\dfrac{r_{\max}}{r_{\min}}\right]^{\frac{1}{n-1}}, \quad n > 1.$ (17)

Consider that $r\,d_{1j} = l1,\ r\,d_{2j} = l2$ and make use of inequality given below in Eq. (15),

$$1-(l1+l2) < e^{-(l1+l2)} < 1-(l1+l2)+\frac{(l1+l2)^2}{2!},\ \ \forall\ l1,\ l2>0,\ l1+l2 \ll 1,$$

One can find that the ξ_j lies in between $(0,\ 1), \forall j$. Since ξ_j is bounded $\forall j$, $Var[Y_t^{(r)}]$ is bounded for $t > 0$. By virtue of self-similarity, the following relations can be adopted: $r_j\,d_{1j} = d_1$ (a constant), $r_j\,d_{2j} = d_2$ (a constant), $j = 1,\ 2,\ \ldots,n,$ where d_1 and d_2 are independent,

$$i.e.,\ d_{1j} = \frac{r_1}{r_i} d_{11}\ ,\ \ d_{2j} = \frac{r_1}{r_i} d_{21}. \tag{18}$$

The above assumptions are based on the fact that self-similar process is identical for all time scales. Now the parameters to be obtained are d_{11}, d_{21}. Once their values are obtained, then the values of d_{1j} and d_{2j} can be generated using (18). By using these values, $\lambda_p(t)$ can be obtained from (10). Finally, the required transition parameters $d_{11},\ d_{21}$ can be determined, such that the value of the integral $\int_{r_{\min}}^{r_{\max}} \left[Var(X^{(r)}) - Var(Y_t^{(r)})\right] dr$ is minimum.

4. Numerical Results

Accuracy of fitting (self similar traffic as of MMPP) is presented using different samples given in Table 1. The samples are pertaining to seminal studies at AT& T Bell labs (Leland, *et al.* 1994). The number of superposed IPPs n is taken to be 4. The sinusoidal arrival rates are given in Table 2. The variance versus time curves of resultant MMPPs and self-similar traffic are shown in Figs.1-12 for $n = 4$ in the time scale ranges $[10, 10^4],\ [10^2, 10^5],\ [10^2, 10^6],\ [10^2, 10^7]$. The n is taken to be 3 in typical time scale range $[10^2, 10^5]$ to represent the effect of the number of superposed components, and the pertinent results are presented in Figs.13-15. The results exhibit good agreement with that of self-similar traffic.

Table 1: Fitting data of samples in time scale range [10^2, 10^5]

Sample Number	Parameter Values	*n*=4		*n*=3	
		d11	***d21***	***d11***	***d21***
Sample 1	$H = 0.7,\ \lambda_w(t) = 1$, and $\sigma^2 = 0.6$	0.22	0.001	0.09	0.0924
Sample 2	$H = 0.8,\ \lambda_w(t) = 1$, and $\sigma^2 = 0.6$	0.5	0.0359	0.385	0.025
Sample 3	$H = 0.9,\ \lambda_w(t) = 1$, and $\sigma^2 = 0.6$	0.040015	0.005	0.202	0.001

Source: Leland, et al. 1994 *(for samples)*

Table 2: Sinusoidal arrival rates

Arrival rate	Value
$\lambda_{11}(t)$	$1+0.1\times\sin(t)$
$\lambda_{12}(t)$	$1+0.3\times\sin(t)$
$\lambda_{13}(t)$	$1+0.5\times\sin(t)$
$\lambda_{14}(t)$	$1+0.7\times\sin(t)$

