Broadband teletraffic characterization using correlated interarrival time Poisson process (CIPP)

R. Manivasakan, U. B. Desai and Abhay Karandikar
Department of Electrical Engineering, Indian Institute of Technology, Bombay 400 076, India
e-mail: {mani, ubdesai, karandi}@ee.iitb.ernet.in

Abstract

In this paper, we propose a new traffic model (with correlation in interarrivals) for composite arrival process of packets in broadband networks and individual source modeling. The traffic model involves a generalization of Poisson process which is referred to as correlated interarrival time Poisson process (CIPP). The CIPP, a stationary counting process, is parametrized by a correlation parameter ρ which represents the degree of correlation in adjacent interarrivals in addition to λ, the intensity of the process. It is shown by simulations that this arrival process models well the real-time traffic source, in particular the video source. The burstiness measure corresponding to this process is evaluated as defined in Saito *et al.* (*IEEJ*, 1991, SAC-9, 359–367). It turns out that although CIPP is *not* strictly a self-similar model, it *does* exhibit self-similar behaviour over a range of time scales. We explore, by simulations, the relation between the positive correlation in interarrivals and the estimated local Hurst parameter.

Keywords: CIPP, interarrivals, burstiness measure, self-similarity, teletraffic modeling.

1. Introduction

Packet-switched networks can be viewed as networks of queues, the most fundamental component being a single server queue. Accordingly, the performance analysis of a single server queue has been a focal point of communication research for several years. One of the important questions often raised by network planners and designers is: Is an accurate and useful traffic model in the form of a simple stochastic process with minimum number of parameters available, which, when fed to a single server queue, gives the same performance as that of a real traffic stream? Such a traffic model will be very useful in network design tools or in tools supporting real-time traffic management. Despite an exhaustive research during the last two decades,[1] there is no consensus on such a model. Nevertheless, significant progress has been made over the years in understanding the characteristics of such a traffic. We intend here to use the fact that the interarrival correlation plays a major role in imitating the self-similar behaviour in increments of the corresponding counting process, and wish to make a contribution towards a consensus on *this* key aspect of broadband teletraffic modeling.

In literature, the following models for cell[a]-level traffic have been reported: Poisson, Bernoulli, and D-MAP.[2] Researchers have used the following analytical models for traffic at burst level: on–off processes, non-renewal Markov chain (MC),[3, 4] interrupted Poisson process

---

[a] *cell* denotes fixed length packet.

(IPP),[5] Markov-modulated Poisson process (MMPP),[6, 7] and gaussian models.[8] The above models are essentially short-range-dependent (SRD) and rely heavily on modulating process for their correlation in increments.

The discovery of long-range dependence (LRD) in Ethernet traffic in Bellcore[9] is considered to be one of the most significant contributions in the area of teletraffic modeling. Further studies confirm that metropolitan area network (MAN) traffic,[10] wide-area network (WAN) traffic, and variable bit rate (VBR) video traffic[2] also exhibit LRD (or self-similar or fractal) characteristics. Moreover, it has been firmly concluded by teletraffic researchers that positive correlation is dominant in real-world Ethernet traces.[11] (Paxson and Floyd[11] state: "If we require fixed rates only over 10-minute intervals, then SMTP and FTPDATA burst arrivals are not terribly far from Poisson, though neither is statistically consistent with Poisson arrivals, and consecutive SMTP interarrival times show *consistent positive correlation*.") Also, it has been empirically established that positive correlation degrades the queueing performance.[12] It is well known that self-similar processes with $0.5 \leq H < 1.0$ (where H is the Hurst parameter*) have positively correlated increments.

In this paper, we propose a model called correlated interarrival time Poisson process (CIPP). In CIPP, the interarrivals form a first-order Markov sequence. We will see in later sections that the CIPP does exhibit self-similarity over a range of time scales of interest. It is also analytically tractable (see, for instance, Manivasakan *et al.*[13]). (Note that the empirical models based on self-similarity and long-range dependence proposed in the literature[11, 14-16] are *not* analytically tractable except that they have some bounds on their queueing performance.) In addition, CIPP has positive correlation in increments. Manivasakan *et al.*[13] have shown *analytically* that in CIPP positive correlation degrades the queueing performance.

In this paper, we will restrict our discussion to modeling broadband teletraffic by CIPP, leaving its implications on queueing theory to a separate paper. We evaluate some properties of CIPP which are of significance to teletraffic modeling like the index of dispersion for counts (IDC), index of dispersion for intervals (IDI) and burstiness measure.[1] We examine the self-similar nature of the CIPP process using wavelet analysis. The paper is organized as follows: In Section 2, we describe the CIPP process. In Section 3, fitting the CIPP model to real-world traffic trace and self-similar property of CIPP are presented. Section 4 concludes the paper.

2. CIPP

2.1. *Mathematical background and properties*

We need to define a counting process $N(t)$ which retains as far as possible all the nice properties of the Poisson process, but has correlated interarrivals. Typically, we would like to have interarrivals being exponentially distributed apart from being correlated. Motivated by the analytical simplicity of the method of generating stationary correlated exponential sequence (first-order exponential autoregressive process (EAR(1))) developed by Gaver and Lewis,[17] we present the following formalism to define the CIPP.

*Hurst parameter quantifies the degree of self-similarity.

Consider the arrivals occurring in time on the interval $(0, \infty]$. For $t > 0$, let $N(t)$ be the number of arrivals that have occurred in the half-open interval $(0, t]$. Consequently, $N(t)$ and $N(t + h) - N(t)$ $(h > 0)$ assume only non-negative integer values. We now state the following axioms.

