

A source independent traffic model for ATM networks

S. Galmés and R. Puigjaner

Dept. of Mathematics and Computer Science

Universitat de les Illes Balears

Cra. de Valldemossa, km. 7.5, 07071 Palma, Spain

Telephone: +34-71-172989, +34-71-173288

Fax: +34-71-173003

E-mail: dmisgo0, putxi @ps.uib.es

Abstract

At present, most of research on ATM traffic modelling is application dependant. This means that each type of B-ISDN application is described by a set of traffic models, but these models are not capable of capturing the characteristics of other types of applications. On the other hand, from the network performance point of view, just some teletraffic features at cell level are finally of interest, no matter what the sources are. In this paper, we propose the idea of constructing a source independent model, that may capture those relevant features for any type of source. As a first attempt, we focus on the on-off model with general distributions as a possible candidate to reach this objective. For this model, we compute exact-closed form expressions of the burstiness and correlations of the inter-arrival times between cells. Finally, we use these results to set up an algorithm for the evaluation of an equivalent on-off model for any ATM source.

Keywords

GMDP, MMBP, statistical multiplexing, average rate, burstiness, autocorrelation

1 INTRODUCTION

A brief overview of the state of the art on the area of ATM traffic modelling shows that most of research has been application dependant (Cosmas 1994), (Cosmas et al., 1994), (Frost, 1994), (Gahr, Tran-Gia, 1992) and (Kuehn, 1995). Different types of traffic sources, such as voice, data or video, are described by different sets of mathematical models. Moreover, some models are analytically tractable, but some are not. This heterogeneity makes it difficult to analyse the aggregate behaviour or the statistical multiplexing effects of different types of sources - see Figure 1. In these cases, the only mathematical tool that can be applied is usually simulation, with the added difficulty that the variety of aggregations that may be considered is enormous.

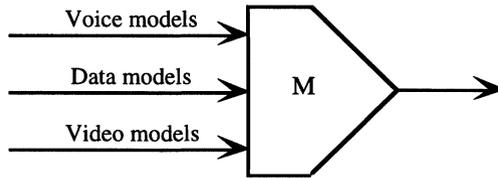


Figure 1 Multiplexing of different types of sources.

Essentially, B-ISDN is a network designed to integrate different types of applications, what means that, within the network, all traffic sources can be viewed just as streams of cells. Therefore, if B-ISDN integrates sources, why not integrate models and construct a general B-ISDN model? This hypothetic source independent model should not try to capture the complete statistical profile of any source, but just the basic traffic characteristics that really affect the internal operation of an ATM network: average rate, burstiness and correlations. Here, not only short-term, but also long-term correlations should be considered. The B-ISDN model should be able to capture these characteristics no matter what the original source is - voice, data or video. In other words, this is a great-scale problem, as distinguished from the classical problem of fitting a few statistical parameters of an ATM source by means of a mathematical model with also a small number of parameters. In the new context, the objective consists of capturing a limited but undefined number of parameters by using new or already known mathematical models.

Just as a way to illustrate the proposed methodology, and with no loss of generality, we have firstly focused on the on-off models with general distributions for the following reasons:

- Because these models have a very general definition. Since no assumption is made on the distributions of the on (active) and off (silence) periods, it seems that their adaptability may be great. On-off models with general distributions are really a class of 2-GMDP (General Modulated Deterministic Processes); consequently, with a just 2-phase model, due to the arbitrariness of the distributions, many parameters

of the input source may be captured (average rate, burstiness and short and long term autocorrelation coefficients). This is not possible with other commonly used and well-known models. For instance, with a 2-MMBP, which also has two phases, there are only four degrees of freedom, according to the four parameters that characterize completely this type of models (transition probabilities between the two phases and their average rates).

- Also, because on-off models with general distributions are analytically tractable, since they are constituted by two renewal processes. Here, renewal theory provides interesting results that will be useful in our characterization.

In summary, our goal is to study the viability of the on-off models with general distributions to capture any combination of the relevant traffic characteristics mentioned above (rates, burstiness and correlation coefficients), just by properly fitting the distributions of the active and silence periods. First, to achieve this objective, we obtain an exact closed-form expression of the burstiness, which was already obtained in (Galmés et al., 1994) in a different way. Next, we end up with a new result, which is the exact closed-form expression of the autocorrelation, for any time lag, of the inter-arrival times between cells. Finally, we use these results to set up an algorithm for the evaluation of an equivalent on-off model for any source.

More specifically, in Section 2 we recall the definition of an on-off model with general distributions. For this model, in Section 3, we construct an exact Markovian characterization of the sequence of successive inter-arrival times between cells. In Section 4, from the results obtained in Section 3, we end up with an exact expression of the burstiness of the on-off model - see (Galmés et al., 1994) for more details, and we also obtain a new exact closed-form expression of the autocorrelation, for any time lag, of the inter-arrival times between cells. Here, we will observe that burstiness and correlations are affected by the complete distribution of the active period of the model, but just by the first and second moments of the distribution of the silence period. We also consider in this section some numerical examples. In Section 5, from the results obtained in Section 4, we set up an algorithm to find an equivalent on-off model for any ATM source, which is characterized by its relevant features: average and peak rates, burstiness and correlations. Finally, in Section 6, we make our conclusions and suggestions for future research.