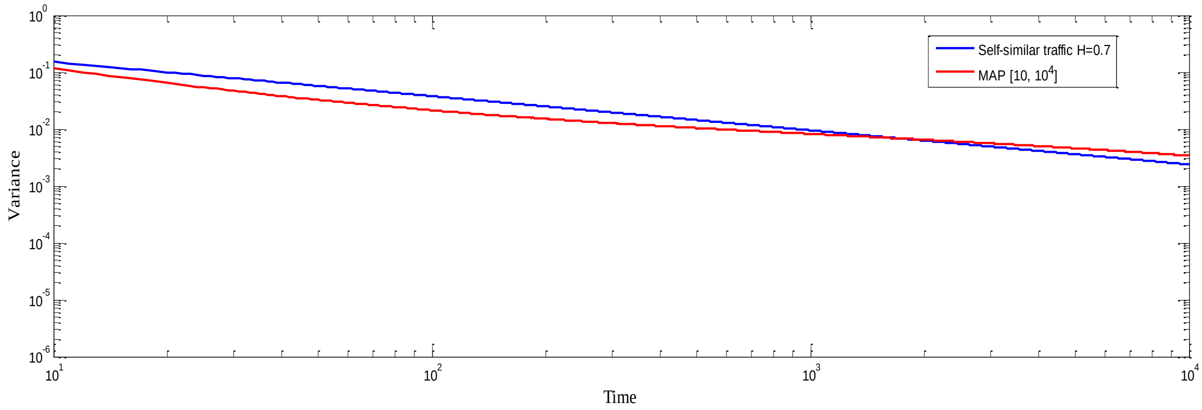

Figure 1: Variance versus time curves with $n=4$ **for sample 1during time scale** $[10, 10^4]$

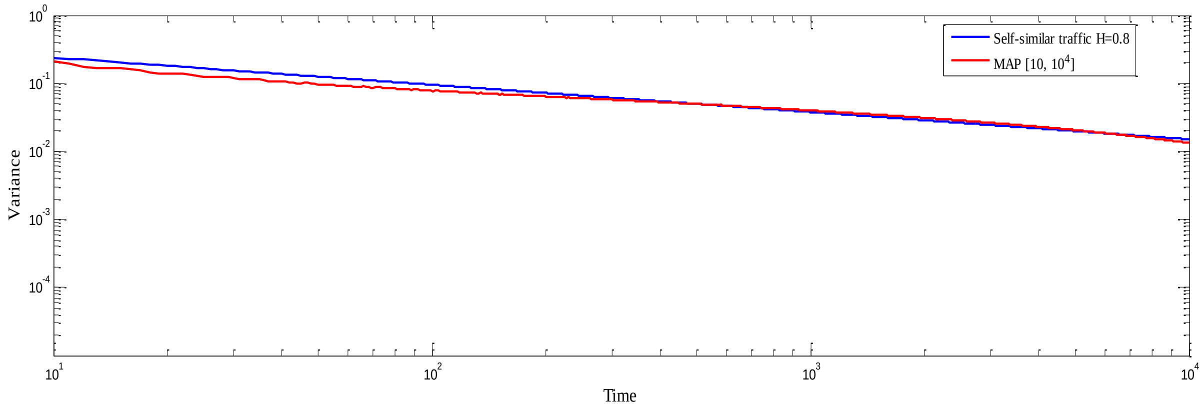

Figure 2: Variance versus time curves with $n=4$ **for sample2 during time scale** $[10, 10^4]$

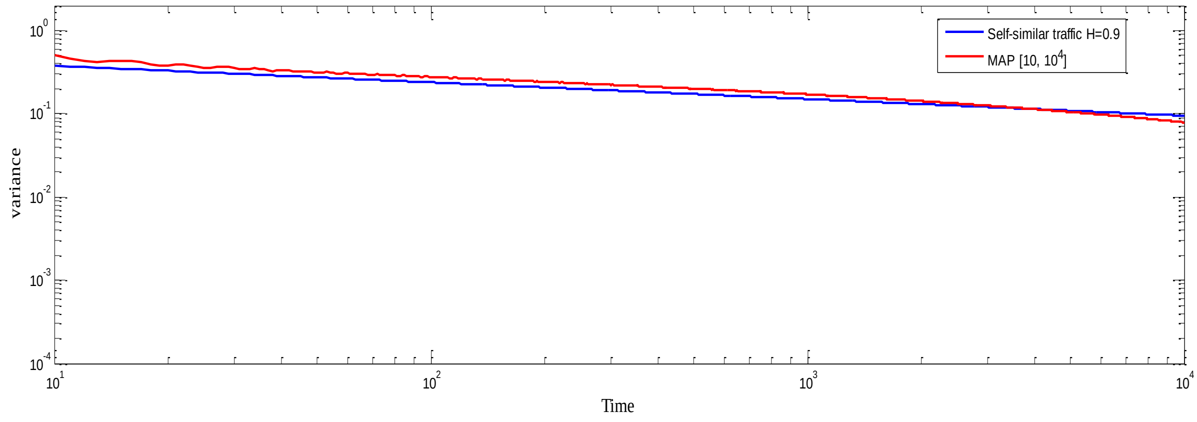

Figure 3: Variance versus time curves with $n = 4$ **for sample3 during time scale** $[10, 10^4]$