Axioms:

1. Since we begin counting arrivals at time 0, we define $N(0) = 0$.

2. Correlation structure for interarrivals: Let X_n be the interarrival time between the *nth* arrival instant (T_n) and $(n - 1)th$ arrival instant (T_{n-1}). Then,

$$X_{n+1} = \rho X_n + \epsilon_{n+1} \qquad 0 \le \rho < 1 \qquad n = 1, 2, 3... \qquad (1)$$

   where $\left\{\epsilon_n\right\}_{n=2}^{\infty}$ is an iid sequence, with $\epsilon_n, n > 1$, being a product of Bernoulli random variable (B) with parameter ρ and exponential random variable (V) with parameter λ. B is statistically independent of V. $\{X_n\}$ forms a stationary sequence with the exponential distribution characterized by parameter λ. Observe that $X_1 = T_1 - T_0$ where T_0 is the 0th arrival instant.

3. The counting process $N(t)$ is strictly stationary, that is, for any $r = 1, 2,...$ the joint distribution of $\{N(h + t_{u1}) - N(t_{l1}), N(h + t_{u2}) - N(t_{l2}),..., N(h + t_{ur}) - N(t_{lr})\}$ is independent of h, for any $h > 0$ and $0 < t_{l1} < t_{u1} < t_{l2} < t_{u2} < ... < t_{lr} < t_{ur} < \infty$.

Definition: A counting process $\{N(t), t > 0\}$ satisfying Axioms 1 through 3 is called a CIPP with parameter λ and ρ.

Remark 1: Axiom 2 is *more specific* and the following properties of the counting process $N(t)$ directly follow from it.

1. For any $t > 0$, $0 < Pr\{N(t) > 0\} < 1$. It means that in any interval (no matter however small) there is a non-zero probability that an arrival will occur.

2. For any $t \ge 0$,

$$\lim_{h \to 0} \frac{Pr\{N(t+h) - N(t) \ge 2\}}{Pr\{N(t+h) - N(t) = 1\}} = 0. \qquad (2)$$

   In other words, in sufficiently small intervals, at most one arrival can occur, i.e. it is not possible for arrivals to happen simultaneously. The process is *orderly*, in the sense of Daley *et al.*[18]

Remark 2: The main difference between the axiomatic definition of a Poisson process and CIPP is the addition of Axiom 2 and the relaxation of the independence assumption in Axiom 3.

Let $P_n(t) \underset{=}{\Delta} Pr\{N(t) = n\}$ be the probability that the number of arrivals in the interval $(0, t]$ is n.

Lemma 1. *CIPP with parameter λ and ρ has the distribution*

$$P_n(t) = \frac{\left(1 + \rho - (1-\rho)\alpha_n\right)}{\times \sum_{j=0}^n \frac{\alpha_j \prod_{k=0}^n \left[\alpha_j - \rho\alpha_k\right]}{\left\{\rho(\alpha_j + 1) - \alpha_n + 1\right\}(\alpha_j - \alpha_n + 1)(\alpha_j - \rho\alpha_n)\prod_{k=0,k\neq j}^n \left[\alpha_j - \alpha_k\right]}} \, e^{-\frac{\lambda t}{\alpha_j}} \quad 0 \leq \rho < 1 \tag{3}$$

for $n = 0, 1, 2,\dots$, $\lambda > 0$, $t \geq 0$, and $\alpha_j = \sum_{i=0}^j \rho^i$.

Proof: See Appendix I for proof.

It is worth noting that the mean of the process, $\mathbf{E}\{N(t)\}$, can be shown to be λt.[19]

Remark 3: The prime reason for assuming (1) for the interarrival sequence is to mimic the self-similar behaviour in increments for the corresponding counting process. Note that to mimic a self-similar process with $0.5 < H < 1$ it is necessary that the increments should have positive correlations.[12] The reason for this particular correlation is that (1) introduces strong positive correlation than other models based on moving average structure like EMA(1) proposed by Lawrance and Lewis[20] and the one proposed by Finch.[21] We will see in later sections[22] that, as ρ increases, the estimated local Hurst parameter (denoted by $\hat{H}$) increases. Secondly, the correlation structure (1) allows one to derive a closed-form expression for $P_n(t)$ with reasonable simplicity which is not possible with the other aforementioned models. Thirdly, the correlation structure (1) is analytically simple which allows one to develop the corresponding queueing theory.[13] Finally, the sequence $\{X_n\}$ is obtained as an additive random linear combination of random variables and is thus easy to simulate on the computer.

Remark 4: The distribution (3) is count-stationary in the sense of Lawrance and Lewis.[20] An interval-stationary version (again in the sense of Lawrance and Lewis) can also be derived. Here, we give only the result.

$$P_n^I(t) = \left(1 + \rho - (1-\rho)\alpha_n\right) \tag{4}$$

Remark 5: For both the forms (3) and (4), as the limit $\rho \to 0$, we obtain the Poisson case. Thus, the Poisson distribution is approached continuously from the CIPP distribution.

Remark 6: We contend here that the introduction of correlation in the interarrival sequence is more appropriate in the context of modeling broadband teletraffic rather than to introduce correlation in the count sequence (as done in Heyman *et al.*[23] and Xu *et al.*[24] for VBR modeling) or to introduce correlation by the modulating Markov chain as in SRD models. The introduction of correlation in the interarrival sequence in the context of broadband teletraffic modeling is a new concept and is the main contribution of this paper. We defend our claim that the interarrival correlation is more appropriate. The correlations in interarrivals *necessarily* imply the correlations in the number of counts; however, the converse of this statement is *not* necessarily true. A good example would be a renewal process. In this context, we would like to mention that the discrete version of (1), the discrete autoregressive model of order 1 (DAR(1)) is used in Heyman *et al.*[23] and Xu *et al.*[24] to model the *frame size* of the VBR traffic. Needless to say that the correlation here is in the count sequence.

