Modeling Self-Similar Traffic through Markov Modulated Poisson Processes over Multiple Time Scales

António Nogueira[1], Paulo Salvador[1], Rui Valadas[1], and António Pacheco[2]

[1] University of Aveiro / Institute of Telecommunications Aveiro
Campus de Santiago, 3810-193 Aveiro, Portugal
{nogueira, salvador}@av.it.pt, rv@det.ua.pt
[2] Instituto Superior Técnico - UTL / Department of Mathematics and CEMAT
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
apacheco@math.ist.utl.pt

Abstract. In recent years several studies have reported peculiar types of traffic behavior, such as long-range dependence and self-similarity, which can have significant impact on network performance. In this paper we propose a novel traffic model and parameter fitting procedure, based on Markov Modulated Poisson Processes (MMPPs), which is able to capture variability over many time scales, a characteristic of self-similar traffic. The fitting procedure matches the complete distribution at each time scale, and not only some of its moments as it is the case in related proposals. Our results show that the proposed traffic model and parameter fitting procedure closely matches the main characteristics of measured traces over the time scales present in data.

keywords: Traffic modeling, self-similar, time scale, Markov Modulated Poisson Process.

1 Introduction

An efficient design and control of data networks needs to take into account the main characteristics of the supported traffic. Due to the growing diversity of services and applications, there is a strong requirement to make frequent measurements of packet flows and to describe them through appropriate traffic models. Since the work by Leland *et al.* [1] several studies have shown that network traffic may exhibit properties of self-similarity and/or long-range dependence (LRD) [1–4], which have significant impact on network performance. Self-similar traffic shows identical statistical characteristics over a wide range of time scales. In general, self-similarity implies long-range dependence, and vice-versa.

Several works have addressed the impact of LRD on network performance. References [4–7] study the case of a single queue and conclude that the buffer occupancy is not affected by autocovariance lags that are beyond the so-called critical time scale (CTS) or correlation horizon (CH), which depends on system parameters such as the buffer capacity. Similar conclusions are observed for the case of tandem queues in [8]. Thus, matching the LRD is only required within the time scales specific to the system under study. One of the consequences of this result is that more traditional

traffic models, such as Markov Modulated Poisson Processes (MMPPs), can still be used to model traffic exhibiting LRD. The use of MMPPs also benefits from the existence of several tools for calculating the queuing behavior and the effective bandwidth.

In this paper we propose a novel traffic model and parameter fitting procedure, which captures self-similar behavior over a range of time scales. The traffic model is a superposition of discrete-time MMPPs (dMMPPs), where each dMMPP represents a specific time scale. The parameter fitting procedure matches, at each time scale, a dMMPP to the empirical probability function characteristic of that time scale. The number of states of each dMMPP is not fixed a priori; it is determined as part of the fitting procedure. The accuracy of the fitting procedure is evaluated by applying it to measured traffic traces that exhibit self-similar behavior: the well-known pOct Bellcore trace and a trace of aggregated IP WAN traffic. We compare the probability function at each time scale, and the queuing behavior (as assessed by the loss probability and average waiting time), corresponding to the measured and to synthetic traces generated from the inferred models. Our results show that the proposed fitting method is very effective in matching the probability function at the various time scales and leads to an accurate prediction of the queuing behavior.

Fitting procedures for MMPPs with an arbitrary number of states mainly concentrate on matching first- and/or second-order statistics, without addressing directly the issue of modeling on multiple time scales [9–13]. Yoshihara *et al.* [14] developed a fitting method for self-similar traffic based on the superposition of 2-MMPPs, that matches the variance at each time scale. In this way, the resulting MMPP reproduces the variance-scale curve characteristic of self-similar processes. Our contribution is to develop a procedure that matches the complete distribution at each time scale (and not only the variance) in order to reproduce accurately self-similar behavior.

The paper is organized as follows. Section 2 gives the required background on MMPPs and self-similarity and presents the various steps of the parameter fitting procedure. Section 3 briefly describes the data traces used in the numerical evaluation and in section 4 we discuss the results. Finally, section 5 presents the main conclusions.