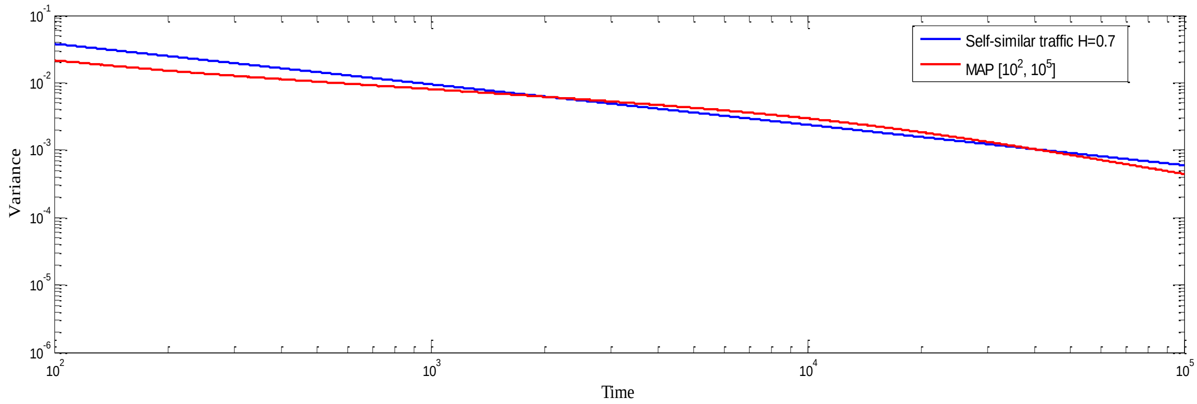

Figure 4: Variance versus time curves with $n = 4$ **for sample1 during time scale** $[10^2, 10^5]$

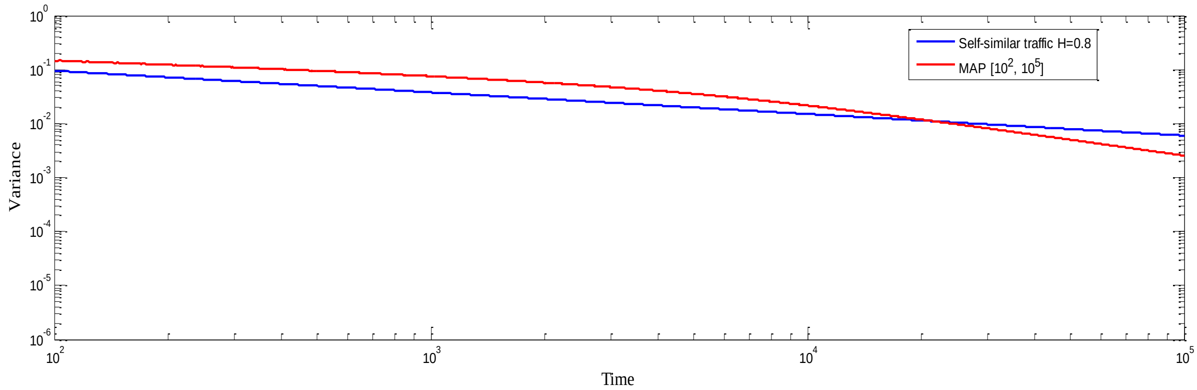

Figure 5: Variance versus time curves with $n = 4$ **for sample2 during time scale** $[10^2, 10^5]$