Remark 7: The number of parameters needed to exhibit the self-similar behaviour is minimal in CIPP as compared to the conventional SRD models like MMPP, DBMAP, ON–OFF, etc. For example, in MMPP, the number of parameters is N^2 for N state MMPP ($N^2 - N$ for infinitesimal generator and N for the intensity matrix Λ) while in CIPP it is always 2. The authors believe that the reason for this phenomena is that in CIPP, self-similar behaviour emanates from interarrival correlation rather than from correlation due to modulating Markov chain.

Remark 8: We will see in Section 3.3 that there exists an empirical relationship between the correlation parameter ρ and the degree of self-similarity (measured by estimated local Hurst parameter ($\hat{H}$) of the corresponding CIPP trace). Note that in sharp contrast, in conventional SRD models too many parameters decide the degree of (pseudo-)self-similarity. Moreover, one does not know which of them plays a prime role in exhibiting self-similar behaviour. The above feature of CIPP makes it more suited for practical purposes, when one needs to fit CIPP to a real-world data sequence (say a video sequence) of given Hurst parameter.

Remark 9: For many reasons, using an excessive number of parameters is undesirable, especially because it increases the uncertainty of the statistical inference and the parameters are difficult to interpret. In our case, CIPP has only two parameters, ρ and λ, and as we mentioned, both of them have a nice physical interpretation. λ denotes the arrival rate while ρ quantifies the correlation in interarrival times. While λ can be estimated from the mean of the process, ρ can be estimated from higher-order statistics. In this work, we use autocorrelation as a measure for fitting the model to capture the dependence structure in the real-world traffic.

Remark 10: CIPP is *not* a special case of *Neuts* process.[25]

Remark 11: We derived CIPP, by first generating correlated exponential sequence and then constructing the counting process over it. The whole exercise can be repeated in discrete time domain, by considering correlated geometric sequence generated by discrete AR(1) model.[26] But the result does not seem to yield a closed-form expression and will not be quoted any further in this paper.

3. Applications of CIPP to broadband teletraffic

In this section, we justify the CIPP model for modeling broadband teletraffic. In particular, we undertake experimental check on the 'goodness of fit'' of the CIPP for various video sequences[27] and for Bellcore Ethernet traffic datasets.[9]

3.1. Fitting the CIPP model to real-traffic data

In order to show that a proposed statistical process models well the real world data, we have to 'fit'' the model to the data. This fitting procedure usually involves the estimation of parameters of the model. One way to accomplish this is to estimate the first/higher-order statistics of data and to equate this to the corresponding moments (which *are* usually a function of parameters of the model). (Note that, here in CIPP case, one may use the elegant method of estimating the parameters, namely ρ and λ, as given in Gaver and Lewis.[17] Here, however, we
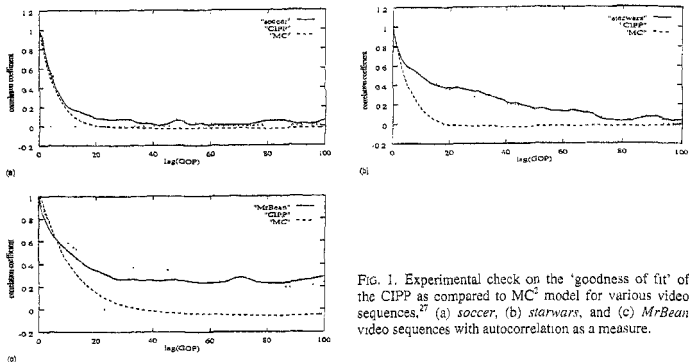
FIG. 1. Experimental check on the 'goodness of fit' of the CIPP as compared to MC² model for various video sequences.[27] (a) *soccer*, (b) *starwars*, and (c) *MrBean* video sequences with autocorrelation as a measure.

are *more* interested in capturing the LRD of the real-world data. Hence, we use correlation measure in some form or the other for our fitting procedure). We estimate the parameters of the model with a measure based on any of the following second-order properties: (1) autocorrelation of counts $\{u_T(i)\}$, (2) index of dispersion for intervals (IDI), and (3) index of dispersion for counts (IDC). Of the above three, CIPP has analytical expression only for IDI.

3.1.1. *Autocorrelation of counts* $\{u_T(i)\}$

Time-dependent statistics are crucial in the case of video traffic because correlations in the video streams can affect the performance of the statistical multiplexer (in a typical broadband network). In particular, a positive correlation in the traffic process (which is the case with CIPP streams) degrades the performance of statistical multiplexer.[12]

Let $\{u_T(i)\}_{i=1}^{\infty}$ denote the number of arrivals in ith slot on the positive real axis where each time slot is of equal length, T time units. Then the correlation coefficient is defined by,

$$\zeta_T(k) = \frac{E\{u_T(i)u_T(i+k)\} - E^2\{u_T(i)\}}{\mathrm{var}(u_T(i))}.$$

Note that for Poisson process $\zeta_T(k) = 0$ for all $k \neq 0$.

We now attempt to fit a CIPP model to MPEG-1-encoded video trace. GOP sizes (sum of frame sizes of one GOP) in the frame size trace from MPEG-1-encoded video sequence can be thought of as the number of (cell) arrivals (after segmentation of frame size trace) $\{u_T(i)\}$ in a GOP time slot (T_{GOP}). We fit a CIPP by using the following algorithm.

---

[b]This is just to avoid computational overflow while estimating the correlation coefficient of the data.

Fitting procedure

1. Segment[b] the bits in each frame into cells (of size 53 bytes).

2. Estimate the correlation coefficient of this count process of cells. This is the correlation coefficient $\varsigma_{T_{\text{GOP}}}^{\text{data}}(.)$ to which we intend to match the correlation coefficient of the CIPP model.

3. *Algorithm to estimate the parameters of CIPP*

   Set the rate λ_{opt} of CIPP equal to the ratio of total number of cells to the duration (total number of GOPs times T_{GOP}) of the video sequence.