2 THE ON-OFF MODELS WITH GENERAL DISTRIBUTIONS

An on-off model is constituted by the alternation of two types of periods: an active or on period, in which a cell is transmitted per slot, and a silence or off period, during which no cells are transmitted. We assume that both periods may have arbitrary distributions and that the duration of any period is independent of the duration of any other period, active or silence. Let p_a and p_s be the probability distributions of the active and silence periods respectively. If MBL is the maximum

number of cells of the active periods (maximum burst length) and *MSP* the maximum number of empty slots of the silence periods (maximum silence period), we can express the average values, \bar{a} and \bar{s} respectively, in the following way:

$$\bar{a} = \sum_{l=1}^{MBL} l \cdot p_a(l), \tag{1a}$$

$$\bar{s} = \sum_{l=1}^{MSP} l \cdot p_s(l). \tag{1b}$$

From the definition of the on-off model, the normalized peak rate (rate of active periods in cells per slot) is 1. The normalized average rate or link utilization *m* is the following:

$$m = \frac{\bar{a}}{\bar{a} + \bar{s}}. \tag{2}$$

To complete the characterization of the on-off models from the teletraffic point of view, we need to obtain the expressions of burstiness and autocorrelation. These parameters are directly related to the statistics of the inter-arrival times between cells. In the next section, we characterize the sequence of successive inter-arrival times as an exact Markov chain.

3 CHARACTERIZATION OF THE INTER-ARRIVAL TIMES

Let's define the inter-arrival time between two successive cells as the time (measured in number of cells) between arrival instants (see Figure 2).

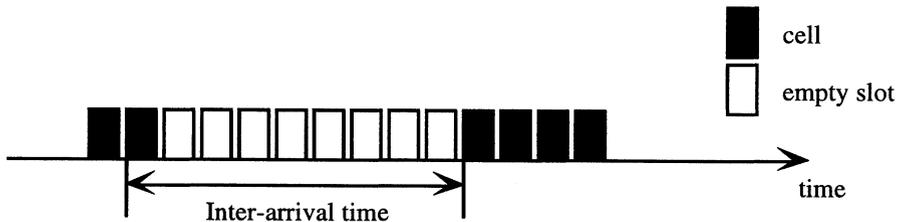


Figure 2 Inter-arrival time between two successive cells in an on-off model.

In Figure 3, we can easily observe that there are just two types of inter-arrival times between cells:

- Inter-arrival time of type A: this type of inter-arrival time takes place between two successive cells that belong to the same active period, and its value is always 1.

- Inter-arrival time of type S : inter-arrival time between the last cell of an active period and the first cell of the following active period. This type of inter-arrival time involves a complete silence period inside and its value can always be expressed as $1+s$, where s is the duration of the inner silence period (expressed in number of cells).

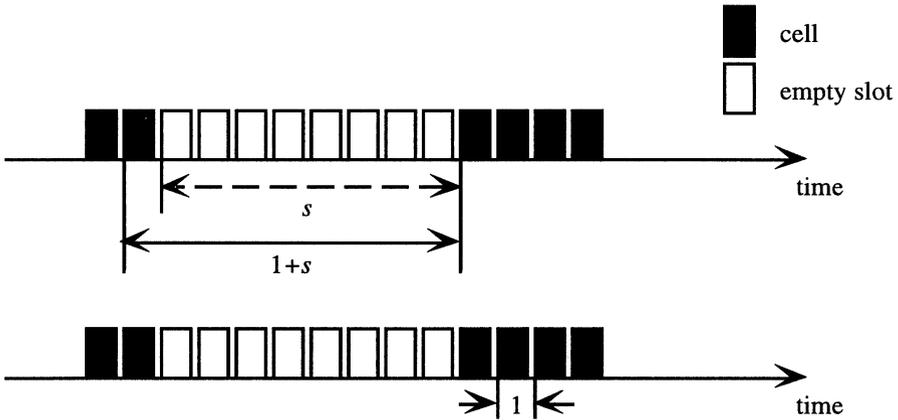


Figure 3 Two types of inter-arrival times in an on-off model.

Neither the active nor the silence period distributions exhibit the memoryless property, because they are general distributions; nevertheless, it is possible to describe the evolution of successive inter-arrival times by means of a Markov chain, if we take into account the actual position of the cells in the case of the active periods. The reason is that the sequence of active periods is, like the sequence of silence periods, a renewal process and, for this class of processes, renewal theory provides some exact results concerning the statistics of the age and residual lifetime of the renewal intervals (Taylor, Karlin, 1984). These results can be used in our context to characterize the transition probabilities of the Markov chain representing the sequence of inter-arrival times. More precisely, in the discrete case, if γ_n is the residual lifetime at time n of a renewal interval of any given renewal process X , the following asymptotic result is provided:

$$\gamma_n \rightarrow \gamma \text{ if } n \rightarrow \infty, \tag{3a}$$

$$\lim_{n \rightarrow \infty} Prob[\gamma_n = k] = Prob[\gamma = k] = \frac{1 - F_x(k-1)}{E(X)}. \tag{3b}$$

In these expressions, γ represents the stationary residual lifetime (we assume the process hits a stationary behaviour), k is any non-negative integer value, n represents the time index and X is the random variable associated to the duration of

the renewal intervals associated with the renewal process X_r . F_x and $E(X)$ are respectively the cumulative distribution function and the expectation of the random variable X . Expression (3a) establishes that the residual lifetime of a renewal process hits a stationary behaviour and expression (3b) shows the corresponding exact stationary distribution.