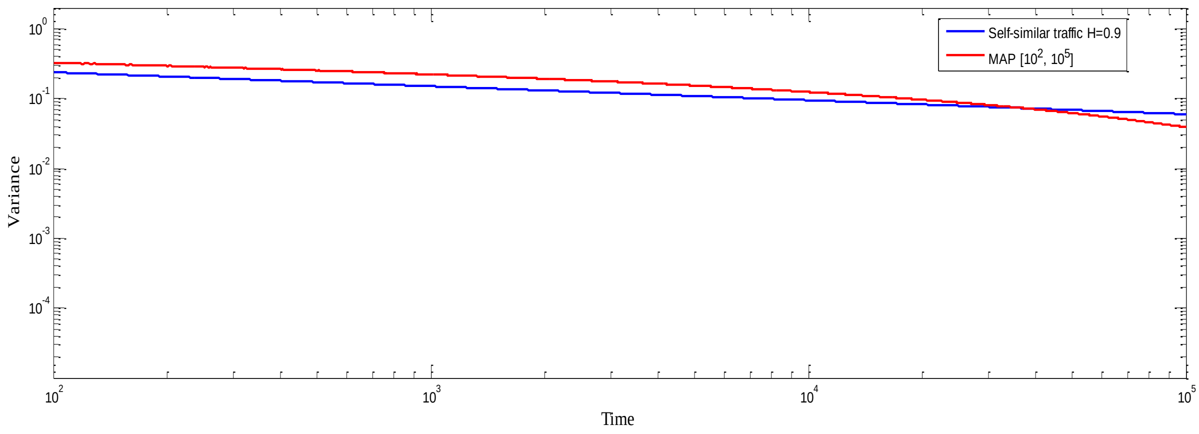

Figure 6: Variance versus time curves with $n = 4$ **for sample3 during time scale** $[10^2, 10^5]$

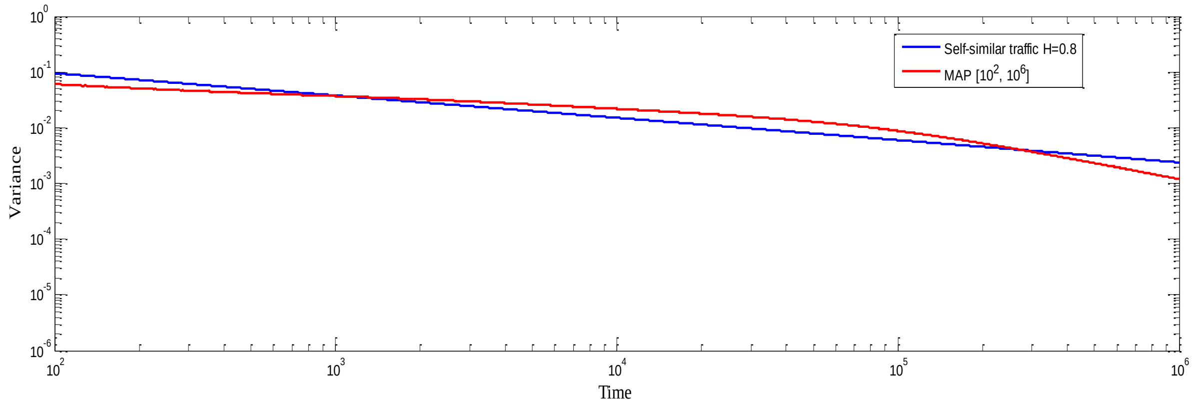

Figure 7: Variance versus time curves with $n = 4$ **for sample1 during time scale** $[10^2, 10^6]$

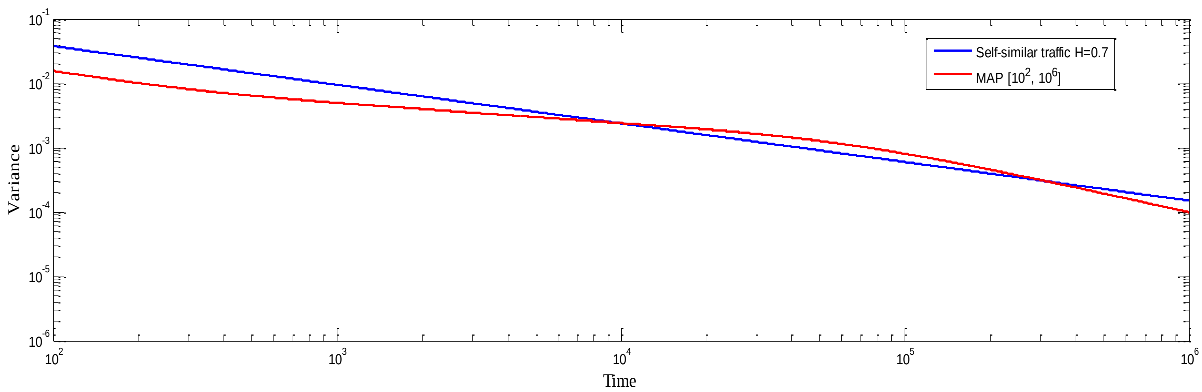

Figure 8: Variance versus time curves with $n = 4$ **for sample2 during time scale** $[10^2, 10^6]$