2 Inference Procedure

Our inference procedure is closely related to the notion of distributional self-similarity. Consider the continuous-time process $Y(t)$ representing the traffic volume (e.g. in bytes) from time 0 up to time t and let $X(t) = Y(t) - Y(t-1)$ be the corresponding increment process (e.g. in bytes/second). Consider also the sequence

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=1}^{m} X((k-1)m + i), k = 1, 2, ... \tag{1}$$

obtained by averaging $X(t)$ over non-overlapping blocks of length m. $Y(t)$ is exactly self-similar when it is equivalent, in the sense of finite-dimensional distributions, to $a^{-H}Y(at)$, for all $t > 0$ and $a > 0$, where H $(0 < H < 1)$ is the Hurst parameter. Clearly, the process $Y(t)$ can not be stationary. However, if $Y(t)$ has stationary

increments then again $X(k) = X^{(1)}(k)$ is equivalent, in the sense of finite-dimensional distributions, to $m^{1-H}X^{(m)}(k)$. This illustrates that a traffic model developed for fitting self-similar behavior must preferably enable the matching of the distribution on several time scales. Note also that, in general, self-similarity implies LRD, and vice-versa. An excellent overview of self-similarity and LRD can be found in [15].

The inference procedure estimates one dMMPP for each time scale that matches a probability mass function (PMF) characteristic of that time scale. The resulting dMMPP is obtained from the superposition of all dMMPPs inferred for each time scale. An (homogeneous) Markov chain $(Y, J) = \{(Y_k, J_k),\ k = 0, 1, \ldots\}$ with state space $\mathbb{N}_0 \times S$ is a dMMPP if and only if Y has non-decreasing sample paths and

$$P(Y_{k+1} = m,\ J_{k+1} = j | Y_k = n,\ J_k = i) = p_{ij}\, e^{-\lambda_i} \lambda_i^{m-n}/(m-n)! \qquad (2)$$

for $k = 0, 1, \ldots, m, n \in \mathbb{N}_0$ with $n \leq m$, and $i, j \in S$, where $\lambda_i, i \in S$ are nonnegative real constants and $\mathbf{P} = (p_{ij})$ is a stochastic matrix. In this case we say that (Y, J) is a dMMPP with set of modulating states S and parameter (matrices) $\mathbf{P}$ and $\mathbf{\Lambda}$, and write $(Y, J) \sim \text{dMMPP}_S(\mathbf{P}, \mathbf{\Lambda})$, where $\mathbf{\Lambda} = (\lambda_{ij}) = (\lambda_i \delta_{ij})$. The matrix $\mathbf{P}$ is the transition probability matrix of the modulating Markov chain J, whereas $\mathbf{\Lambda}$ is the matrix of Poisson arrival rates. If S has cardinality r, we say that (Y, J) is a dMMPP of order r (r-dMMPP). When, in particular, $S = \{1, 2, \ldots, r\}$ for some $r \in \mathbb{N}$, then we write simply that $(Y, J) \sim \text{dMMPP}_r(\mathbf{P}, \mathbf{\Lambda})$.

The superposition of independent dMMPPs is still an dMMPP. More precisely, if $(Y^{(l)}, J^{(l)}) \sim \text{dMMPP}_{r_l}(\mathbf{P}^{(l)}, \mathbf{\Lambda}^{(l)})$, $l = 1, 2, \ldots, L$, are independent, then their superposition $(Y, J) = (\sum_{l=1}^{L} Y^{(l)}, (J^{(1)}, J^{(2)}, \ldots, J^{(L)}))$ is a $\text{dMMPP}_S(\mathbf{P}, \mathbf{\Lambda})$ where $S = \{1, 2, \ldots, r_1\} \times \ldots \times \{1, 2, \ldots, r_L\}$,