Another important variable considered in renewal theory is the age or current lifetime of the renewal intervals, δ_n , which also hits a stationary behaviour but, as we will show later on, its statistics can always be reduced to the statistics of the residual lifetime.

Let's define the states of the inter-arrival times between cells as follows:

- A_i : inter-arrival time between the i th cell and the $(i+1)$ th cell of an active (A) period; the subscript i takes values between 1 and $MBL-1$, where MBL is the maximum burst length. All these inter-arrival times take a value of 1.
- S_j : inter-arrival time that contains a silence (S) period of length j . The value of this inter-arrival time is $1+j$. The subscript takes the values between 1 and MSP , where MSP is the maximum silence period.

Note that the meaning of the subscript in the definition of the states is different depending on the type of them. However, this formulation will make it possible to develop an exact Markov chain to model the sequence of inter-arrival times of an on-off model with general distributions. This Markov chain will be constituted by $MBL+MSP-1$ states, and its transition probabilities will be obtained from the independence properties of the process and the results of renewal theory set up before in this section. We define these transition probabilities as:

$$Prob[A_i \rightarrow A_{i'}] = P_{AA}(i, i'), \tag{4a}$$

$$Prob[A_i \rightarrow S_j] = P_{AS}(i, j), \tag{4b}$$

$$Prob[S_j \rightarrow A_i] = P_{SA}(j, i), \tag{4c}$$

$$Prob[S_j \rightarrow S_{j'}] = P_{SS}(j, j'). \tag{4d}$$

The probability defined in (4a) is non-zero only if its two subscripts are consecutive. It can be evaluated taking into account the concepts of age and residual lifetime and expressions (3a) and (3b). Let's consider δ_{An} and γ_{An} as the age and the residual lifetime at time n of the renewal process constituted by the sequence of active periods; the probability defined in (4a) is, for the non-zero cases:

$$\begin{aligned} P_{AA}(i, i+1) &= \lim_{n \rightarrow \infty} Prob[\gamma_{An} \geq 1 / \delta_{An} = i+1] \\ &= \lim_{n \rightarrow \infty} \frac{Prob[\gamma_{An} \geq 1, \delta_{An} = i+1]}{Prob[\delta_{An} = i+1]} \end{aligned}$$

$$= \lim_{n \rightarrow \infty} \frac{\text{Prob}[\gamma_{A_n} \geq 1, \delta_{A_n} = i + 1]}{\text{Prob}[\gamma_{A_n} \geq 0, \delta_{A_n} = i + 1]}. \quad (5)$$

On the other hand, taking into account the definition of the current and residual lifetimes, the following events are equivalent:

$$\{\gamma_{A_n} \geq i, \delta_{A_n} \geq j\} \Leftrightarrow \{\gamma_{A_{n-j}} \geq i + j\}. \quad (6)$$

Expression (6) makes it possible to simplify the problem to just the statistics of the residual lifetime. If we define γ_A as the stationary residual lifetime - expression (3a), and take into account the equivalence (6), expression (5) can be evaluated after some computations:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\text{Prob}[\gamma_{A_n} \geq 1, \delta_{A_n} = i + 1]}{\text{Prob}[\gamma_{A_n} \geq 0, \delta_{A_n} = i + 1]} &= \frac{\lim_{n \rightarrow \infty} \text{Prob}[\gamma_{A_n} \geq 1, \delta_{A_n} = i + 1]}{\lim_{n \rightarrow \infty} \text{Prob}[\gamma_{A_n} \geq 0, \delta_{A_n} = i + 1]} \\ &= \frac{\lim_{n \rightarrow \infty} \{\text{Prob}[\gamma_{A_n} \geq 1, \delta_{A_n} \geq i + 1] - \text{Prob}[\gamma_{A_n} \geq 1, \delta_{A_n} \geq i + 2]\}}{\lim_{n \rightarrow \infty} \{\text{Prob}[\gamma_{A_n} \geq 0, \delta_{A_n} \geq i + 1] - \text{Prob}[\gamma_{A_n} \geq 0, \delta_{A_n} \geq i + 2]\}} \\ &= \frac{\lim_{n \rightarrow \infty} \{\text{Prob}[\gamma_{A_{n-i-1}} \geq i + 2] - \text{Prob}[\gamma_{A_{n-i-2}} \geq i + 3]\}}{\lim_{n \rightarrow \infty} \{\text{Prob}[\gamma_{A_{n-i-1}} \geq i + 1] - \text{Prob}[\gamma_{A_{n-i-2}} \geq i + 2]\}} \\ &= \frac{\text{Prob}[\gamma_A \geq i + 2] - \text{Prob}[\gamma_A \geq i + 3]}{\text{Prob}[\gamma_A \geq i + 1] - \text{Prob}[\gamma_A \geq i + 2]} \\ &= \frac{\text{Prob}[\gamma_A = i + 2]}{\text{Prob}[\gamma_A = i + 1]}. \quad (7) \end{aligned}$$