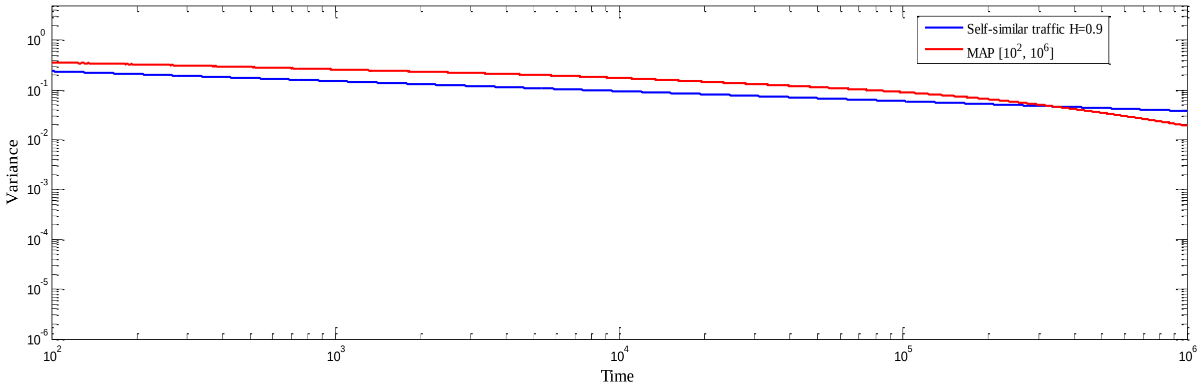

Figure 9: Variance versus time curves with $n = 4$ for sample3 during time scale $[10^2, 10^6]$

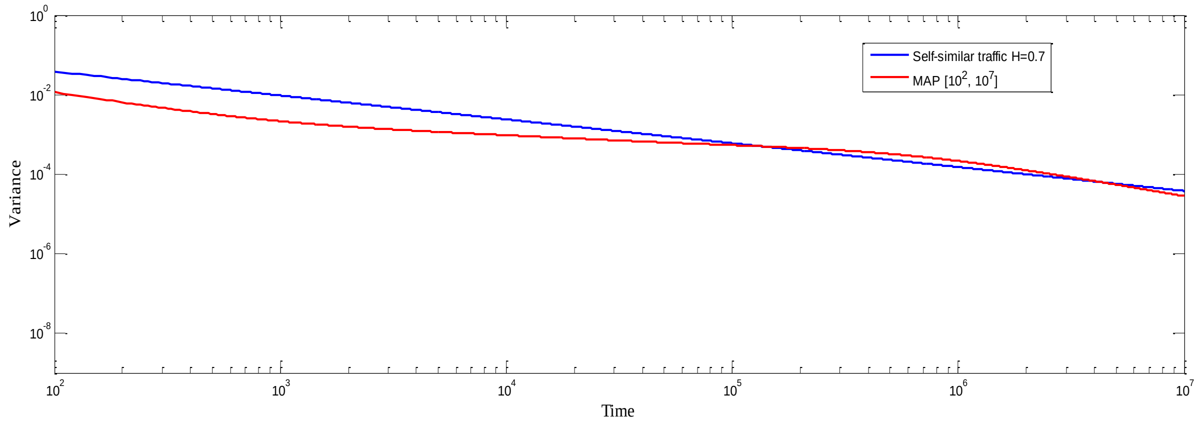

Figure 10: Variance versus time curves with $n = 4$ for sample1 during time $[10^2, 10^7]$