   Algorithm to estimate ρ_{opt} of CIPP

   begin

       min = 99999999.9

       step = 0.0001

       for i = 1:9999

           ρ = step*i

           estimate the correlation coefficient $\varsigma_{T_{\text{GOP}}}^{\text{data}}(.)$ for CIPP by reasonably large realizations.

           compute the mean square error (P) between $\varsigma_{T_{\text{GOP}}}^{\text{CIPP}}(.)$ and $\varsigma_{T_{\text{GOP}}}^{\text{data}}(.)$

           if(min > P)

           min = P

           $\rho_{\text{opt}} = \rho$

           $\varsigma_{T_{\text{GOP}}}^{\text{fit}}(.) = \varsigma_{T_{\text{GOP}}}^{\text{CIPP}}(.)$

       end

       end

   end

The correlation coefficient $\varsigma_{T_{\text{GOP}}}^{\text{fit}}(.)$ is plotted along with the correlation coefficient $\varsigma_{T_{\text{GOP}}}^{\text{data}}(.)$. For comparison, we also plot the result by fitting a Markov chain (MC)[2] to the GOP size process (our procedure is very similar to Roberts *et al.*[2]). The number of states M in MC is taken to be G_{max}/σ_G, where G_{max} denotes the size of the largest GOP and σ_G the standard deviation of the GOP sizes. Thus, the size of quantization interval is σ_G. Note that it is conventional to consider the autocorrelation function (or equivalently correlation coefficient) of the GOP sizes since it is difficult to get a clear picture of the long-range correlations of video traffic stream from frame-level correlations.

We consider the frame-size trace from MPEG-1-encoded video sequence corresponding to video sequences with various H parameters (ranging from 0.77 to 0.96) like *soccer* ($H = 0.77$), *starwars* ($H = 0.85$) and *MrBeans* ($H = 0.96$)[27] (Fig. 1). One can notice that for
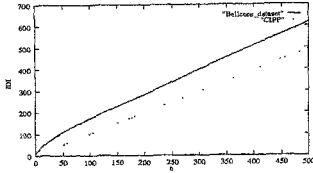
Fig. 2. IDI of 'Bellcore dataset' and the approximating
CIPP model.

soccer ($H = 0.77$), although both the models show almost the same good approximation quality, the MC model is more optimistic; however, for higher H-parameter video sequences like *starwars* and *MrBean*, the CIPP performs substantially better than the MC model.

3.1.2. *Discussion*

Note that the correlation function corresponding to MC model decays very fast which is a feature of Markov models. In contrast, the count sequence, $\left\{u_T(i)\right\}_1^\infty$ in CIPP case, may have more complex correlation structure (may not even be a first-order Markov chain), despite the fact that the interarrivals form the first-order Markov chain. We strongly believe that this is the reason for non-exponential decay of autocorrelation of counts for CIPP (and we will see in a later section that this also accounts for self-similar behaviour to a range of time scales.)

3.1.2. *IDI*

We use the variance of the sum of n random variables normalized by the factor $nE^2(X)$ as a measure of the variability of packet arrival processes.[28] The sequence of values

$$J_n \triangleq \frac{\text{var}(X_{i+1}+\ldots+X_{i+n})}{nE^2(X)}$$

with $n = 1, 2,\ldots$ is called IDI. It can be shown that for CIPP, J_n is given by

$$J_n = \frac{1}{n(1-\rho)^2}\left[n\left(1-\rho^2\right)+2\rho^{n+1}-2\rho\right].\qquad(5)$$

Then,

$$J \triangleq \lim_{n\to\infty} J_n = \frac{1+\rho}{1-\rho}.\qquad(6)$$

Since the CIPP has an analytical expression for J_n, the modeling procedure becomes simple. We have to just estimate ρ from the data using (5). Figure 2 shows the IDI plot for Bellcore August traffic dataset[9] and CIPP model with $\rho = 0.997$. CIPP does show reasonably good approximation to the Bellcore dataset.
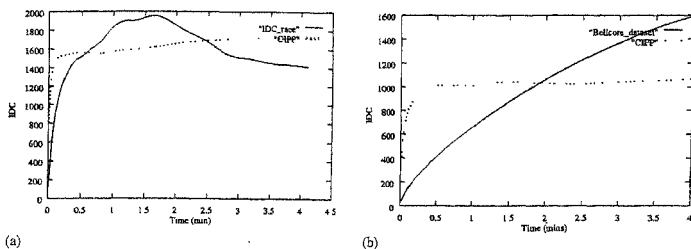
FIG. 3. IDC of (a) *race* video sequence and (b) *Bellcore* dataset and their corresponding approximating CIPP model.

3.1.3. *IDC*

For packet counts, we can define a function similar to the index of dispersion for intervals.[28] IDC at time t is the variance of the number of arrivals in an interval of length t divided by the mean number of arrivals in t:

$$I_t = \frac{\text{var}\big(N(t)\big)}{E\big(n(t)\big)} \tag{7}$$

where $N(t)$ indicates the number of arrivals in an interval of length t. IDC has been defined in this manner so that its value for a Poisson process is 1, for all t.

In estimating the IDC of the measured arrival processes, we will only consider the time at discrete, equally spaced instants, τ_i $(i \geq 0)$. Denoting the number of arrivals in $T = \tau_i - \tau_{i-1}$, by $u_T(i)$, we have,

$$I_n(T) = \frac{\text{var}\big(\sum_{i=1}^n u_T(i)\big)}{E\big(\sum_{i=1}^n u_T(i)\big)} \approx \frac{\text{var}\big(u_T(i)\big)}{E\big(u_T(i)\big)} \left[1 + 2\sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)\varsigma_T(j)\right] \tag{8}$$

where $\zeta_T(j)$ is the autocorrelation coefficient of the sequence $\{u_T(i)\}$ at lag j. Note that, in general, I_n will not be a constant for renewal processes where counts in the disjoint intervals
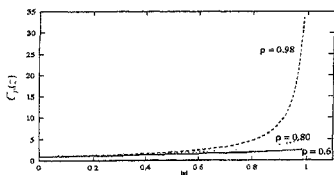

FIG. 4. $C_p(z)$ vs $|z|$ for CIPP.

are correlated (exception is the Poisson case). While we can estimate $\text{var}(N_t)/E(N_t)$, the representation on the RHS of (8) is valid only if the data are stationary.

