A Framework based on Markov Modulated Poisson Processes for Modeling Traffic with Long-Range Dependence

Paulo Salvador† and Rui Valadas‡

Institute of Telecommunications
University of Aveiro
Campus Universitário, Aveiro, Portugal

ABSTRACT

This paper proposes a novel fitting procedure for Markov Modulated Poisson Processes (MMPPs), consisting of the superposition of N 2-MMPPs, that is capable of capturing the long-range characteristics of the traffic. The procedure matches both the autocovariance and marginal distribution functions of the rate process. We start by matching each 2-MMPP to a different component of the autocovariance function. We then map the parameters of the model with N individual 2-MMPPs (termed superposed MMPP) to the parameters of the equivalent MMPP with 2^N states that results from the superposition of the N individual 2-MMPPs (termed generic MMPP). Finally, the parameters of the generic MMPP are fitted to the marginal distribution, subject to the constraints imposed by the autocovariance matching. Specifically, the matching of the distribution will be restricted by the fact that it may not be possible to decompose a generic MMPP back into individual 2-MMPPs. Overall, our procedure is motivated by the fact that direct relationships can be established between the autocovariance and the parameters of the superposed MMPP and between the marginal distribution and the parameters of the generic MMPP.

We apply the fitting procedure to traffic traces exhibiting LRD including (i) IP traffic measured at our institution and (ii) IP traffic traces available in the Internet such as the well known, publicly available, Bellcore traces. The selected traces are representative of a wide range of services/protocols used in the Internet. We assess the fitting procedure by comparing the measured and fitted traces (traces generated from the fitted models) in terms of (i) Hurst parameter; (ii) degree of approximation between the autocovariance and marginal distribution curves; (iii) range of time scales where LRD is observed using a wavelet based estimator and (iv) packet loss ratio suffered in a single buffer for different values of the buffer capacity. Results are very clear in showing that MMPPs, when used in conjunction with the proposed fitting procedure, can be used to model efficiently Internet traffic in the relevant time scales, even when exhibiting LRD behavior.

Keywords: Traffic modeling, traffic characterization, MMPP, long-range dependence, time scales, QoS

1. INTRODUCTION

Traffic modeling plays an increasingly important role in the management and planning of modern telecommunications networks. In order to make an efficient use of network resources, operators are required to perform frequent traffic measurements and to derive traffic models capable of describing rigorously its data. When selecting a stochastic model to describe a traffic source, there is the need to consider the fitting procedures available for parameter estimation. The design of the fitting procedure is a trade-off between computational complexity and accuracy and requires careful consideration of the model parameters that have more impact on the performance metrics of interest. This integrated approach has driven several works[1–7] .

† e-mail: salvador@av.it.pt; ‡ e-mail: rv@det.ua.pt
phone: +351 234377900; fax: +351 234377901; Institute of Telecommunications, University of Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

Published in Internet Performance and Control of Network Systems II, Proceedings SPIE vol. 4523, Robert D. van der Mei and Frank Huebner-Szabo de Bucs, Eds., pp. 221-232, August 2001

In recent years it has been clearly shown through experimental evidence, that network traffic may exhibit properties of self-similarity and long-range dependence (LRD)[1,8–13] . These characteristics have significant impact on network performance. However, as pointed out in Ref. 4, matching the LRD is only required within the time scales of interest to the system under study. For example, in order to analyze queuing behavior, the selected traffic model needs only to capture the correlation structure of the source up to the so-called correlation horizon, which is directly related to the maximum buffer size. 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.

Providing a good match of the LRD behavior is not enough for accurate prediction of the queuing behavior. The first-order statistics need also careful consideration. The work in Ref. 14 discusses the limitations of using only the mean and the autocorrelation function, as statistical descriptors of the input process for the purpose of analyzing queuing performance. The authors show that the mean queue length can vary substantially when the parameters of the input process are varied, subject to the same mean and autocorrelation function. We also include an illustration of this fact latter in the paper. While this issue is of major importance, we believe it remains largely neglected. The main goal of the present work is to develop a parameter fitting procedure for MMPPs that matches closely both the autocovariance and the complete marginal distribution (as opposed to matching only the mean). A similar objective was pursued in Ref. 15, in the context of Circulant Modulated Poisson Processes (CMPPs).

There have been several proposals of fitting procedures for MMPPs[1,3,5,11,15–18] . Most of these procedures apply to 2-MMPPs and do not address the simultaneous matching of the autocovariance and marginal distribution functions. We consider a model consisting of the superposition of N 2-MMPPs. We start by matching each 2-MMPP to a different component of the autocovariance function. We then map the parameters of the model with N individual 2-MMPPs (termed superposed MMPP) to the parameters of the equivalent MMPP with 2^N states that results from the superposition of the N individual 2-MMPPs (termed generic MMPP). Finally, the parameters of the generic MMPP are fitted to the marginal distribution, subject to the constraints imposed by the autocovariance matching. Specifically, the matching of the distribution will be restricted by the fact that it may not be possible to decompose a generic MMPP back into individual 2-MMPPs. One consequence is that our fitting procedure favors matching the autocovariance, as opposed to the marginal distribution matching. We believe that, in general, this agrees with the relative importance of these two statistics in terms of queuing behavior. A similar prioritization was adopted in Ref. 15. Overall, our procedure is motivated by the fact that direct relationships can be established between the autocovariance and the parameters of the superposed MMPP and between the marginal distribution and the parameters of the generic MMPP.

We apply the fitting procedure to traffic traces exhibiting LRD, including the well known, publicly available, Bellcore traces. The LRD characteristics are analyzed using the wavelet based estimator of Ref. 19. Results show that the MMPPs obtained through the fitting procedure are capable of modeling the LRD behavior present in data. The fitting procedure is also assessed in terms of queuing behavior. Results show a very good agreement between the packet loss ratio obtained with the original data traces and with traces generated from the fitted MMPPs.

This paper is organized as follows. In section 2 we give some background on the superposition of MMPPs and on the mapping relations between superposed and generic MMPPs. In section 3 we describe the fitting procedure. In section 4 we present numerical results, which include applying the fitting procedure to measured traffic traces. In section 5, we present some related work. Finally, in section 6 we conclude the paper.