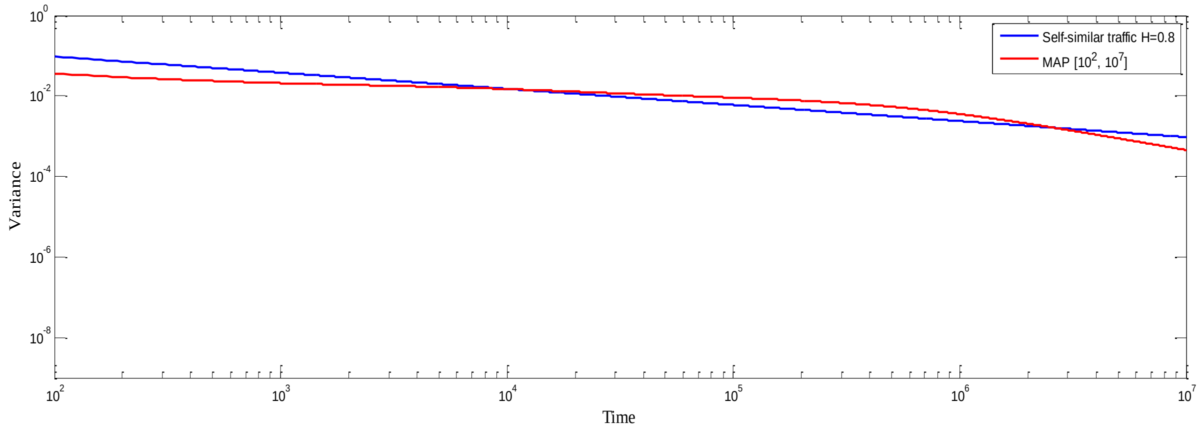

Figure 11: Variance versus time curves with $n = 4$ for sample2 during time $[10^2, 10^7]$

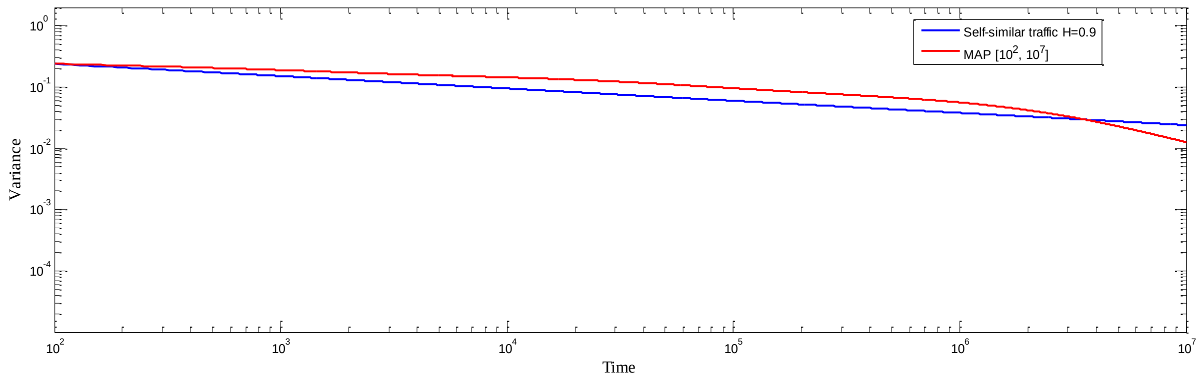

Figure 12: Variance versus time curves with $n = 4$ for sample3 during time $[10^2, 10^7]$

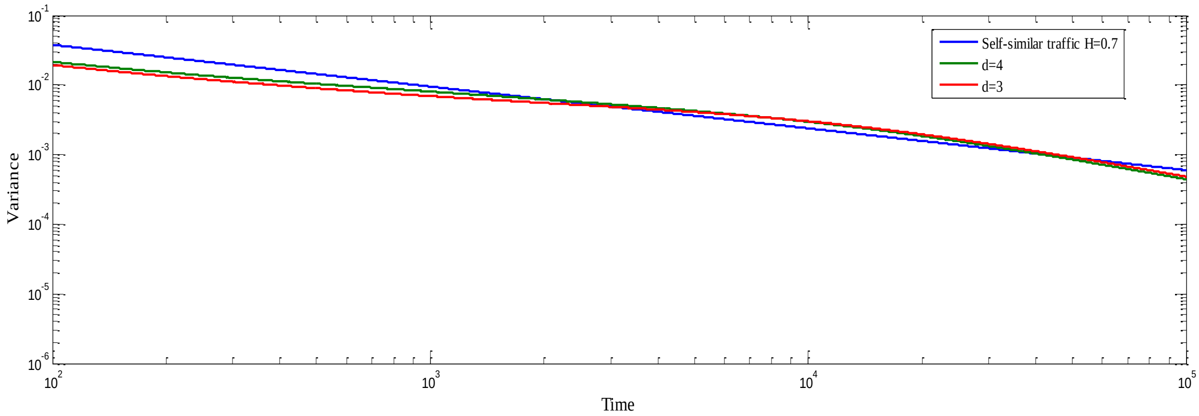

Figure 13: Variance vs time curves with $n = 3, 4$ over time-scale $[10^2, 10^5]$ for sample1