Figure 3a shows the IDC plot for *race* video sequence and the corresponding CIPP match. CIPP again exhibits better approximation property. The corresponding plot for Bellcore data-

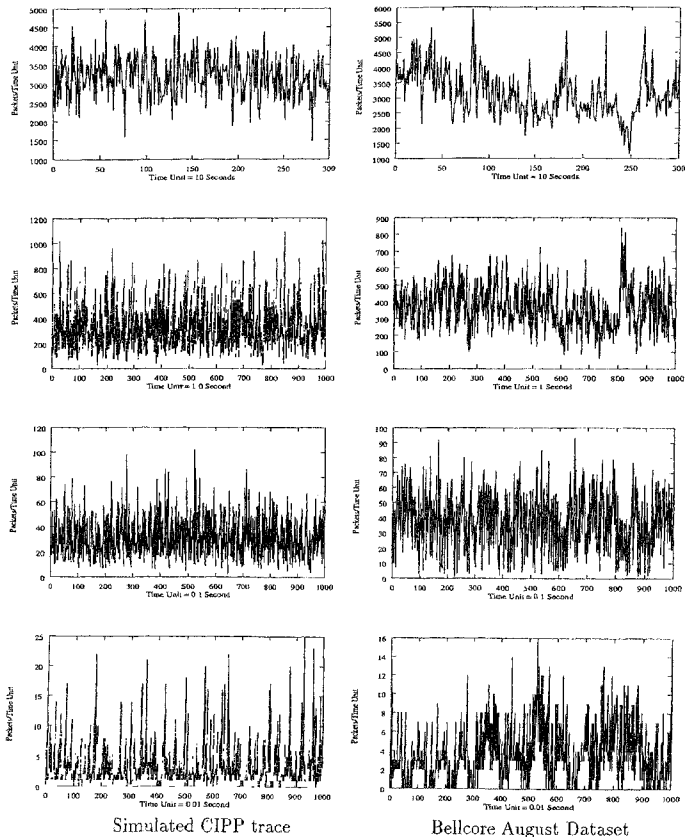Simulated CIPP trace   Bellcore August Dataset

FIG. 5. 'Visual' justification of self-similarity.

set is shown in Fig. 3b. Here the approximation is not very good, possibly because CIPP is a pseudo-self-similar process, as discussed in Section 3.3.

3.2. *Burstiness measure*

We consider the burstiness measure as defined in Saito *et al.*[1] We repeat it here for convenience

$$C_p(z) = \sum_{k=0}^{\infty} C_k z^k / \left(\lambda^{-1}\right)^2, \quad |z| < 1 \tag{9}$$

where λ^{-1} is the mean interarrival time. $C_k = E\{(X_1 - \lambda^{-1})(X_{k+1} - \lambda^{-1})\}$ denotes the correlation with lag k for the interarrival times and $X_1, X_2,...$ form a sequence of interarrival times. $C_p(z)$ includes the complete information for the second moment of interarrival times. For Poisson process, $C_p(z) = 1$ for all z. For a renewal arrival process, $C_p(z) =$ (squared coefficient of variation in the renewal process) for all z. For an MMPP, burstiness measure has been discussed in Saito *et al.*[1] For CIPP, it can be shown that the correlation with lag k of interarrival times (C_k) is given by $C_k = \rho^k$, $|\rho| < 1$. As a result, $C_p(z) = \frac{1}{1-\rho z}$. Figure 4 shows $C_p(z)$ versus z. One can see that CIPP does display burstiness for large ρ.

3.3. *Self-similar property of CIPP*

To compare the behaviour of CIPP with Bellcore August dataset (*pAug.TL*), we generate samples which are equal in length and with the same mean value as the various traces analysed in Leland *et al.*[14] and Misra and Gong.[29] For the synthetic traces we generate, we use $\rho = 0.98$ and $\lambda = 318.18$ arrivals per second. The λ selected here is equal to the mean packet arrival rate in Bellcore August dataset. To get an idea of self-similarity of CIPP, we resort to a pictorial representation analogous to those given in Leland[14] and Misra and Gong.[29] We select, at random, sections of the simulated traffic and the Bellcore August dataset and plot them side by side at the same resolution level (Fig. 5). The simulated CIPP process exhibits a similar bursty behaviour as the Bellcore dataset over all plotted scales. Note that the length of the Bellcore August dataset limits the number of points we can plot at the coarsest scale.

Next, we generate CIPP with different ρ and estimate the Hurst parameter of the synthetic trace generated above by the wavelet-based estimator proposed by Abry *et al.*[30] In Fig. 6, we plot the estimated Hurst parameter for CIPP versus the correlation parameter. Note that
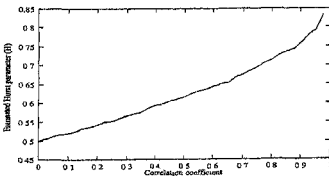

FIG. 6 Estimated local hurst parameter (*H*) versus the correlation parameter.
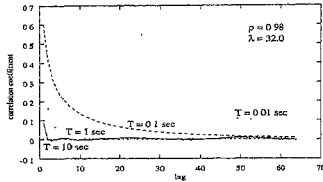
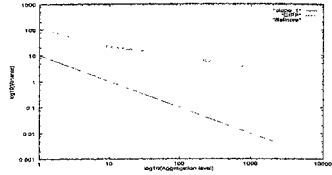FIG. 7. Autocorrelation of CIPP at various scales

FIG. 8. The variance and aggregation level plot for CIPP and the Bellcore dataset

the estimated Hurst parameter is 0.5, corresponding to $\rho = 0$, which is true for the Poisson process. The plot shows that we can model teletraffic with higher H by using a CIPP with larger ρ.