$$\mathbf{P} = \mathbf{P}^{(1)} \otimes \mathbf{P}^{(2)} \otimes \ldots \otimes \mathbf{P}^{(L)} \qquad \text{and} \qquad \mathbf{\Lambda} = \mathbf{\Lambda}^{(1)} \oplus \mathbf{\Lambda}^{(2)} \oplus \ldots \oplus \mathbf{\Lambda}^{(L)} \qquad (3)$$

with $\oplus$ and $\otimes$ denoting the Kronecker sum and product, respectively. In our approach L, the number of considered time scales , is fixed *a priori* and the dimensions of the dMMPPs, $r_1, r_2, \ldots, r_L$, are computed as part of the fitting procedure.

The flowchart of the inference method is represented in figure 1 where, basically, four steps can be identified: (i) compute the data vectors (corresponding to the average number of arrivals per time interval) at each time scale; (ii) calculate the empirical PMF at the largest time scale and infer its dMMPP; (iii) for all other time scales (going from the largest to the smallest one), calculate the empirical PMF, deconvolve it from the empirical PMF of the previous time scale and infer a dMMPP that matches the resulting PMF; and (iv) calculate the final dMMPP through superposition of the dMMPPs inferred for each time scale. We will now describe these steps in detail.

2.1 Aggregation process

Having defined the sampling interval, Δt, the number of time scales, L, and the level of aggregation, a, the aggregation process starts by computing the data vector corresponding to the average number of arrivals in the smallest time scale (where the interval length equals the sampling interval), $N^{(1)}(k)$. Then, it calculates the data

vectors of remaining time scales, $N^{(l)}(k)$, $l = 2, ...L$, corresponding to the average number of arrivals in intervals of length $\Delta t \mathrm{a}^{(\mathrm{l}-1)}$. This is given by

$$N^{(l)}(k) = \left\lceil \frac{1}{a} \sum_{i=0}^{a-1} N^{(l-1)}(k+i) \right\rceil, \quad \frac{k-1}{a} \in I\!N_0 \tag{4}$$

and $N^{(l)}(k) = N^{(l)}(k-1)$, for $\frac{k-1}{a} \notin I\!N_0$, where $\lceil x \rceil$ is the ceiling function. Note that the block length of equation (1) is related with a and l by $m = a^{l-1}$. The empirical distribution of $N^{(l)}(k)$ will be denoted by $\hat{p}^{(l)}(x)$.

2.2 Calculation of the empirical PMFs

This step infers the PMFs that, at each time scale, must be fitted to a dMMPP. For the largest time scale, this PMF is simply the empirical one. For all other time scales l, $l = 1, 2, ..., L-1$, the associated dMMPP will model only the traffic components due to that scale. For time scale l, these traffic components can be obtained through deconvolution of the empirical PMFs of this time scale and of previous time scale, i.e., $\hat{f}_p^{(l)}(x) = [\hat{p}^{(l)} \otimes^{-1} \hat{p}^{(l+1)}](x)$. However, this may result in negative arrival rates for the dMMPP$^{(l)}$, which will occur whenever $\min\left\{x : \hat{p}^{(l+1)}(x) > 0\right\} < \min\left\{x : \hat{p}^{(l)}(x) > 0\right\}$. To correct this, the dMMPP$^{(l)}$ will be fitted to

$$\hat{f}^{(l)}(x) = \hat{f}_p^{(l)}(x + e^{(l)}) \tag{5}$$

where $e^{(l)} = \min\left(0, \min\left\{x : \hat{f}_p^{(l)}(x) > 0\right\}\right)$, which assures $\hat{f}^{(l)}(x) = 0$, $x < 0$. These additional factors are removed in the final step of the inference procedure.

2.3 Approximation of the empirical PMFs by a weighted sum of Poisson distributions and inference of the dMMPP parameters

Before inferring the dMMPP$^{(l)}$ parameters, $l = 1, 2, ..., L$, function $\hat{f}^{(l)}$ is approximated by a weighted sum of Poisson probability functions, using an algorithm that progressively subtracts a Poisson probability function from $\hat{f}^{(l)}$. The most important steps of this algorithm are depicted in the flowchart of figure 2 and will be explained in the next paragraphs.