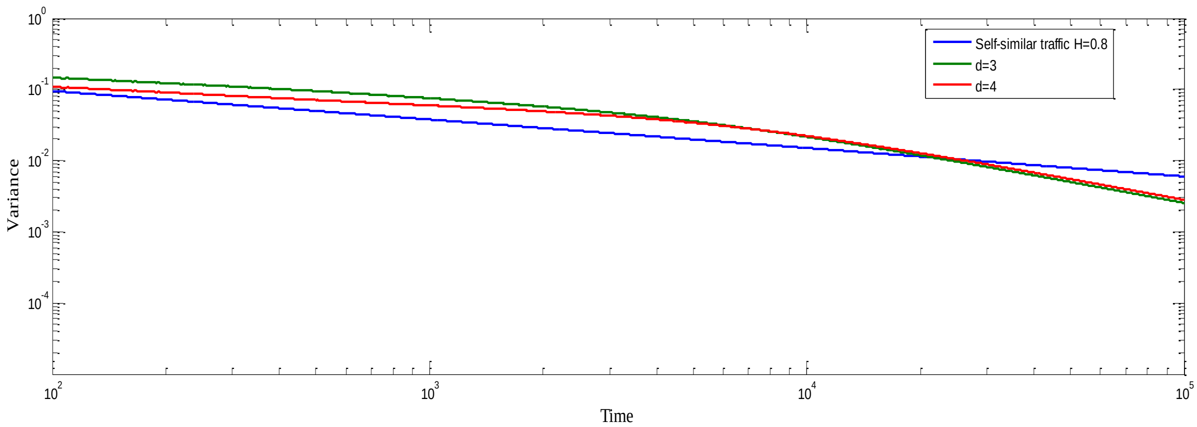

Figure 14: Variance vs time curves with $n = 3, 4$ over time-scale $[10^2, 10^5]$ for sample2

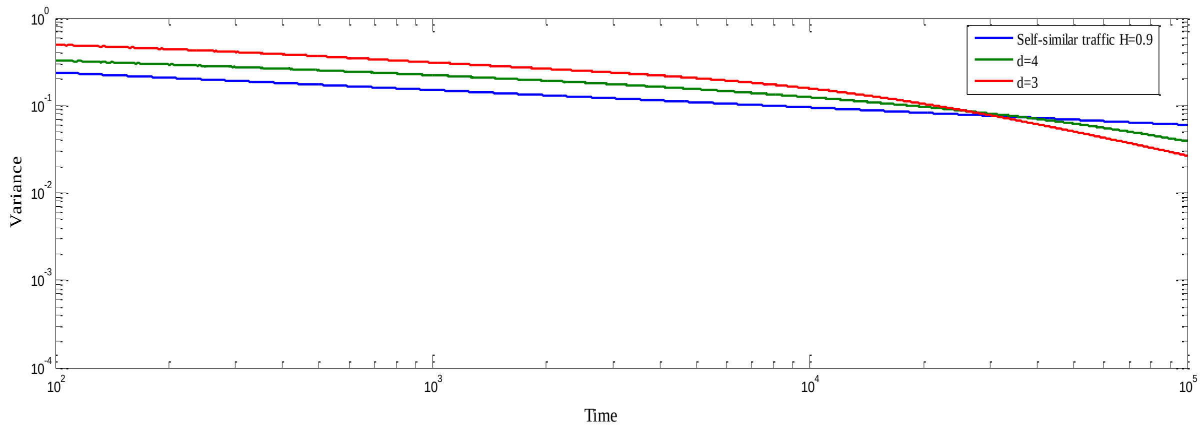

Figure 15: Variance vs time curves with $n = 3, 4$ over time-scale $[10^2, 10^5]$ for sample3

5. Conclusion

Self-similar traffic models proposed earlier are independent of time (homogeneous), and they do not work for time dependent queueing analysis. Here, variance based Markovian fitting procedure is presented using time dependent arrival rates. For the validation of fitting variance–time curves are presented, which show how the resultant MMPPs exhibit legitimate agreement with that of self-similar traffic in specified time scales. In addition, it is seen that the accuracy improved as number of MMPPs in superposition increases. This model is useful for time dependent queuing based performance analysis.