2. BACKGROUND

The MMPP with M states is fully characterized by the infinitesimal generator matrix, $\mathbf{Q}$, and by the diagonal matrix of the Poisson arrival rates, $\mathbf{\Lambda}$,

$$\mathbf{Q} = \begin{bmatrix} -\sigma_1 & \sigma_{12} & \dots & \sigma_{1M} \\ \sigma_{21} & -\sigma_2 & \dots & \sigma_{2M} \\ \dots & \dots & \dots & \dots \\ \sigma_{M1} & \sigma_{M2} & \dots & -\sigma_M \end{bmatrix} \qquad \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_M \end{bmatrix}$$

where $\sigma_i = -\sum_{i \neq j} \sigma_{ij}$, σ_{ij} is the transition rate from state i to state j, and li is the Poisson rate in state i. We also represent diagonal matrix $\mathbf{\Lambda}$ by vector $\vec{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_M]$. The steady-state probability vector of $\mathbf{Q}$ is the solution of the following system of equations: and , where is a unit column vector. Consider the superposition of N MMPPs,

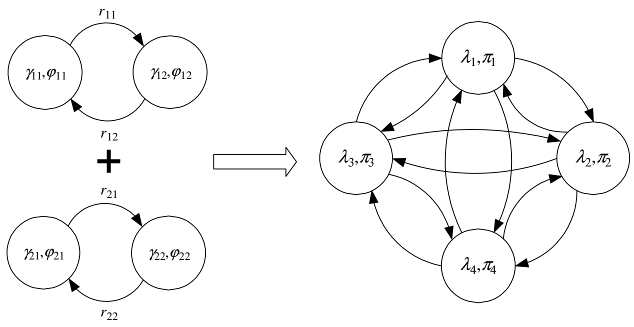

Figure 1. Superposed (left) and generic (right) MMPPs (N = 2).

each characterized by matrixes $\mathbf{Q}_j$ and $\mathbf{\Lambda}_j$. This process is also an MMPP, with infinitesimal generator matrix $\mathbf{Q}$ and arrival rate matrix $\mathbf{\Lambda}$

$$\mathbf{Q} = \mathbf{Q}_1 \oplus \mathbf{Q}_2 \oplus ... \oplus \mathbf{Q}_N \tag{1}$$

$$\mathbf{\Lambda} = \mathbf{\Lambda}_1 \oplus \mathbf{\Lambda}_2 \oplus ... \oplus \mathbf{\Lambda}_N \tag{2}$$

where $\oplus$ denotes the Kronecker sum. We will restrict ourselves to the case of the superposition of 2-MMPPs. We represent diagonal matrix $\mathbf{\Lambda}_j$ of each individual 2-MMPP by vector $\vec{\gamma}_j = [\gamma_{j1}, \gamma_{j2}]$, the steady-state probabilities by vector $\vec{\varphi}_j = [\varphi_{j1}, \varphi_{j2}]$ and the infinitesimal generator matrix by

$$\mathbf{Q}_j = \begin{bmatrix} -r_{j1} & r_{j1} \\ r_{j2} & -r_{j2} \end{bmatrix} \tag{3}$$

In our approach we construct the MMPP from the superposition of N individual 2-MMPPs, as in Ref. 1. The proposed fitting procedure works within the state space of two equivalent models: the model of N individual 2-MMPPs and the model of a MMPP with 2^N states. We will refer to the first model as the superposed MMPP and to the second model as the generic MMPP. These are represented in figure 1. We work with the autocovariance and marginal distribution of the rate process. The autocovariance of the rate process of a MMPP with 2^N states can be described as a weighted sum of N exponentials,i.e.,

$$C(t) = \sum_{l=1}^{N} C_l\,(t) = \sum_{l=1}^{N} \alpha_l e^{-\beta_l t}$$

where each exponential can be associated to the autocovariance of a 2-MMPP. For the j-th 2-MMPP, the autocovariance is given by

$$C_l(t) = d_j^2 \varphi_{j1}\,(1 - \varphi_{j1})\, e^{-(r_{j1}+r_{j2})t}, \qquad j, l = 1, ..., N$$

where $d_j = \gamma_{j2} - \gamma_{j1}$, represents the difference between the Poisson rates of the two states. Thus,

$$\alpha_l = d_j^2 \varphi_{j1}\,(1 - \varphi_{j1})\,, \qquad j, l = 1, ..., N \tag{4}$$

$$\beta_l = r_{j1} + r_{j2}, \qquad j, l = 1, ..., N \tag{5}$$

These equations describe the constraints imposed by the autocovariance matching on the parameters of the superposed model. Note that there are 2^N possible associations of autocovariance exponential terms (indexed by l) and individual 2-MMPPs of the superposed model (indexed by j). Also, from equation (4), there are two d_j solutions for each α_l, one positive and another negative, giving a total of 2^N possible solutions. Thus, overall there are a total of 2^{2N} degrees of freedom, which can alleviate the above mentioned constraints.

In order to define the mapping between the states of the generic and superposed MMPP, let index $j, l = 1, ..., N$ represent the states of the generic MMPP and let indexes (j, k), $j = 1, ..., N$ and $k \in \{1, 2\}$, represent state k of the

j-th 2-MMPP of the superposed MMPP. There is a correspondence between each state of the generic MMPP and a set of N states of the superposed MMPP, where each set includes only one state from each 2-MMPP. Let this set be denoted by E_i. We define

$$E_i = \left\{ (j,k) : j = 1, 2, \ldots, N; k = 2 - \text{mod}\left(\left\lceil \frac{i}{2^{N-j}} \right\rceil, 2 \right) \right\} \tag{6}$$

where mod(x,2) represents x modulus 2 and $\lceil y \rceil$ the lowest integer greater than y. For example, in the case of N = 2, it results E_1 = {(1,1), (2,1)}, E_2 = {(1,1), (2,2)}, E_3 = {(1,2), (2,1)}, E4 = {(1,2), (2,2)}. Thus state 1 of the generic MMPP corresponds to having both individual 2-MMPPs of the superposed MMPP on state 1; state 2 corresponds to having the first 2-MMPP on state 1 and the second on state 2. Using this definition, the arrival rates and steady-state probabilities of the generic MMPP can be obtained from those of the superposed MMPP by

$$\lambda_i = \sum_{(j,k) \in E_i} \gamma_{jk}, \qquad i = 1, ..., N \tag{7}$$

$$\pi_i = \prod_{(j,k) \in E_i} \varphi_{jk}, \qquad i = 1, ..., N \tag{8}$$

It also results from the mapping relations defined by (6) that, one arrival rate of the generic MMPP, that we denote by λ_Δ, and N arrival rate differences of the superposed MMPP, completely determine the remaining $2^N - 1$ arrival rates of the generic MMPP, according to

$$\lambda_i = \lambda_\Delta + \sum_{j=1}^{N} \left(d_j \left(1 - \text{mod}\left(\left\lceil \frac{i}{2^{N-j}} \right\rceil, 2 \right) \right) \right), i = 1, ..., 2^N \tag{9}$$