Next, the autocorrelation function at various scales is estimated using a large number of realizations. Figure 7 shows autocorrelation at various scales. One can see, as T (block size over which the counts process $u_T(i)$ is computed from the original point process) increases, autocorrelation tends to approach that of the Poisson case. This is a typical feature of pseudo-self-similar processes (see Roberts *et al.*[2], p. 338).

The next plot (Fig. 8) is the log (variance) versus log (aggregation level) plot for CIPP and the Bellcore dataset. The lower-most curve is a reference curve with slope-1. The behaviour is strikingly similar, with variances for both the processes decaying at the same rate, slower than the curve with slope-1, with increasing aggregation level.

We finally present the wavelet analysis of CIPP. Wornell[31] has proved that as long as the analysing mother wavelet has at least one vanishing moment, the *log of variance* of the wavelet coefficients for a self-similar process increases linearly with the scale with the slope being the Hurst parameter. This interpretation[29] of wavelet transform gives a convenient way of looking at the variance scale plots and is used here. Figure 9 shows such a plot for various processes and CIPP.

We now compare CIPP with Bellcore August dataset using wavelet analysis. We have analysed the two signals using the Daubechies-4 wavelet.[32] The plot of log of variance of
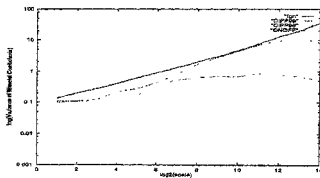

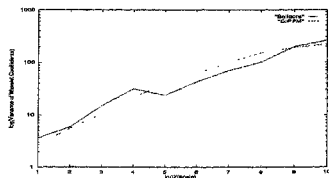FIG. 9. Variance scale plots for CIPP and other processes.

FIG. 10. Variance scale plots for Bellcore dataset and the simulated CIPP process.

wavelet coefficient versus log2 (scale) for the simulated CIPP and Bellcore August dataset shows a strong resemblance in the properties of the two processes as seen in Fig. 10.

4. Conclusions

We have proposed a new traffic model with correlations in interarrivals. It has been shown by simulations that this arrival process models well the real-time traffic source, in particular the video source. The model is motivated by the argument that a strong positive correlation in interarrivals leads to the count process which might be self-similar over a range of time scales. This range of time scale depends on the choice of parameter of the model. Simulation runs of the model also show a close match to the observed dataset as far as 'self-similarity' goes.

We summarize below some findings of the CIPP process:

1. With strong positive correlation in interarrivals one gets strong positive correlation in the counts sequence $u_T(i)$. The magnitude of the correlation depends on ρ.

2. This strong positive correlation in the interarrival sequence results in some complex correlation structure for the counts sequence $u_T(i)$. This feature could be responsible for the self-similar behaviour of CIPP to a range of time scales.

3. CIPP does exhibit some scale-invariant behaviour. This range of time scales depends on the correlation parameter ρ.

4. The relationship between the ρ of CIPP and the Hurst parameter H, given in Section 3.3, will be very useful in network design tools or in tools supporting real-time traffic management.

5. Since CIPP is analytically tractable, it can be used as a full-fledged model in a measurement-based system.

Currently, we are investigating the Markov-modulated CIPP (MMCIPP) in which CIPP is modulated by a continuous-time Markov chain. Preliminary results show that MMCIPP is more self-similar on a much wider scale than that of MMPP or CIPP. The motivation for applying CIPP even when the real traffic might be long-range dependent is that powerful tools have been developed for calculating performance measures for infinite-capacity queues for Markov and deterministic service times.[13, 33]

Acknowledgements

References

1. ROBERTS, J., MOCCI, U. AND VIRTAMO, J.   *Broadband network teletraffic: Performance evaluation and design of broadband multiservice networks*, Springer-Verlag, 1996.

2. BLONDIA, C. AND CASALS, O.   Statistical multiplexing of VBR sources: A matrix analytical approach, *Performance Evaluation*, 1992, **16**, 5–20.

3. LI, S. Q. AND MARK, J. W.   Traffic characterisation of integrated services network, *IEEE Trans.*, 1990, **C-3**, 1231–1243.

4. LI, S. Q.   A general solution technique for discrete queueing analysis of multimedia traffic on ATM, *IEEE Trans.*, 1991, **C-39**, 1115–1132

5. HEFFES, H.   A class of data traffic processes: Covariance function characterisation and related queueing results, *Bell Syst. Tech. J.*, 1980, **59**, 897–929

6. SAITO, H., KAWARASAKI, M AND   An analysis of statistical multiplexing in an ATM transport net-
   YAMADA, H   work, *IEEE J.*, 1991, **SAC-9**, 359–367.

7. HEFFES, H. AND LUCANTONI, D. M.   A Markov modulated characterisation of packetised voice and data traffic and related statistical multiplexer performance, *IEEE J.*, 1986, **SAC-4**, 856–868.

8. ADDIE, R. G. AND ZUKERMAN, M.   An approximation for performance evaluation of stationary single server queues, *IEEE Trans.*, 1994, **C-42**, 3150–3160.

9   http://bellcore.com, 1992.

10. ADDIE, R. G., ZUKERMAN, M. AND   Fractal traffic: Measurements, modeling and performance evalua-
    NEAME, T.   tion, *Proc. IEEE INFOCOM'95*, Apr 1995.