Finally, according to expression (3b), we can write the following exact result for the transition probability between two consecutive states of type A:

$$P_{AA}(i, i + 1) = \frac{1 - F_a(i + 1)}{1 - F_a(i)}; \quad i = 1, \dots, MBL - 1. \quad (8)$$

For the other probabilities of the Markov chain, we can make use of basic concepts of statistics and the independence properties of the on-off model:

$$P_{AS}(i, j) = [1 - P_{AA}(i, i + 1)]p_s(j) = \frac{p_a(i + 1)}{1 - F_a(i)} p_s(j), \quad (9)$$

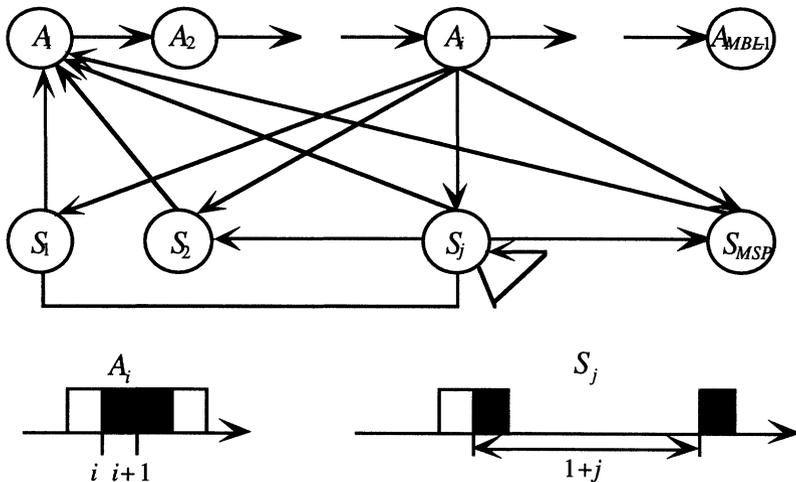
where $i = 1, \dots, MBL - 1$; $j = 1, \dots, MSP$.

The only non-zero probabilities contained in expression (4c) are those in which i is equal to 1, since, obviously, any active period starts at age 1. The probability the Markov chain shifts from any state S to the first state A is the probability that the next arriving cell belongs to an active period with length strictly greater than 1. Therefore, the result is a constant value:

$$P_{SA}(j, 1) = 1 - p_a(1), \quad \forall j. \tag{10}$$

The transition between two successive inter-arrival times of type S involves always an active period of length 1 in the middle of two silence periods; in this case, independence assumptions make it very easy to end up with the following result:

$$P_{SS}(j, j') = p_a(1)p_s(j'); \quad j' = 1, \dots, MSP. \tag{11}$$



	A_1	A_2	A_3	...	$AMBL1$	S_1	S_2	...	$SMSP$
A_1	0	$P_{AA}(1,2)$	0		0	$P_{AS}(1,1)$	$P_{AS}(1,2)$		$P_{AS}(1,MSP)$
A_2	0	0	$P_{AA}(2,3)$		0	$P_{AS}(2,1)$	$P_{AS}(2,2)$		$P_{AS}(2,MSP)$
A_3	0	0	0		0	$P_{AS}(3,1)$	$P_{AS}(3,2)$		$P_{AS}(3,MSP)$
$AMBL2$	0	0	0		$P_{AA}(MBL2,MBL1)$	$P_{AS}(MBL2,1)$	$P_{AS}(MBL2,2)$		$P_{AS}(MBL2,MSP)$
$AMBL1$	0	0	0		0	$P_{AS}(MBL1,1)$	$P_{AS}(MBL1,1)$		$P_{AS}(MBL1,1)$
S_1	$P_{SA}(1,1)$	0	0		0	$P_{SS}(1,1)$	$P_{SS}(1,2)$		$P_{SS}(1,MSP)$
S_2	$P_{SA}(2,1)$	0	0		0	$P_{SS}(2,1)$	$P_{SS}(2,2)$		$P_{SS}(2,MSP)$
$SMSP$	$P_{SA}(MSB)$	0	0		0	$P_{SS}(MSB)$	$P_{SS}(MSB,2)$		$P_{SS}(MSB,MSB)$

Figure 4 Markovian characterization of the sequence of inter-arrival times of a general on-off model. The matrix is constituted by four different regions.

Expressions (8), (9), (10) and (11) are the only non-zero probabilities of the Markov chain that represents the sequence of inter-arrival times of an on-off model with general distributions. These expressions are all exact, and are based on the probability distributions of the active and silence periods of the model. Figure 4 shows the state diagram of this Markov chain and the associated transition probability matrix, $\mathbf{T}[1]$, whose elements have already been evaluated. We can observe clearly four regions in the matrix, depending on the type of states between which the transitions take place. Also, from expressions (8), (9), (10) and (11), it is not difficult to show that this matrix is stochastic.