Let the i^{th} Poisson probability function, with mean $\varphi_i^{(l)}$, be represented by $g_{\varphi_i^{(l)}}(x)$ and define $h_{(i)}^{(l)}(x)$ as the difference between $\hat{f}^{(l)}(x)$ and the weighted sum of Poisson probability functions at the i^{th} iteration. Initially, we set $h_{(1)}^{(l)}(x) = \hat{f}^{(l)}(x)$ and, in each step, we first detect the maximum of $h_{(i)}^{(l)}(x)$. The corresponding x-value, $\varphi_i = [h_{(i)}^{(l)}]^{-1}\left(\max h_{(i)}^{(l)}(x)\right)$, will be considered the i^{th} Poisson rate of the dMMPP$^{(l)}$. We then calculate the weights of each Poisson probability function, $\boldsymbol{w}_i^{(l)} = \left[w_{1i}^{(l)}, w_{2i}^{(l)}, ..., w_{ii}^{(l)}\right]$, through the following set of linear equations:

$$\hat{f}^{(l)}(\varphi_m^{(l)}) = \sum_{j=1}^{i} w_{ji}^{(l)} g_{\varphi_j^{(l)}}(\varphi_m^{(l)}) \tag{6}$$

for $m = 1, ..., i$ and $l = 1, ..., L$. This assures that the fitting between $\hat{f}^{(l)}(x)$ and the weighted sum of Poisson probability functions is exact at $\varphi_m^{(l)}$ points, for $m = 1, 2, \ldots, i$. The final step in each iteration is the calculation of the new difference function

$$h_{(i)}^{(l)}(x) = \hat{f}^{(l)}(x) - \sum_{j=1}^{i} w_{ji}^{(l)} g_{\varphi_j^{(l)}}(x). \tag{7}$$

The algorithm stops when the maximum of $h_{(i)}^{(l)}(x)$ is lower than a pre-defined percentage of the maximum of $\hat{f}^{(l)}(x)$ and r_l, the number of states of the dMMPP, is made equal to i.

Note that the number of states of each dMMPP depends on the level of accuracy employed in the approximation of the empirical PMF by the weighted sum of Poisson probability functions.

After r_l has been determined, the parameters $p_{ij}^{(l)}$ and $\lambda_j^{(l)}$, $j = 1, 2, \ldots, r_l$, of the $r_l -$ dMMPP, are set equal to

$$\pi_j^{(l)} = w_{jr_l}^{(l)} \qquad \text{and} \qquad \lambda_j^{(l)} = \varphi_j^{(l)}. \tag{8}$$

The next step is to associate, for each time scale l, one of the dMMPP$^{(l)}$ states with each time slot of the arriving process. The state that is assigned to a time interval is calculated randomly according to the probability vector $\boldsymbol{\theta}^{(l)}(k) = \left\{\theta_1^{(l)}(k), \ldots, \theta_{r_l}^{(l)}(k)\right\}$, with

$$\theta_i^{(l)}(k) = g_{\lambda_i^{(l)}}\left(N^l(k)\right) \Big/ \sum_{j=1}^{r_l} g_{\lambda_j^{(l)}}\left(N^l(k)\right), \tag{9}$$

where $i = 1, ..., r_l$, $l = 1, ..., L$ and $g_\lambda(y)$ represents a Poisson probability distribution function with mean λ. The elements of this vector represent the probability that the state j had originated the number of arrivals $N^{(l)}(k)$ at time slot k from time scale l.