Since the autocovariance imposes only the arrival rate differences, as seen from equation (4), λ_Δ will give an additional degree of freedom for matching the distribution. Although it is always possible to map the superposed MMPP into a generic MMPP, the opposite may not be true, i.e., the generic MMPP may not be decomposed into N 2-MMPPs. There will be one (or more) solutions for mapping the generic MMPP into a superposed MMPP only if matrices Q_j and Λ_j, $j = 1, ..., N$, can be determined from equations (1) and (2). For example, in the case of a generic 4-MMPP there will only be a solution for Λ_j, $j = 1, ..., 4$, if $\lambda_1 + \lambda_4 = \lambda_2 + \lambda_3$.

3. INFERENCE PROCEDURE

The inference procedure starts by defining the number of states of the MMPP, which is required to be a power of 2. It can be divided in three parts: (i) approximation of the autocovariance by a weighted sum of exponentials, (ii) matching of the marginal distribution with simultaneous parameter fitting and (iii) calculation of the final MMPP parameters. We work with a discretized version of the rate process, which is obtained from data by defining a sampling interval of duration τ, and making the rate constant and equal to the number of arrivals divided by τ, in each interval. The inference procedure requires two input parameters: (i) the sampling interval and (ii) the number of 2-MMPP sources. The duration of the sampling interval has a lower limit determined by the need to have at least 2^N arrival rates. Also, the duration should not be too long in order to allow capturing any burstiness behavior present in data. Our experiments indicate that a reasonable value for the sampling interval is one that allows a maximum of 100 arrivals per interval. The number of 2-MMPP sources is a tradeoff between the precision of the matching process and the computational complexity.

3.1. Approximation of the autocovariance

The first part of the procedure is the approximation of the empirical autocovariance of the rate process by a sum of N exponentials with real positive weights and negative real time constants. This is accomplished through a modified Prony algorithm[20] and the approximation is validated using the methods of Ref. 21. The Prony algorithm returns two vectors,

$$\vec{\alpha} = [\ \alpha_1 \quad \alpha_2 \quad ... \quad \alpha_N\ ] \qquad \vec{\beta} = [\ \beta_1 \quad \beta_2 \quad ... \quad \beta_N\ ]$$

3.2. Approximation of the distribution and parameter fitting

In this part of the fitting procedure we seek to define the MMPP parameters that best match the marginal distribution, given the constraints imposed by the approximation of the autocovariance described in previous section. The marginal distribution can be determined by the arrival rates of the generic MMPP $\vec{\lambda}$ and the steady-state probabilities of the superposed MMPP $\vec{\varphi}_j$, $j = 1, ..., N$. From equations (4) and (9) it results that $\vec{\lambda}$ is fully determined by $\vec{\varphi}_j$, $j = 1, ..., N$ and λ_Δ. Thus we will consider these two variables to be our working variables in the process of matching the marginal distribution. Matching the distribution requires the minimization of the error between the steady-state probabilities of the generic MMPP, $\vec{\pi}$, and the empirical steady-state probabilities, denoted by $\vec{p}\left(\vec{\lambda}\right)$. The empirical teady-state probabilities can be obtained from the empirical marginal distribution of the rate process $F(x)$ by

$$\vec{p}\left(\vec{\lambda}\right) = \begin{cases} F\left(\lambda_1\right), & i = 1 \\ F\left(\lambda_{i+1}\right) - F\left(\lambda_i\right), & i = 2, ..., 2^N \end{cases} \tag{10}$$

where $\lambda_{i+1} \geq \lambda_i$.

Variables $\vec{\varphi}_j$, $j = 1, ..., N$, and λ_Δ can be calculated through a minimization process, which is constrained by the mapping restrictions between the superposed and generic MMPP, defined by equations (8) and (9), and by the autocovariance approximation, defined by equation (4). It can be formulated as follows:

$$\min_{\vec{\varphi}_j, \lambda_\Delta} \left| \vec{p}\left(\vec{\lambda}\right) - \vec{\pi} \right|$$

subject to

$$\pi_i = \prod_{(j,k)\in E_i} \varphi_{jk}, \qquad i = 1, ..., 2^N$$

$$d_j = \pm\sqrt{\frac{\alpha_l}{\varphi_{j1}\left(1 - \varphi_{j1}\right)}}, \qquad j, l = 1, ..., N$$

$$\lambda_i = \lambda_\Delta + \sum_{j=1}^{N} \left(d_j \left(1 - \bmod\left(\left\lceil \frac{i}{2^{N-j}} \right\rceil, 2 \right) \right) \right), \qquad i = 1, ..., 2^N$$

$$\lambda_i \geq 0, \qquad i = 1, ..., 2^N$$

We have defined an iterative algorithm for calculating variables $\vec{\varphi}$, $j = 1, \ldots, N$, and λ_Δ, which is represented in the flow diagram of figure 2. The algorithm starts by determining initial values for the arrival rates and for the corresponding empirical probabilities. It then calculates the steady-state probabilities of the superposed MMPP, under the constraints imposed by the mapping relations between the two models. According to (8), these steady-state probabilities uniquely determine new steady-state probabilities for the generic MMPP. After this, the algorithm determines the arrival rate differences of the superposed model and adjusts the mean arrival rate. This gives rise to new arrival rates for the generic MMPP, and the process can be iterated from this point. A detailed description of the algorithm follows.