Finally, it remains to calculate the steady-state probabilities. If \mathbf{p} is a vector containing these probabilities, we need to solve the following equation:

$$\mathbf{p} = \mathbf{pT}[1]. \quad (12)$$

Starting from this equation and taking into account expressions (8) and (10), we can easily show that the steady-state probabilities of the states of type A obey this simple relation:

$$Prob[A_i] = [1 - F_a(i)]C; \quad \forall i = 1, \dots, MBL - 1, \quad (13)$$

where C is a constant defined as follows:

$$C = \sum_{j=1}^{MSP} Prob[S_j]. \quad (14)$$

On the other hand, it follows from equation (12) that the steady-state probabilities of the states of type S can be expressed in this general way:

$$Prob[S_j] = p_s(j)C; \quad \forall j = 1, \dots, MSP. \quad (15)$$

Finally, C can be obtained from the constraint on the sum of all steady-state probabilities:

$$\sum_{i=1}^{MBL-1} Prob[A_i] + \sum_{j=1}^{MSP} Prob[S_j] = 1 \Rightarrow C = 1/\bar{a}. \quad (16)$$

4 ANALYSIS OF BURSTINESS AND AUTOCORRELATION

In the previous section, we have analysed the sequence of successive inter-arrival times of an on-off model with general distributions, and developed an exact Markovian representation. In this section, we use this formulation to end up with

the exact closed-form expressions of the burstiness and autocorrelation of those models. We also consider some numerical examples.

Burstiness

Assuming the squared coefficient of variation of the inter-arrival times as a definition of burstiness, we can set up the following:

$$b = \frac{Var[T_n]}{E^2[T_n]} = m^2 \cdot E[T_n^2] - 1. \tag{17}$$

The second moment of the distribution of the inter-arrival times can be obtained starting from its definition, since we already know the steady-state probabilities:

$$E[T_n^2] = 1^2 \cdot \sum_{i=1}^{MBL-1} Prob[A_i] + \sum_{j=1}^{MSP} (1+j)^2 \cdot Prob[S_j] = 1 + \frac{\bar{s}}{a} [2 + \bar{s}(1 + c_s^2)], \tag{18}$$

where c_s^2 is the squared coefficient of variation of the silence period distribution. If we substitute expression (18) in (17), we can obtain the following closed-form expression:

$$b = m(1 - m) \left[(1 + c_s^2) \bar{s} - \frac{1 - m}{m} \right]. \tag{19}$$

This expression is exact. Note that the greater the variability of the silence periods, the burstier the traffic is. Also note that the variability of the active periods does not affect the burstiness. For more details see (Galmés et al., 1994), where expression (19) was already obtained in a different way.

Autocorrelation structure

The autocorrelation function is defined as follows:

$$\phi(k) = \frac{E[T_n T_{n+k}] - m^{-2}}{Var[T_n]}. \tag{20}$$

Here, T_n and T_{n+k} are the inter-arrival times at instants n and $n+k$. Therefore, to obtain the complete autocorrelation structure of the model, we need to characterize the sequence of inter-arrival times for any lag k . If $\mathbf{T}[1]$ is the state transition matrix that characterizes the sequence of successive inter-arrival times (see Figure 4), the corresponding matrix for any given lag k is the following:

$$\mathbf{T}[k] = \mathbf{T}^k[1]; \quad \forall k \geq 1. \quad (21)$$

Let's call $aa^{(k)}(i, i')$, $as^{(k)}(i, j)$, $sa^{(k)}(j, i)$ and $ss^{(k)}(j, j')$, with $i, i' = 1, \dots, MBL-1$ and $j, j' = 1, \dots, MSP$, the elements of the matrix $\mathbf{T}[k]$ for any value of the lag k , including $k=1$. The names of these elements are intended to specify the type of states among which the transitions take place; for instance, $aa^{(k)}(i, i')$ denotes the transition probability from the state A_i to the state $A_{i'}$, when there are $k-1$ intermediate states between them. Each type of element corresponds to one of the four regions that appear in all matrices $\mathbf{T}[k]$. According to this notation, the expectation contained in (20) can be expressed as follows:

$$\begin{aligned} E[T_n T_{n+k}] &= \sum_{i, i'=1}^{MBL-1} aa^{(k)}(i, i') Prob[A_i] + \sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} (1+j) as^{(k)}(i, j) Prob[A_i] \\ &+ \sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} (1+j) sa^{(k)}(j, i) Prob[S_j] + \sum_{j, j'=1}^{MSP} (1+j)(1+j') ss^{(k)}(j, j') Prob[S_j]. \end{aligned} \quad (22)$$