After this step, we infer the dMPPP$^{(l)}$ transition probabilities, $p_{ij}^{(l)}, i, j = 1, ..., r_l$, by counting the number of transitions between each pair of states. If $n_{ij}^{(l)}$ represents the number of transitions from state i to state j, corresponding to the dMPPP$^{(l)}$, then

$$p_{ij}^{(l)} = n_{ij}^{(l)} / \sum_{m=1}^{r_l} n_{mj}^{(l)}, j = 1, ..., r_l \tag{10}$$

The transition probability and the Poisson arrival rate matrices of the dMPPP$^{(l)}$ are then given by $\mathbf{P}^{(l)} = (p_{ij}^{(l)})$ and $\mathbf{\Lambda}^{(l)} = (\lambda_i^{(l)} \delta_{ij})$, for $l = 1, ..., L$.

2.4 Final dMMPP model construction

The final dMMPP process is constructed using equation (3), where the matrices $\mathbf{\Lambda}^{(l)}$ and $\mathbf{P}^{(l)}$, $l = 1, ..., L$, were calculated in the last subsection. However, the additional factors introduced in sub-section 2.2 must be removed. Thus, $\mathbf{\Lambda} = \mathbf{\Lambda} - \sum_{l=1}^{L-1} e^{(l)} \cdot \mathbf{I}$ where $\mathbf{I}$ is the identity matrix.

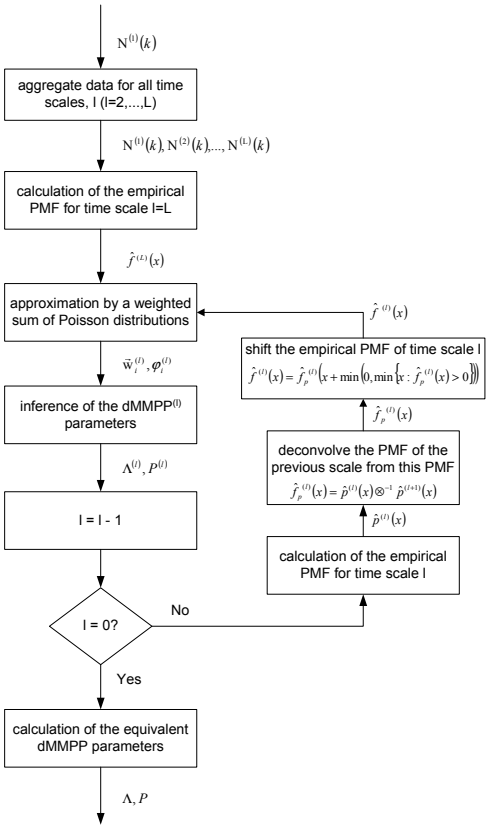

Fig. 1. Flow diagram of the inference procedure.

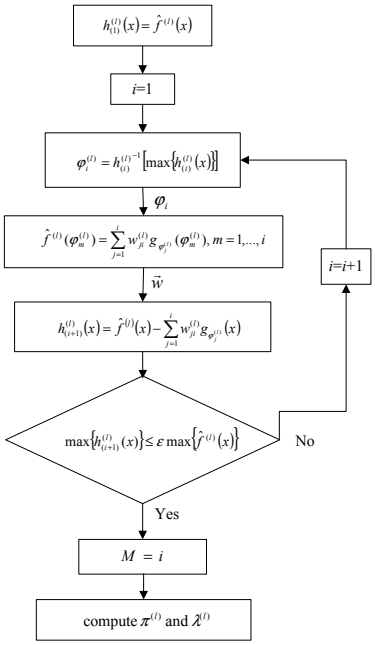

Fig. 2. Algorithm for calculating the number of states and the Poisson arrival rates of the dMMPP$^{(l)}$.

3 Overview of the traffic traces

We consider two traffic traces to evaluate the fitting procedure: (i) the well-known and publicly available pOct LAN trace from Bellcore [1] and (ii) a trace corresponding to the downstream Internet access traffic of approximately 65 simultaneous users, measured at the access link of a Portuguese ISP (to an ADSL network). The traffic analyzer was a 1.2 GHz AMD Athlon PC, with 1.5 Gbytes of RAM and running WinDump; we recorded the arrival instant and the IP header of each packet. The main characteristics of the selected traces are described in Table 1.