Step 1: Define initial $\vec{\lambda}$ such that $d_j = \lambda_{max} - \lambda_{min}/N, \forall j$ and set $\lambda_\Delta = \lambda_{min}$; λ_{max} and λ_{min} are the maximum and minimum empirical rates, respectively. This option assures that the initial $\vec{\lambda}$ verifies the restrictions imposed by the mapping between the superposed and generic MMPPs, as defined by equation (9).

Step 2: Calculate $\vec{p}\left(\vec{\lambda}\right)$, from (10).

Step 3: Calculate $\vec{\varphi}_j$, $j = 1, ..., N$. From (8), the empirical steady-state probabilities of the generic MMPP can be related to the steady-state probabilities of the individual 2-MMPPs by the following set of non-linear equations:

$$p_i = \prod_{(j,k)\in E_i} \varphi_{jk}, \qquad i = 1, 2, ..., 2^N \tag{11}$$

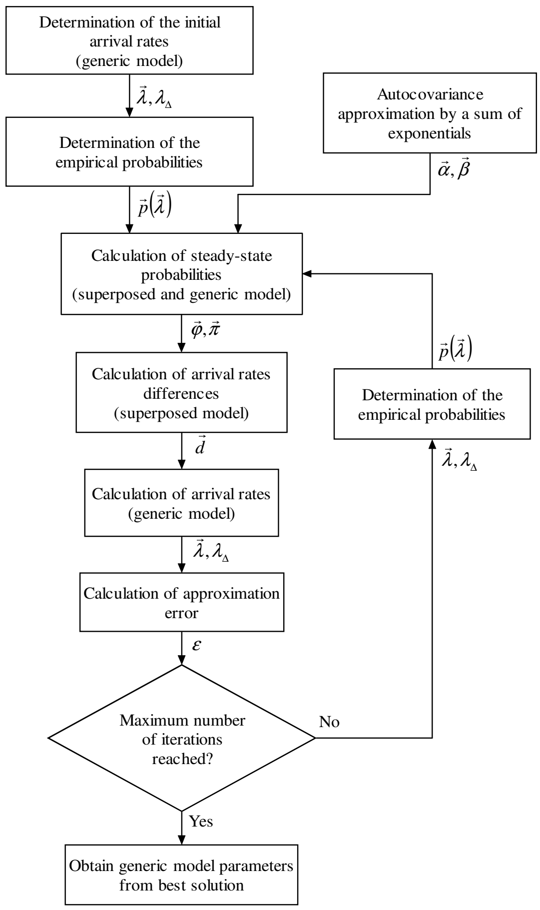

Figure 2. Flow diagram of fitting procedure.

where $\varphi_{j2} = 1 - \varphi_{j1}$, $j = 1, 2, ..., N$. There may not be a solution for obtaining the φ_{j1}, $j = 1, ..., N$, from the p_i, $i = 1, ..., 2^N$. For example, in the case of $N = 2$, the following set of equations result: $p_1 = \varphi_{11}\varphi_{21}$; $p_2 = \varphi_{11}(1-\varphi_{21})$; $p_3 = (1-\varphi_{11})\varphi_{21}$; $p_4 = (1-\varphi_{11})(1-\varphi_{21})$. Hence, there will only be a solution for φ_{11} and φ_{21} if $p_1p_4 = p_2p_3$. In the general case, there are 2^N conditions and only N independent variables. Thus, there may not be a solution that satisfies all conditions. We have adopted the following approximation

$$\varphi_{j1} = \sum_{k=1}^{2^{j-1}} \sum_{l=1}^{2^{N-j}} p_{2^{N-j+1}(k-1)+l}; \;\; \varphi_{j2} = 1 - \varphi_{j1}, \qquad i = 1, ..., N$$

which is based on N conditions derived from the 2^N available in (11). Our numerical experiments indicate that this approximation achieves good results. This step ends by calculating $\vec{\pi}$ from (8).

Step 4: Calculate $\vec{d}$ from

$$d_j = \sqrt{\frac{\alpha_j}{\varphi_{j1}\left(1 - \varphi_{j1}\right)}}, \qquad j = 1, ..., N$$

which resorts to (4). However, we restricted dj to be positive and adopted the mapping between autocovariance exponential terms and individual 2-MMPPs that corresponds to set $j = l$ in equation (4). Note also that the minimum value of d_j is $2\sqrt{\alpha_j}$, which restricts the fitting of the marginal distribution.

Step 5: Calculate new $\vec{\lambda}$ from $\vec{d}$ of step 4 and previous λ_Δ (from step 1 or step 6), using (7).

Step 6: Adjust the mean by calculating a new value for λ_Δ. This will be given by

$$\lambda_\Delta = \min \vec{\lambda} + \bar{\lambda}_e - \bar{\lambda}$$

where $\bar{\lambda}_e$ is the empirical mean arrival rate and $\bar{\lambda}$ is the mean arrival rate of the MMPP, given by $\vec{\pi} \bullet \vec{\lambda}$.

Step 7: Calculate $\vec{\lambda}$ from $\vec{d}$ of step 4 and λ_Δ of step 6, using equation (9).

Step 8: Calculate mean square error

$$\varepsilon = \frac{1}{2^N} \sum_{i=1}^{2^N} \left(p_i\left(\lambda_i\right) - \pi_i\right)^2$$

Step 9: Return to step 2.

It is not possible to assure convergence of this iterative algorithm. However, our numerical studies indicate that iterating a few times (less than 10) and saving the best result according to the mean square error criterion, defined in step 8, gives excellent results in most cases.

3.3. Calculation of the final MMPP parameters

The fitting procedure finishes by calculating the remaining parameters of the generic MMPP, and by constructing matrices $\mathbf{\Lambda}$ and $\mathbf{Q}$. We first calculate the transition rates of the individual 2-MMPPs from φ_{j1}, β_j, $j = 1, ..., N$, calculated previously and

$$r_{j1} = (1 - \varphi_{j1})\beta_j \qquad r_{j2} = \varphi_{j1}\beta_j$$

We then construct matrix $\mathbf{Q}$ from (1) and (3). The construction of $\mathbf{\Lambda}$ requires the calculation of vectors $\vec{\gamma}_j$, $j = 1, ..., N$. This can be done from (7) and (9), using the values of $\vec{d}$ and λ_Δ calculated previously. These equations allow more than one solution, but have to verify $\sum_{j=1}^{N} \gamma_{j1} = \lambda_c$, where λ_c represents one arrival rate of the generic MMPP. We have assumed $\lambda_c = \lambda_\Delta$ and $\gamma_{j1} = \lambda_\Delta / N$, $j = 1, ..., N$, which implies $\gamma_{j2} = \gamma_{j1} + d_j$. After obtaining the $\mathbf{\Lambda}_j$ we can then construct matrix $\mathbf{\Lambda}$ from (2).