Each component of expression (22) can be evaluated separately:

$$\sum_{i, i'=1}^{MBL-1} aa^{(k)}(i, i') Prob[A_i] = \sum_{i=1}^{MBL-1} Prob[A_i] \sum_{i'=1}^{MBL-1} aa^{(k)}(i, i') = \sum_{i=1}^{MBL-1} Prob[A_i] - \sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} Prob[A_i] as^{(k)}(i, j),$$

$$\sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} (1+j) as^{(k)}(i, j) Prob[A_i] = \sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} as^{(k)}(i, j) Prob[A_i] + \sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} j \cdot as^{(k)}(i, j) Prob[A_i],$$

$$\begin{aligned} \sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} (1+j) sa^{(k)}(j, i) Prob[S_j] &= \sum_{j=1}^{MSP} (1+j) Prob[S_j] \sum_{i=1}^{MBL-1} sa^{(k)}(j, i) \\ &= \sum_{j=1}^{MSP} (1+j) Prob[S_j] - \sum_{j, j'=1}^{MSP} (1+j) ss^{(k)}(j, j') Prob[S_j] \\ &= \sum_{j=1}^{MSP} Prob[S_j] + \sum_{j=1}^{MSP} j \cdot Prob[S_j] \\ &\quad - \sum_{j, j'=1}^{MSP} ss^{(k)}(j, j') Prob[S_j] - \sum_{j, j'=1}^{MSP} j \cdot ss^{(k)}(j, j') Prob[S_j] \end{aligned}$$

and

$$\begin{aligned} \sum_{j, j'=1}^{MSP} (1+j)(1+j') ss^{(k)}(j, j') Prob[S_j] &= \sum_{j, j'=1}^{MSP} ss^{(k)}(j, j') Prob[S_j] + \sum_{j, j'=1}^{MSP} j' \cdot ss^{(k)}(j, j') Prob[S_j] \\ &\quad + \sum_{j, j'=1}^{MSP} j \cdot ss^{(k)}(j, j') Prob[S_j] + \sum_{j, j'=1}^{MSP} j \cdot j' \cdot ss^{(k)}(j, j') Prob[S_j] \end{aligned}$$

Now, grouping again all the components, the cross expectation among the inter-arrival times can finally be developed as follows:

$$\begin{aligned}
 E[T_n T_{n+k}] &= 1 + \sum_{i=1}^{MBL-1} \sum_{j=1}^{MSP} j \cdot as^{(k)}(i, j) Prob[A_i] + \sum_{j=1}^{MSP} j \cdot Prob[S_j] \\
 &\quad + \sum_{j,j'=1}^{MSP} j' \cdot ss^{(k)}(j, j') Prob[S_j] + \sum_{j,j'=1}^{MSP} j \cdot j' \cdot ss^{(k)}(j, j') Prob[S_j] \\
 &= 1 + 2 \sum_{j=1}^{MSP} j \cdot Prob[S_j] + \sum_{j,j'=1}^{MSP} j \cdot j' \cdot ss^{(k)}(j, j') Prob[S_j]. \tag{23}
 \end{aligned}$$

From expression (23), we can observe that we just need to know one of the four regions of all matrices $T[k]$, that is, the region constituted by the elements $ss^{(k)}(j, j')$. Also, expression (23) only involves the stationary probabilities of the states of type S of the inter-arrival times between cells.

It is convenient to rename the elements of $T[k]$ in a generic way, such as $t^{(k)}(l, m)$:

$$t^{(k)}(l, m) = aa^{(k)}(l, m); \quad l, m = 1, \dots, MBL - 1, \tag{24a}$$

$$t^{(k)}(l, m) = as^{(k)}(l, m - MBL + 1), \tag{24b}$$

where $l = 1, \dots, MBL - 1; m = MBL, \dots, MSP + MBL - 1$,

$$t^{(k)}(l, m) = sa^{(k)}(l - MBL + 1, m) \tag{24c}$$

where $l = MBL, \dots, MSP + MBL - 1; m = 1, \dots, MBL - 1$, and finally

$$t^{(k)}(l, m) = ss^{(k)}(l - MBL + 1, m - MBL + 1); \quad l, m = MBL, \dots, MSP + MBL - 1. \tag{24d}$$

For any lag k , the elements $t^{(k)}(l, m)$ can be expressed in terms of the elements of the matrix $T[1]$, which are already known, in the following way:

$$t^{(k)}(l, m) = \sum_{i_1=1}^{MSP+MBL-1} \sum_{i_2=1}^{MSP+MBL-1} \dots \sum_{i_{k-1}=1}^{MSP+MBL-1} t^{(1)}(l, i_1) t^{(1)}(i_1, i_2) \cdot \dots \cdot t^{(1)}(i_{k-1}, m). \tag{25}$$

The special structure of the matrix $T[1]$ makes it not difficult to develop expression (25) for the elements we are interested in - expression (24d) -. Starting from the observation of some particular cases, we will derive a general rule. For instance, for $k = 1$ and considering $l, m = MBL \dots MSP + MBL - 1$, we can set up the following:

$$t^{(1)}(l, m) = P_{ss}(l - MBL + 1, m - MBL + 1) = p_a(1)p_s(m - MBL + 1). \tag{26}$$

In the case of $k = 2$, more transition probabilities of the original Markov chain appear. Taking into account their expressions, it is neither difficult to obtain the final result. It shows that only the first and second samples of the probability mass function of the active period affect the autocorrelation up to lag 2:

$$\begin{aligned}
 t^{(2)}(l, m) &= \sum_{i_1=1}^{MSP+MBL-1} t^{(1)}(l, i_1) t^{(1)}(i_1, m) = t^{(1)}(l, 1) \cdot t^{(1)}(1, m) + \sum_{i_1=MBL}^{MSP+MBL-1} t^{(1)}(l, i_1) t^{(1)}(i_1, m) \\
 &= P_{SA}(l - MBL + 1, 1) P_{SS}(l - MBL + 1, m - MBL + 1) \\
 &\quad + \sum_{i_1=MBL}^{MSP+MBL-1} P_{SS}(l - MBL + 1, i_1 - MBL + 1) P_{SS}(i_1 - MBL + 1, m - MBL + 1) \\
 &= [1 - p_a(1)] p_a(1) p_s(m - MBL + 1) + \sum_{i_1=MBL}^{MSP+MBL-1} p_a(1) p_s(i_1 - MBL + 1) p_a(1) p_s(m - MBL + 1) \\
 &= [p_a^2(1) + p_a(2)] p_s(m - MBL + 1). \tag{27}
 \end{aligned}$$