We assess the self-similar (LRD) behavior through the semi-parametric estimator developed in [16]. Here, one looks for alignment in the so-called Logscale Diagram (LD), which is a log-log plot of the variance estimates of discrete wavelet transform coefficients, against scale, complete with confidence intervals about these estimates at each scale. Traffic is said to be LRD if, within the limits of the confidence intervals, the log of the variance estimates fall on a straight line, in a range of scales from some initial

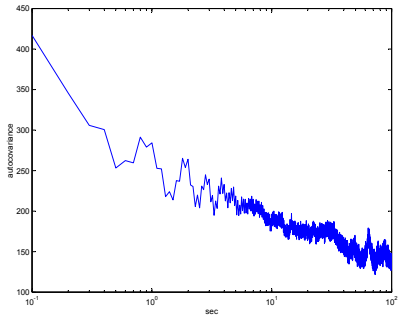

Fig. 3. Autocovariance of packet counts, trace pOct.

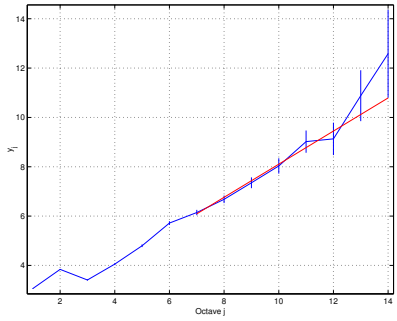

Fig. 4. Second order Logscale Diagram, trace pOct.

value j_1 up to the largest one present in data and the slope of the straight line, which is an estimate of the scaling exponent α, lies in $(0, 1)$.

Both traces exhibit self-similar (LRD) behavior. For example, taking the case of trace pOct, the analysis of its autocovariance function (Figure 3) lead us to suspect that it exhibits LRD, due to the slow decay for large time lags. This is confirmed by the scaling analysis, since the y_j values in the Logscale Diagram are aligned between a medium octave (7) and octave 14, the highest one present in data (Figure 4). A similar analysis was made for the other trace, also revealing the same self-similar (LRD) behavior.

4 Numerical Results

We assess the suitability of the proposed dMMPP fitting procedure using several criteria: (i) comparing the Hurst parameters of the original and synthesized (from the inferred dMMPP) data traces; (ii) comparing the probability functions of the average number of arrivals in different time scales, obtained for the original and synthesized traces and (iii) comparing the queuing behavior, in terms of packet loss ratio (PLR), of the original and synthesized traces, using trace-driven simulation. All simulations were carried out using a fixed packet length equal to the mean packet length of the trace. For all traces, the sampling interval of the counting process was chosen to be 0.1s and three different time scales were considered: 0.1s, 1s and 10s. For each trace, the estimation procedure took less than 1 minute, using a MATLAB implementation running in the PC described above, which shows that the procedure is computationally very efficient.

In order to verify that the proposed fitting approach captures the self-similar behavior, we compare in Table 2 the Hurst parameters estimated for the original and

Trace name	Capture period	Trace size (pkts)	Mean rate (byte/s)	Mean pkt size (bytes)
October	Bellcore trace	1 million	322790	568
ISP	10.26pm to 10.49pm, October 18^{th}2002	1 million	583470	797

Table 1. Main characteristics of measured traces.

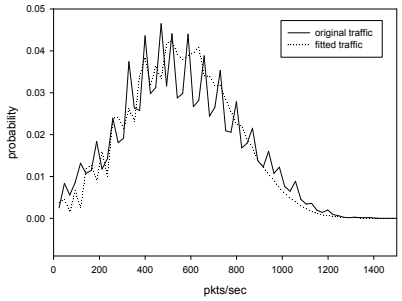

Fig. 5. Probability mass function at the smallest time scale, trace pOct.