Acknowledgement

The authors wish to acknowledge Council of Scientific and Industrial Research (CSIR), Government of India, for their funding under the Major Research Project (MRP) scheme (File. No: 25(0301)/19/EMR-II).

References

Abate, J. and Whitt, W. (1988). Transient behavior of the M/M/1 queue via Laplace transforms. *Advances in Applied Probability*, **20**(**1**), 145-178.

Andersen, A. T. and Nielsen, B. F. (1998). A Markovian approach for modeling packet traffic with long-range dependence. *IEEE Journal on Selected Areas in Communications*, **16**(**5**), 719-732.

Coddington, A. E. and Levinson, N. (1987). *Theory of Ordinary Differential Equations*. 9th Reprint, Tata McGraw-Hill.

Crovella, M. E. and Bestavros, A. (1997). Self-similarity in world wide web traffic: evidence and possible causes. *IEEE/ACM Transactions on Networking*, **5**(**6**), 835-846.

Eick, S. G., Massey, W. A. and Whitt, W. (1993). $M_t/G/\infty$ Queues with Sinusoidal Arrival Rates. *Management Science*, **39**(**2**), 241-252.

Erramilli, A., Narayan, O. and Willinger, W. (1996). Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Transactions on Networking*, **4**(**2**), 209-223.

Fisher, W. and Hellstern, M. K. (1993). The Markov-modulated Poisson process (MMPP) cook book. *Performance Evolution*, **18**(**2**), 149-171.

Heffes, H. and Lucantoni, D. (1986). A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE Journal on Selected Areas in Communications*, **4**(**6**), 856-868.

Jennings, O. B. and Massey, W. A. (1997). A modified offered load approximation for nonstationary circuit switched networks. *Telecommunication Systems*, **7**, 229-251.

Kasahara, S., Yoshihara, T. and Takahasi, Y. (2001). Practical time-scale fitting of self-similar traffic with Markov-modulated Poisson process. *Telecommunication Systems*, **17,** 185-211.

Kelton, W. D. and Law, A. M. (1985). The transient behavior of M/M/s queue with implications for steady state simulation. *Operation Research*, **33**(**2**), 378-396.

Leland, W. E., Taqqu, M. S., Willinger, W. and Wilson, D. V. (1994). On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transanctions on Networking*, **2**(**1**), 1-15.

Liu, Y. and Whitt, W. (2014). Stabilizing performance in networks of queue with time-varying arrival rates. *Probability in the Engineering and Informational Sciences*, **28**(**4**), 419-449.

Massey, W. A. (2002). The analysis of queues with time-varying rates for telecommunication models. *Telecommunication Systems,* **21**, 173-204.

Misra, S., Banerjee, B. and Wolfinger, B. E. (2012). A learning automata-based uplink scheduler for supporting real-time multimedia interactive traffic in IEEE 802.16 WiMAX networks. *Computer Communications,* **35**(**15**), 1871-1881.

Neuts, M. (1979). A versatile Markovian point process. *Journal of Applied Probability*, **16**(**4**), 764-779.

Norros, I. (1994). A storage model with self-similar input. *Queueing Systems*, **16**, 387-396.

Pant, A. and Ghimire, R. (2016). M(t)/M/1 queueing system with sinusoidal arrival rate. *Journal of the Institute of Engineering*, **11**(**1**), 120-127.

Paxson, V. and Floyd, S. (1995). Wide area traffic: The failure of Poisson modeling. *IEEE/ACM Transactions on Networking*, **3**(**3**), 226-244.

Qian, Y. and Tipper, D. (2004). Adaptive channel allocation scheme for next generation wireless networks. *IEEE 60th Vehicular Technology Conference, VTC 2004-Fall*, Los Angels, CA, USA, **7,** 4918-4922.

Reddy, P. M., Shao, S. K., Tsai, M. G., Tsao, H. W. and Wu, J. (2005). Generalized variance-based Markovian fitting for self-similar traffic modeling. *IEICE Transactions on Communications*, **E88-B**(**4**), 1493-1502.