4. NUMERICAL RESULTS

We apply our fitting procedure to several traffic traces: the publicly available Bellcore LAN traces[9] and Internet traces measured at our institution. The traces measured at our institution are representative of Internet traffic produced within a large educational and research environment. We assess the inference procedure by comparing the marginal density, the marginal distribution and the autocovariance of the original data traces and of simulated traces obtained from the fitted MMPPs. We analyze the presence of LRD behavior, in both original and fitted data traces, using the method described in Ref. 19. This method resorts to the so-called Logscale Diagram which consists in the graph of y_j against j, together with confidence intervals about the y_j, where y_j is a function of the wavelet discrete transform coefficients at scale j. Traffic is said to be LRD if, within the limits of the confidence intervals, the y_j fall on a straight line, in a range of scales from some initial value j_1 up to the largest one present in data. We also analyze the queuing behavior by comparing the packet loss ratio, obtained through trace-driven simulation, using again the original data traces and the simulated traces obtained from the fitted MMPP. To calculate the packet loss ratio (Bellcore and our institution) we assume a fixed packet size equal to mean packet size. The sampling interval for building the rate process was 0.1 seconds for all traces.

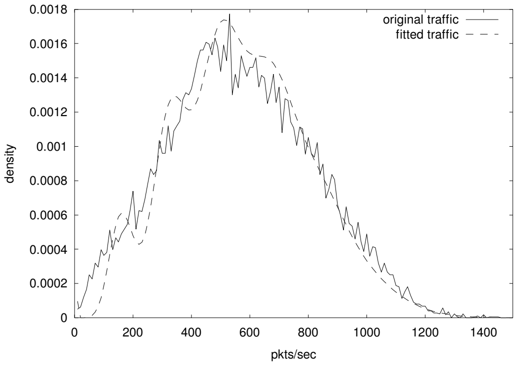

Figure 3. Marginal density, pOct.TL.

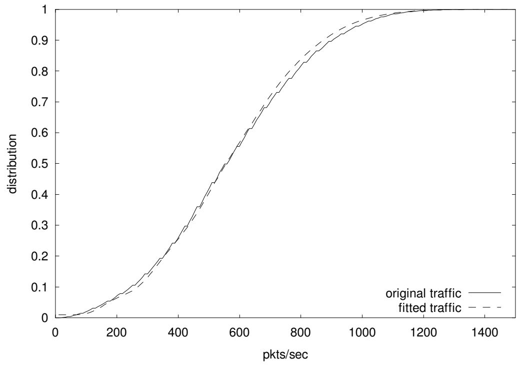

Figure 4. Marginal distribution, pOct.TL.

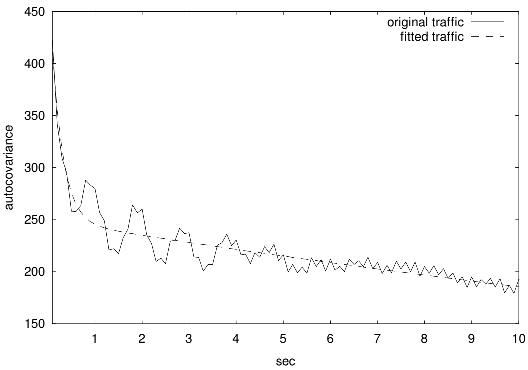

Figure 5. Autocovariance, pOct.TL.

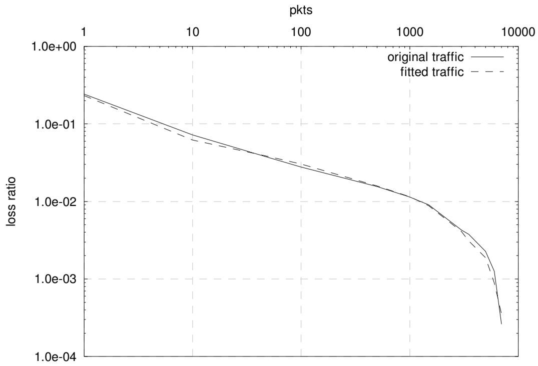

Figure 6. Packet loss ratio, pOct.TL.

4.1. Bellcore traces

The fitting procedure was applied to Bellcore/Telcordia traces pOct.TL and pAug.TL, both with 1 million samples. The pOct.TL trace was fitted to a 128-MMPP model (corresponding to the superposition of seven 2-MMPPs) and the pAug.TL trace was fitted to a 64-MMPP model (corresponding to the superposition of six 2-MMPPs).

Figure 3 and figure 4show the fitting results for the first-order statistics (pOct.TL trace). In this case, the fitting was performed with a very small approximation error.

From figure 5 it can be seen that the autocovariance of the original data has an oscillatory behavior. The fitting procedure captures only the average behavior of the autocorrelation but, as shown latter, this is sufficient for assessing queuing performance. Figure 7 and figure 8 show that both traces exhibit LRD, since the y_j values are aligned between octave 5 and octave 11, the highest octave present in data. The estimated Hurst parameters are $\widehat{H} = 0.847$ for the original data trace and $\widehat{H} = 0.824$ for the fitted data trace. To analyze the queuing behavior we considered a queue with a service rate of 3.85 Mb/s. The buffer size was varied from 1 to 7000 packets. The average packet size for this trace is 638 bytes. Figure 6 shows that the loss ratios of original and fitted traces are almost coincident for all buffer sizes. This confirms the good matching obtained in both first and second order statistics.