11. PAXSON, V. AND FLOYD, S.   Wide area traffic: The failure of Poisson modeling, *IEEE/ACM Trans. Networking*, 1995, **3**, 226–243.

12. LIVNY, M., MELAMED, B. AND   The impact of autocorrelation function on queueing systems, *Mgmt*
    TSIOLIS, A. K.   *Sci.*, 1993, **39**, 322–339.

13. MANIVASAKAN, R., DESAI, U. B. AND CIPP/M/1 queue, *Natn. Conf. on Communications (NCC'99)*, India,
    KARANDIKAR, A.   Jan 1999.

14. LELAND, W. E., TAQQUE, M S.,   On the self-similar nature of ethernet traffic (extended version),
    WILLINGER, W. AND WILSON, D. V.   *IEEE/ACM Trans Networking*, Feb. 1994, **2**, 1–15.

15. NORROS, J. AND VIRTAMO, J. T.   Who loses cells in case of burst scale congestion, *Proc. 13th Int Teletraffic Cong.*, 1991.

16. TSYBAKOV, B. AND GEORGANAS, N. D. On self-similar traffic in ATM queues: Definitions, overflow prob-
                                      ability bound and cell delay distribution, *IEEE/ACM Trans. Networking*, 1997, **5**, 397–409.

17. GAVER, D. P. AND LEWIS, P. A. W. First-order autoregressive gamma sequences and point processes, *Adv. Appl Prob.*, 1980, **12**, 727–745.

18. DALEY, D. J. AND VERE-JONES, D.   *An introduction to the theory of point processes*, Springer-Verlag, 1988.

19. McFADDEN, J. A.   On the lengths in a stationary point process, *J. R. Statist. Soc. B*, 1993, **24**, 364–382.

20. LAWRANCE, A. J. AND LEWIS, P. A. W. An exponential moving average sequence and point process emal process, *J. Appl. Prob.*, 1977, **14**, 98–113.

21. FINCH, P. D.   The single server queueing system with non-recurrent input-process and Erlang service time, *J. Aust. Math. Soc.*, 1963, **3**, 220–236.

22. MANIVASAKAN, R., DESAI, U. B. AND Correlated interarrival time Poisson Process (CIPP) for modeling
    KARANDIKAR, A.   broadband teletraffic, *Fifth Biennial Conf. on Signal Processing*

		and Communications (SPCOM'99), Indian Institute of Science, Bangalore, July 1999.
23.	HEYMAN, D. P., TABATABAI, A. AND LAKSHMAN, T. V.	Statistical analysis and simulation study of video teleconference traffic in ATM networks, *IEEE Trans.*, 1992, **CSVT-2**, 49–59.
24.	XU, S., HUANG, Z. AND YAO, Y.	A theoretic analysis model for VBR video traffic in ATM networks, *Int. Conf on Communications, 1997 (ICC'97)*, Vol. 2, pp. 844–848.
25.	NEUTS, M. F	A versatile Markovian point process, *J. Appl Prob.*, 1979, **16**, 764–779.
26.	JACOB, P. A. AND LEWIS, P. A. W.	Time series generated by mixtures, *J. Time Ser. Analysis*, 1983, **4**, 19–36.
27.		ftp-info3.informatik.uni-wuerzburg.de/pub/mpeg/, 1995.
28.	GUSELLA, R.	Characterising the variability of arrival processes with indexes of dispersion, *IEEE J.*, 1991, **SAC-9**, 203–211.
29.	MISRA, V. AND GONG, W. B.	*Markovian hierarchical on-off processes for teletraffic modeling*, Tech. Rep., Dept. of ECE, University of Massachusetts, Amherst, 1998, Technical Report No.-TRCCS9801.
30.	ABRY, P. AND VEITCH, D.	Wavelet analysis of long range dependent traffic, *IEEE Trans.*, 1998, **IT-44**, 2–15.
31.	WORNELL, G. W.	Wavelet-based representations for the $1/f$ family of fractal processes, *Proc. IEEE*, 1995, **105**, 15–30.
32.	DAUBECHEIS, I.	Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, 1998, **41**, 909–996.
33.	MANIVASAKAN, R., DESAI, U. B. AND KARANDIKAR, A.	*CIPP/D/1 queue*, Technical Report No-spann.mani.1998.3, SPANN Lab, Indian Institute of Technology, Bombay, May 1998.
34.	DALEY, D. J.	Stationary point processes with Markov dependent intervals and infinite intensity, *J. Appl. Prob. A*, 1982, **19**, 313–320.
35.	COX, D. R. AND ISHAM, V.	*Point processes*, Chapman and Hall, 1980.
36.	PARZEN, E.	*Stochastic processes*, Holyden-Day, 1962.

Appendix I

Construction of stationary CIPP process

For stationarity in the counting variable, the initial point of the interval of the counting must be chosen in a particular probabilistic way.[c] These are the stationary initial conditions which we will give; they can be used to construct a count stationary point process. By doing so, we are exploiting Axiom 3.