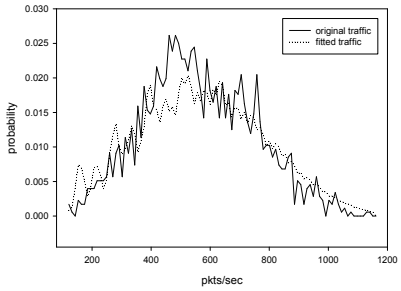

Fig. 6. Probability mass function at the intermediate time scale, trace pOct.

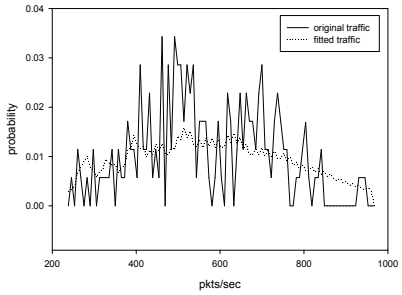

Fig. 7. Probability mass function at the largest time scale, trace pOct.

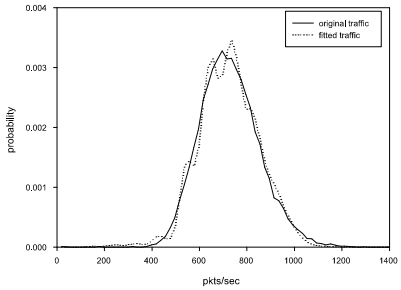

Fig. 8. Probability mass function at the smallest time scale, trace ISP.

fitted traffic, for each selected data trace. Table 2 also includes the range of time scales where the y_j follow a straight line, inside brackets near to the corresponding Hurst parameter value. There is a very good agreement between the Hurst parameter values of the original and fitted traffic, so LRD behavior is indeed well captured by our model.

The next evaluation criteria is based on the comparison between the probability functions of the original and fitted traces, for different time scales. Starting with trace pOct, it can be seen in Figures 5, 6 and 7 that there is a good agreement between the probability functions of the original and fitted traces, for all time scales. Recall that the fitting procedure explicitly aimed at matching the probability function at the various time scales; these results confirm that the procedure is effective in performing this task. Due to space limitations, for the case of the ISP trace we only show the comparison between the probability functions at the smallest time scale in Figure 8. However, a good agreement also exists at the three times scales.

Considering now the queuing behavior, we compare the PLR obtained, through trace-driven simulation, with the original and fitted traces. Two different sets of utilization ratios were used in the simulations: for trace pOct, we used $\rho = 0.6$ and $\rho = 0.7$ and, for trace ISP, the selected values were $\rho = 0.8$ and $\rho = 0.9$. This is due to the lower burstiness of the ISP traffic, which leads to lower packet losses for

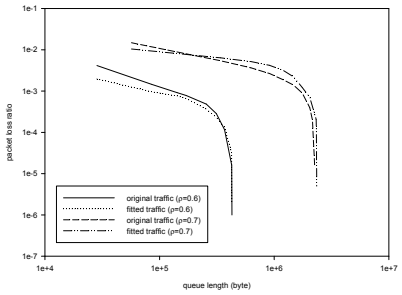

Fig. 9. Packet loss ratio, trace pOct.

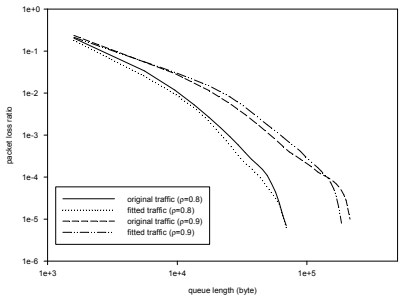

Fig. 10. Packet loss ratio, trace ISP.

the same link utilization. From figures 9 and 10, it can be seen that the PLR is very well approximated by the fitted dMMPPs for both utilization ratios. Nevertheless, as the utilization ratio increases the deviation slightly increases, because the sensitivity of the metrics variation to a slight difference in the buffer size is higher.

As a final remark, we can say that the proposed fitting approach provides a close match of the Hurst parameters and probability mass functions at each time scale, and this agreement reveals itself sufficient to drive a good queuing performance in terms of packet loss ratio. The computational complexity of the fitting method is also very small.