As shown in figure 9 and in figure 10, the fitting results of the first-order statistics for the pAug.TL trace were not so good as for the pOct.TL trace. Essentially, the fitting procedure did not produce an accurate match of the density tail. Also, the shape at low rates of the fitted density shows lower probabilities. This is due to the restrictions imposed by the autocovariance fitting. The results of the autocovariance matching were similar to the ones obtained with the pOct.TL trace, as seen in figure 11.

Again, as in previous trace, both original and fitted data traces exhibit LRD. In figure 13 and figure 14, it can be seen that the y_j values are aligned between octave 6 and the highest octave present in data. The estimated Hurst parameters are $\widehat{H} = 0.814$ for the original data trace and $\widehat{H} = 0.895$ for the fitted data trace.

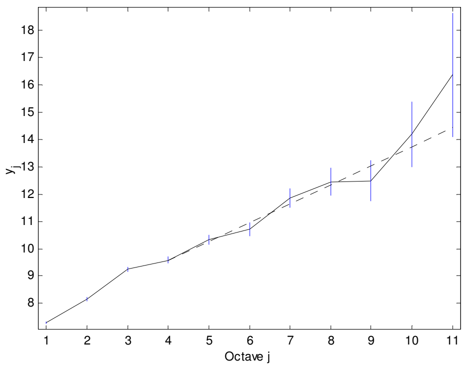

Figure 7. LRD wavelet analysis, pOct.TL.

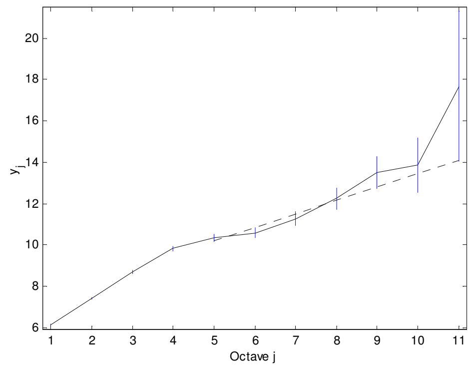

Figure 8. LRD wavelet analysis, pOct.TL fitted traffic.

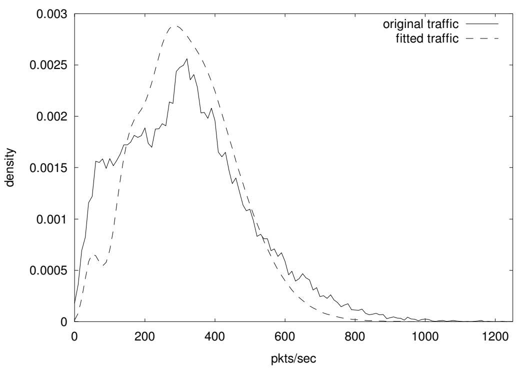

Figure 9. Marginal density, pAug.TL.

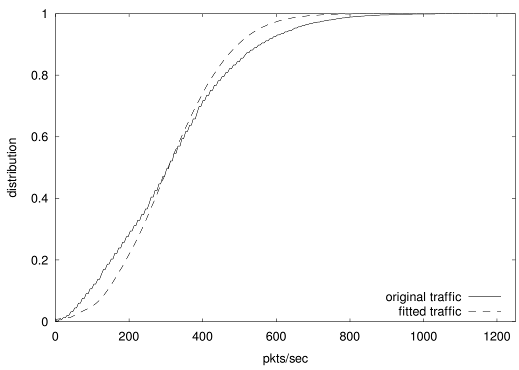

Figure 10. Marginal distribution, pAug.TL.

The queuing behavior reflects the worse matching of the first-order statistics, when compared with the pOct.TL case. As seen in figure 12, the packet loss ratio of the original traffic is higher than that of the fitted traffic, for all buffer sizes. This can be attributed to the lower tail of the fitted density. In this case the simulations were performed with a service rate of 1.14 Mb/s and the buffer size was varied from 1 to 3500 packets. The average packet size for this trace is 434 bytes. We note that a similar relative performance of fitting pOct.TL and pAug.TL traces was obtained in Ref. 1.

An interesting observation regarding these results is that although a similar LRD behavior was obtained for both original and fitted data traces, there is a clear deviation in terms of queuing behavior for the same traces. This reinforces our initial claim that a good matching of the marginal distribution is required for accurate prediction of the packet loss ratio.

4.2. Internet traces measured at our institution

The inference procedure was applied to a traffic trace measured at our institution (the University of Aveiro - UA). This is a trace with approximately 500000 packets, measured in December 2000 at the Internet Access Point of our campus network. We fitted the trace to a 16-MMPP model (corresponding to the superposition of four 2-MMPPs). As with the pOct.TL Bellcore traces the fitted results were very good, as can be seen from figure 15, figure 16, figure 17 and figure 18. Figure 19 and figure 20 demonstrate the existence of LRD in both traces. The estimated Hurst parameters are $\widehat{H} = 0.808$ for the original data trace and $\widehat{H} = 0.871$. The service rate was 4 Mb/s and the buffer size was varied from 1 to 4000 packets.The average packet size of this trace is 610 bytes.

5. RELATED WORK

In this section, we restrict our attention to fitting procedures for MMPPs. We start by noting that most procedures only apply to 2-MMPPs[5,16–18] . While 2-MMPPs can capture traffic burstiness, the number of states is in general

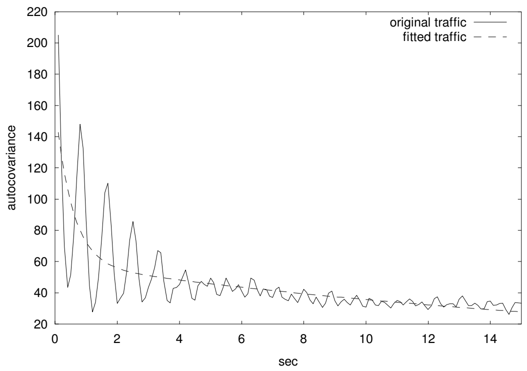

Figure 11. Autocovariance, pAug.TL.