We next use Axiom 2. Since our aim is to construct a process whose interarrivals are correlated, we first generate correlated exponential sequence $\{X_0, X_2, X_3,...,\}$ as given[d] in Gaver and Lewis[17] where X_0 is the length-biased distribution with respect to exponential distribution

[c] see, for instance, Daley,[34] Daley and Vere–Jones,[18] and Cox and Isham.[35]
[d] Note the absence of X_1 in the sequence.

with parameter λ and $X_1 = UX_0$ where again U is uniformly distributed in $[0, 1]$. Now form the partial sum,

$$T_n = X_1 + X_2 + ... + X_n \quad n = 1, 2, 3.... \tag{A1}$$

Note that the sequence $\{T_n\}$ gives the arrival instants. Substituting,

$$X_{n+1} = \rho X_n + \varepsilon_{n+1} \qquad 0 \le \rho < 1 \quad n = 1, 2, 3... \tag{A2}$$

and eliminating all X_i's, except X_0, we get,

$$T_n = (U + \alpha_{n-1} - 1)X_0 + \alpha_{n-2}\varepsilon_2 + ... + \alpha_0\varepsilon_n \tag{A3}$$

where X_0 is the length-biased distribution with respect to exponential distribution with parameter λ and $\alpha_j = \sum_{i=0}^{j} \rho^i$ for $0 \le \rho < 1$, and $j = 1, 2, 3,....$ We use $\Phi_W(s)$ to denote the moment-generating function of a random variable W. Define $Y = (U + \alpha_{n-1})X_0$, then,

$$\Phi_Y(s) = \frac{\lambda^2}{\left(\lambda + s(\alpha_{n-1} - 1)\right)\left(\lambda + s\alpha_{n-1}\right)}. \tag{A4}$$

Then exploiting independence among random variables

$$\Phi_{T_n}(s) = \frac{\lambda}{\left(\lambda + s(\alpha_{n-1} - 1)\right)\left(\lambda + s\alpha_{n-1}\right)} \prod_{i=2}^{n} \Phi_\varepsilon\left(\alpha_{n-i}s\right) \tag{A5}$$

where

$$\Phi_\varepsilon(s) = \rho + \frac{(1-\rho)\lambda}{\lambda + s}.$$

See Gaver and Lewis[17] for details. Substituting this, and upon some manipulation, we get,

$$\Phi_{T_n}(s) = \frac{\lambda}{\left(\lambda + s\rho\alpha_{n-2}\right)\left(\lambda + s\alpha_{n-1}\right)} \prod_{i=0}^{n-2} \Phi_\varepsilon\left(\alpha_i S\right) \tag{A6}$$

and upon further reducing , we have

$$\Phi_{T_n}(s) = \frac{(\lambda)^2 \prod_{i=0}^{n-3}\left(\lambda + s\rho\alpha_i\right)}{\prod_{i=0}^{n-1}\left(\lambda + s\alpha_i\right)}. \tag{A7}$$

Rewriting the above,

$$\Phi_{T_n}(s) = \frac{\lambda^2 \rho^{n-2} \prod_{k=0}^{n-3}\left(s + \frac{\lambda}{\rho\alpha_k}\right)}{\alpha_{n-2}\alpha_{n-1} \prod_{k=0}^{n-1}\left(s + \frac{\lambda}{\alpha_k}\right)}. \tag{A8}$$

Inverting the above expression to get the density function $f_{T_n}(t) \underset{=}{\Delta} \mathcal{L}^{-1}\left(\Phi_{T_n}(s)\right)$ and $F_{T_n}(t) \underset{=}{\Delta} \int_0^t f_{T_n}(\tau)d\tau$ we have for distribution $F_{T_n}(t)$,

$$F_{T_n}(t) = \left\{ 1 - \sum_{j=0}^{n-1} \frac{\alpha_j \prod_{k=0}^{n-1}\left[\alpha_j - \rho\alpha_k\right]}{\left(\alpha_j - \rho\alpha_{n-2}\right)\left(\alpha_j - \rho\alpha_{n-1}\right)\prod_{k=0,k\neq j}^{n-1}\left[\alpha_j - \alpha_k\right]} e^{-\frac{\lambda t}{\alpha_j}} \right. \quad 0 \leq \rho < 1. \quad (A9)$$

Next, we derive the distribution (see Parzen[36], p. 133)

$$P_n(t) = F_{T_n}(t) - F_{T_{n+1}}(t)$$

$$= \left\{ \begin{array}{l} \displaystyle\sum_{j=0}^{n} \frac{\alpha_j \prod_{k=0}^{n}\left[\alpha_j - \rho\alpha_k\right]}{\left(\alpha_j - \rho\alpha_{n-1}\right)\left(\alpha_j - \rho\alpha_n\right)\prod_{k=0,k\neq j}^{n}\left[\alpha_j - \alpha_k\right]} e^{-\frac{\lambda t}{\alpha_j}} \\[3ex] \displaystyle -\sum_{j=0}^{n-1} \frac{\alpha_j \prod_{k=0}^{n-1}\left[\alpha_j - \rho\alpha_k\right]}{\left(\alpha_j - \rho\alpha_{n-2}\right)\left(\alpha_j - \rho\alpha_{n-1}\right)\prod_{k=0,k\neq j}^{n-1}\left[\alpha_j - \alpha_k\right]} e^{-\frac{\lambda t}{\alpha_j}} \end{array} \right. \quad 0 \leq \rho < 1.$$

After some manipulations, we get[e] the required result.

$$P_n(t) = \left\{ \begin{array}{l} \left(1 + \rho - (1-\rho)\alpha_n\right) \\[1ex] \displaystyle\sum_{j=0}^{n} \frac{\alpha_j \prod_{k=0}^{n}\left[\alpha_j - \rho\alpha_k\right]}{\left(\rho(\alpha_j + 1) - \alpha_n + 1\right)\left(\alpha_j - \alpha_n + 1\right)\left(\alpha_j - \rho\alpha_n\right)\prod_{k=0,k\neq j}^{n}\left[\alpha_j - \alpha_k\right]} e^{-\frac{\lambda t}{\alpha_j}} \end{array} \right. \quad 0 \leq \rho < 1$$

$$(A10)$$

for $n = 0, 1, 2, 3..., \lambda > 0, t \geq 0$, and $\alpha_j = \sum_{i=0}^{j} \rho^i$.

---

[e]It is worth noting that the value of $P_n(t)$ at $\rho = 0$ is obtained by the limiting value $\lim_{\rho \to 0} P_n(t)$ (which is the Poisson process).