Trace	original	fitted
October	0.941 (6,12)	0.962 (6,12)
ISP	0.745 (6,13)	0.784 (4,13)

Table 2. Comparison between Hurst parameter values

5 Conclusions

In this paper we have proposed a novel traffic model and parameter fitting procedure, based on Markov Modulated Poisson Processes (MMPPs), which is able to capture self-similarity over a range of time scales. The fitting procedure matches the complete distribution at each time scale, and not only some of its moments as it is the case in related proposals. We evaluated the procedure by comparing the probability function at each time scale, and the queuing behavior (as assessed by the loss probability), corresponding to measured traffic traces and to traces synthesized according to the proposed model. Two traffic traces were considered, all exhibiting self-similar behavior: the well-known pOct Bellcore trace and a trace of aggregated IP WAN traffic. Our results show that the proposed traffic model and parameter fitting procedure closely matches the main characteristics of the measured traces over the time scales present in data.

Acknowledgements: This research was supported in part by Fundação para a Ciência e a Tecnologia, the project POSI/42069/CPS/2001, and the grant BD/19781/99.

References

1. W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, Feb 1994.
2. J. Beran, R. Sherman, M. Taqqu, and W. Willinger, "Long-range dependence in variable-bit rate video traffic," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 1566–1579, 1995.
3. M. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: Evidence and possible causes," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835–846, Dec 1997.
4. B. Ryu and A. Elwalid, "The importance of long-range dependence of VBR video traffic in ATM traffic engineering: Myths and realities," *ACM Computer Communication Review*, vol. 26, pp. 3–14, Oct 1996.
5. D.P.Heyman and T.V.Lakshman, "What are the implications of long range dependence for vbr video traffic egineering?," *IEEE/ACM Transactions on Networking*, vol. 4, no. 3, pp. 301–317, June 1996.
6. A. Neidhardt and J. Wang, "The concept of relevant time scales and its application to queuing analysis of self-similar traffic," in *Proceedings of SIGMETRICS'1998/PERFORMANCE'1998*, 1998, pp. 222–232.
7. M. Grossglauser and 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.
8. A. Nogueira and R. Valadas, "Analyzing the relevant time scales in a network of queues," in *SPIE's International Symposium ITCOM 2001*, August 2001.
9. P. Skelly, M. Schwartz, and S. Dixit, "A histogram-based model for video traffic behaviour in an ATM multiplexer," *IEEE/ACM Transactions on Networking*, pp. 446–458, Aug 1993.
10. S. Li and C. Hwang, "On the convergence of traffic measurement and queuing analysis: A statistical-match and queuing (SMAQ) tool," *IEEE/ACM Transactions on Networking*, pp. 95–110, Fev 1997.
11. A. Andersen and B. Nielsen, "A Markovian approach for modeling packet traffic with long-range dependence," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp. 719–732, Jun 1998.
12. P. Salvador and R. Valadas, "A fitting procedure for Markov modulated Poisson processes with an adaptive number of states," in *Proceedings of the 9th IFIP Working Conference on Performance Modelling and Evaluation of ATM & IP Networks*, June 2001.
13. P. Salvador, A. Pacheco, and R. Valadas, "Multiscale fitting procedure using Markov modulated Poisson processes," *Telecommunications Systems*, vol. 23, no. 1-2, pp. 123–148, June 2003.
14. T. Yoshihara, S. Kasahara, and Y. Takahashi, "Practical time-scale fitting of self-similar traffic with Markov-modulated Poisson process," *Telecommunication Systems*, vol. 17, no. 1-2, pp. 185–211, 2001.
15. K. Park and W. Willinger, "Self-similar network traffic: an overview," in *Self-Similar Network Traffic and Performance Evaluation*, K. Park and W. Willinger, Eds. Wiley-Interscience, 2000.
16. D. Veitch and P. Abry, "A wavelet based joint estimator for the parameters of LRD," *IEEE Transactions on Information Theory*, vol. 45, no. 3, Apr 1999.