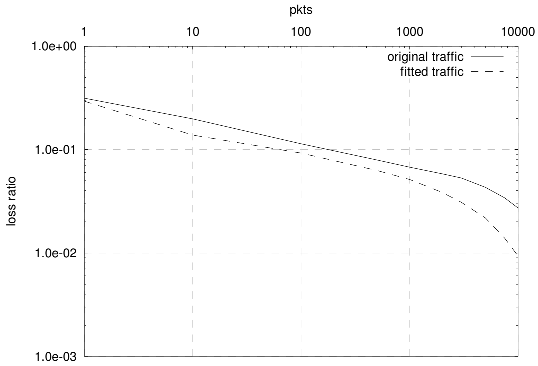

Figure 12. Packet loss ratio, pAug.TL.

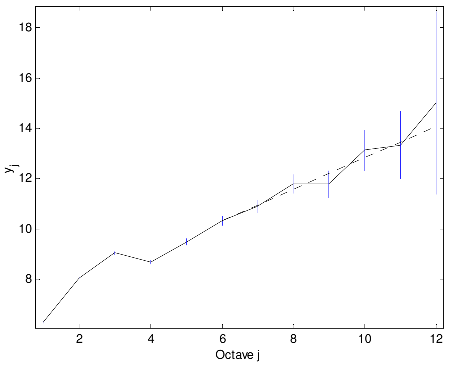

Figure 13. LRD wavelet analysis, pAug.TL.

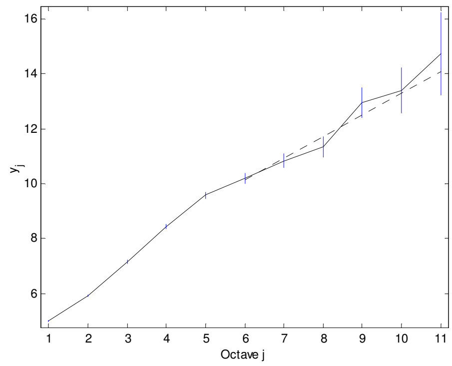

Figure 14. LRD wavelet analysis, pAug.TL fitted traffic.

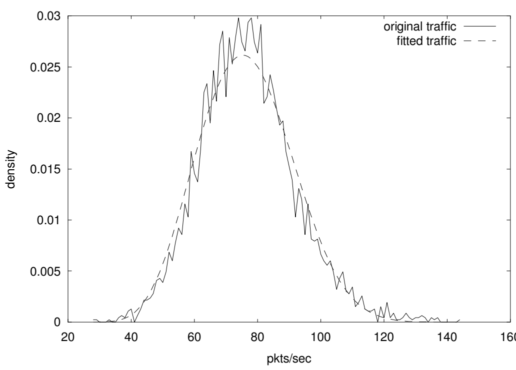

Figure 15. Marginal density, UA trace.

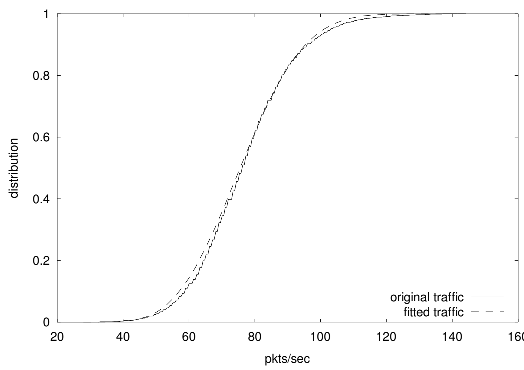

Figure 16. Marginal distribution, UA trace.

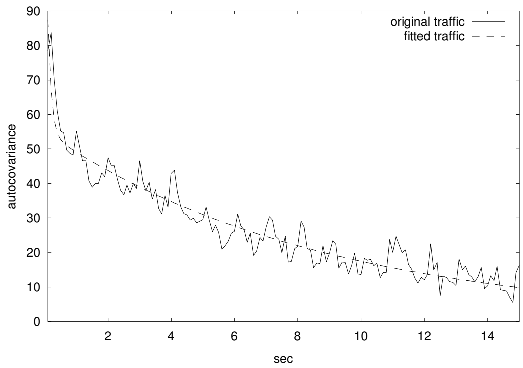

Figure 17. Autocovariance, UA trace.

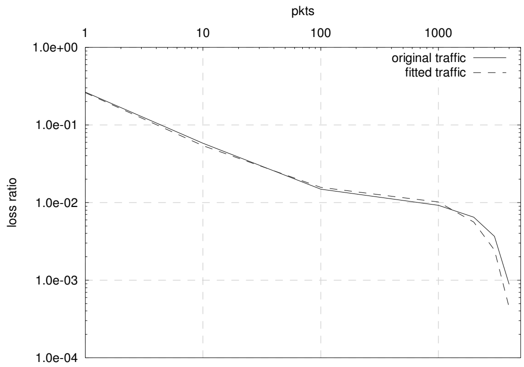

Figure 18. Packet loss ratio, UA trace.

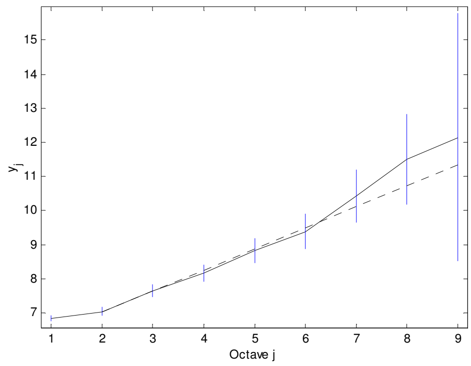

Figure 19. LRD wavelet analysis, UA trace.

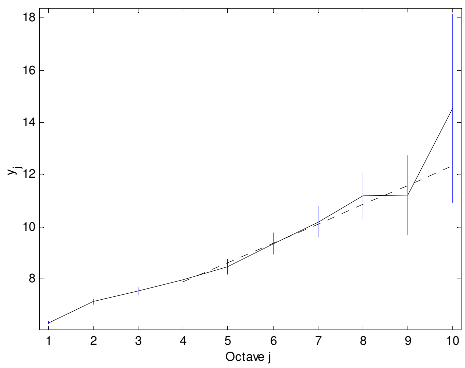

Figure 20. LRD wavelet analysis, UA trace fitted traffic.

not enough to provide a good match of the marginal distribution when the traffic shows variability on a wide range of arrival rates.