For $k = 3$ we can list all the combinations of subscripts that generate non-zero products and evaluate these products in terms of the distributions of the active and silence periods. Table 1 shows the results. Grouping all the factors, we can obtain this final expression for $t^{(3)}(l, m)$:

$$t^{(3)}(l, m) = [p_a^3(1) + 2p_a(1)p_a(2) + p_a(3)] p_s(m - MBL + 1). \tag{28}$$

Table 1 Factors in $t^{(3)}(l, m)$. The brackets mean $[MBL \dots MSP+MBL-1]$

l	i_1	i_2	m	Factor
[...]	1	2	[...]	$p_a(3) \cdot p_s(m - MBL + 1)$
		[...]	[...]	$p_a(1) \cdot p_a(2) \cdot p_s(m - MBL + 1)$
	[...]	1	[...]	$p_a(1) \cdot p_a(2) \cdot p_s(m - MBL + 1)$
		[...]	[...]	$p_a^3(1) \cdot p_s(m - MBL + 1)$

From the observation of these three particular cases, we can see that subscript l never appears in the final result, and subscript m always appears in the same way. We could continue these computations for increasing values of k and infer, by observation, that the general expression of the elements $t^{(k)}(l, m)$ obeys, for $l, m = MBL, \dots, MSP+MBL-1$ and for any k , the following exact recursive formulation:

$$t^{(k)}(l, m) = t^{(k)} p_s(m - MBL + 1), \tag{29a}$$

$$t^{(k)} = \sum_{\lambda=1}^k p_a(\lambda) t^{(k-\lambda)}. \tag{29b}$$

Here, $t^{(0)} = 1$ has been assumed for convenience. Taking into account that, for $l, m = MBL, \dots, MSP+MBL-1$, the elements $t^{(k)}(l, m)$ correspond to the elements

$ss^{(k)}(j, j')$ in our original notation, with $j, j' = 1, \dots, MSP$, we can write the following statement:

$$ss^{(k)}(j, j') = t^{(k)} p_s(j'). \tag{30}$$

Taking into account expressions (23), (30) and the steady-state probabilities obtained in the previous section, we can end up with the following exact formula, valid for any value of the correlation lag k :

$$E[T_n T_{n+k}] = 1 + \frac{\bar{s}}{a} [2 + \bar{s} \cdot t^{(k)}]. \tag{31}$$

Finally, the exact closed-form expression of the autocorrelation function between the inter-arrival times for any lag k can be expressed as follows:

$$\phi(k) = \frac{t^{(k)} - 1/\bar{a}}{1 + c_s^2 - 1/\bar{a}}; \quad \forall k \geq 1. \tag{32}$$

Expression (32) shows that correlations depend on the distribution of the active period and only on the first and second moments of the distribution of the silence period. This expression and the expression of burstiness are remarkably simple.

Numerical examples

In this section, we evaluate expression (32) for different distributions of the active period, such as constant, uniform or geometric. These distributions are extensively used in ATM as the base of more complex models. The reason is that they are very simple and, simultaneously, they contribute to capture the diversity of traffic behaviour. Referring to the squared coefficient of variation, which basically affects as a compression factor in (32), we have assumed $c_s^2=1$ in all examples, since small values produce correlation functions with higher variability. The variety of correlation patterns we will obtain suggests that on-off models with general distributions may really accommodate a wide set of traffic sources.

The simplest example is probably an on-off model with constant active period distribution. In fact, its autocorrelation function is periodic, with peaks of positive correlation at multiple values of the burst length B , and close to zero in the rest:

$$\phi(k) = \begin{cases} \frac{1 - 1/B}{1 + c_s^2 - 1/B} & \text{if } k = \dot{B} \\ \frac{-1/B}{1 + c_s^2 - 1/B} & \text{if } k \neq \dot{B} \end{cases} \tag{33}$$

Figure 5 plots this correlation function for different values of the burst length. The burst length controls the position of the correlation peaks, while c_s^2 controls inversely their magnitude. Particularly, for any non-zero value of c_s^2 and $B=1$, the model is completely uncorrelated (all inter-arrival times are of type S , and silence periods constitute a renewal process).