Skelly *et al.*[22] propose a method for estimating the parameters of a generic MMPP that only matches the first-order statistics: the Poisson arrival rates are inferred from the empirical probability function and the state transition rates from a direct measurement of the observed trace. The main limitation is that second-order characteristics of the traffic, such as the autocovariance, are not taken into account.

Andersen and Nielsen[1] use 2-MMPPs to model several time-scales of the autocovariance function. Each of the time scales is fitted to an exponential function, resulting in a model that corresponds to the superposition of several 2-MMPPs, as in our case. However, the fitting of the first-order statistics is very poor, since only the mean is matched.

The work by Li and Hwang[15] is closely related to ours, in that it also matches both the autocovariance and the marginal distribution. The fitting procedure applies to CMPPs, which are a special case of MMPPs where the steady-state probabilities are the same for all states. As opposed to ours, the fitting procedure is able to capture pseudoperiodic components present in data, since the infinitesimal generator matrix of a CMPP can have complex eigenvalues. However, our experiments indicate that the pseudoperiodic components have a small importance in what concerns queuing performance. Moreover, there is less flexibility in adjusting the marginal distribution, since CMPP states are equiprobable and the CMPP fitting procedure matches separately the autocovariance and the marginal distribution.

6. CONCLUSIONS

This paper proposed a fitting procedure for Markov Modulated Poisson Processes (MMPPs) that matches both the autocovariance and marginal distribution functions of the rate process. The MMPP is constrained to be decomposable into a superposition of N two-state MMPPs (2-MMPPs). The fitting procedure works simultaneously in the state space of N individual 2-MMPPs and on the state space of its equivalent MMPP with 2^N states. The procedure starts by approximating the autocovariance to a weighted sum of exponentials. It then fits the MMPP parameters in order to match the distribution, within the constraints imposed by the autocovariance matching and by the mapping relations between the two equivalent models. Our numerical results, which include fitting traffic traces that exhibit long-range dependence, show that the fitting procedure matches closely both the autocovariance and the marginal distribution, and achieves a very good performance in terms of queuing behavior, as assessed by the packet loss ratio suffered by original data and fitted traces. The results illustrate that MMPP models, although not being intrinsically long-range dependent, can capture this type of behavior for limited time scales. We also give an example that demonstrates the importance of fitting the complete marginal distribution, as opposed to matching only the mean, for accurate prediction of the packet loss ratio.

ACKNOWLEDGMENTS

This work was part of project 34826/99 SCALE "Statistical Characterization of Telecommunications Traffic", funded by Fundação para a Ciência e Tecnologia, Portugal. We also acknowledge the support of Portugal Telecom Inovação. P. Salvador wishes to thank Fundação para a Ciência e a Tecnologia, Portugal, for support under grant BD/19781/99.

REFERENCES

1. A. Andersen and B. Nielsen, "A markovian approach for modeling packet traffic with long-range dependence," *IEEE Journal on Selected Areas in Communications* **16**, pp. 719–732, Jun 1998.
2. H. Che and S. Li, "Fast algorithms for measurement-based traffic modeling," *IEEE Journal on Selected Areas in Communications* **16**, Jun 1998.
3. L. Deng and J. Mark, "Parameter estimation for markov modulated poisson processes via the em algorithm with time discretization," *Tel. Systems* (1), pp. 321–338, 1993.
4. M. Grossglauser and J. Bolot, "On the relevance of long-range dependence in network traffic," *IEEE/ACM Transactions on Networking* **7**, pp. 629–640, Oct 1999.
5. S. Kang and D. Sung, "Two-state mmpp modelling of atm superposed traffic streams based on the characterisation of correlated interarrival times," *IEEE GLOBECOM'95* , pp. 1422–1426, Nov 1995.
6. S. Robert and J. L. Boudec, "A modulated markov process for self-similar traffic," *Proceedings of Saarbrucken Schloss Dagstuhl, Germany* , Jan 1995.
7. S. Robert and J. L. Boudec, "New models for self-similar traffic," *Performance Evaluation* **30**, Jul 1997.
8. M. Crovella and A. Bestavros, "Self-similarity in world wide web traffic: Evidence and possible causes," *IEEE/ACM Transactions on Networking* **5**, pp. 835–846, Dec 1997.
9. W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking* **2**, Feb 1994.
10. V. Paxson and S. Floyd, "Wide-area traffic: The failure of poisson modeling," *IEEE/ACM Transactions on Networking* **3**, pp. 226–244, Jun 1995.
11. 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* **26**, pp. 3–14, Oct 1996.
12. W. Willinger, M. Taqqu, R. Sherman, and D. Wilson, "Self-similarity through high-variability: Statistical analysis of ethernet lan traffic at the source level," *IEEE/ACM Transactions on Networking* **5**, pp. 71–86, Feb 1997.
13. W. Willinger, V. Paxson, and M. Taqqu, *Self-similarity and Heavy Tails: Structural Modeling of Network Traffic*, A Practical Guide to Heavy Tails: Statistical Techniques and Applications, Birkhauser, 1998.
14. B. Hajek and L. He, "On variations of queue response for inputs with the same mean and autocorrelation function," *IEEE/ACM Transactions on Networking* **6**(5), pp. 588–598, 1998.
15. 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.
16. R. Grünenfelder and S. Robert, "Which arrival law parameters are decisive for queueing system performance," in *ITC 14*, 1994.
17. K. Meier-Hellstern, "A fitting algorithm for markov-modulated poisson process having two arrival rates," *European Journal of Operational Research* **29**, 1987.

18. C. Nunes and A. Pacheco, "Parametric estimation in MMPP(2) using time discretization," *Proceedings of the 2nd Internation Symposium on Semi-Markov Models: Theory and Applications* , Dec 1998.
19. D. Veitch and P. Abry, "A wavelet based joint estimator for the parameters of LRD," *"Special issue on Multiscale Statistical Signal Analysis and its Applications", IEEE Transactions on Information Theory* .
20. M. Osborne and G. Smyth, "A modified prony algorithm for fitting sums of exponential functions," *SIAM J. Sci. Statist. Comput.* **16**, pp. 119–138, 1995.
21. A. Feldmann and W. Whitt, "Fitting mixtures of exponentials to long-tail distributions to analyze network," *Performance Evaluation* **31**(3-4), pp. 245–279, 1998.
22. 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.