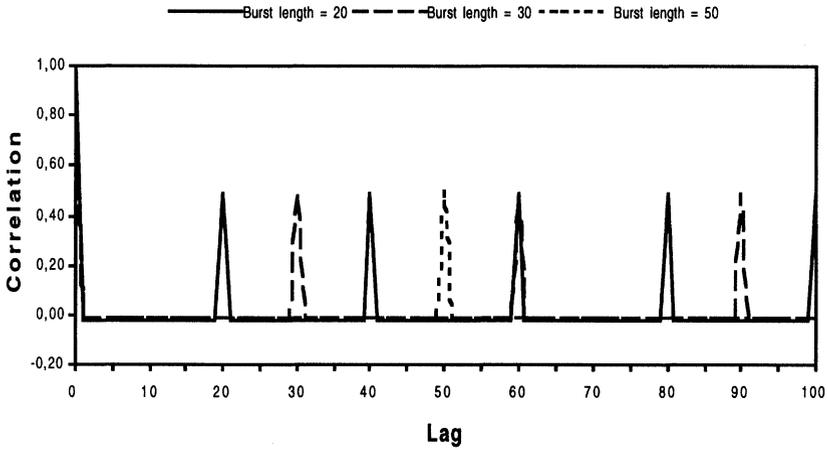


Figure 5 Correlation structure of on-off models with constant active period.

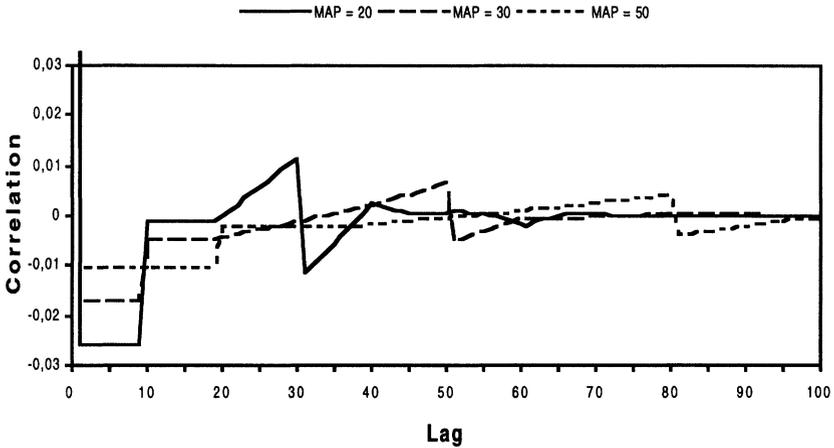


Figure 6 Correlation structure of on-off models with uniform active period distribution.

If the active period is uniformly distributed, the curves are more erratic, as it is shown in Figure 6. In this figure, we have considered different values of the mean active period MAP . As it is shown, negative correlations are captured, with small absolute values that decrease as lag increases.

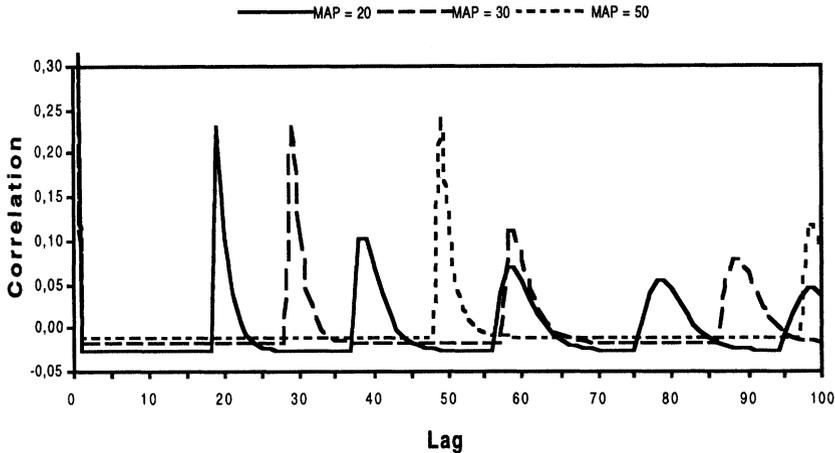


Figure 7 Correlation structure of on-off models with geometric active period distribution.

Finally, if the distribution of the active period is geometric, the correlation function shows decreasing peaks at multiple values of the mean active period, as depicted in Figure 7. As it was in the case of the constant active period distribution, the mean value controls the position of the peaks along the axis, while the squared coefficient of variation of the silence periods affects their magnitude. Now, in comparison with the constant case, the correlation peaks are asymmetric and wider.

In the following section, we address the problem of finding an on-off model that captures the teletraffic characteristics of any given ATM source. To achieve this goal, we make use of the exact expressions of the burstiness and the autocorrelation that we have obtained in this paper.

5 THE SOURCE INDEPENDENT ALGORITHM

The idea of proposing the on-off model with general distributions as a possible source independent model, can be better outlined now. The reason is that we already know its relevant teletraffic features from the network performance point of view: average rate, burstiness and autocorrelation function of the inter-arrival times.

Given any source (voice, data, video) characterized by its relevant features $[m, b, \phi(k)]$, an equivalent model algorithm can be established by inverting the expressions of burstiness and correlations presented in this paper. Certainly, there is no problem to deal with the expression of burstiness, which is quite simple. However, the expression of the autocorrelation function involves a recursive formula in which the unknowns have to satisfy a set of constraints, since they correspond to a probability distribution. Therefore, the algorithm we propose in this paper will capture the average rate and burstiness in an exact way and, as far as the autocorrelation is concerned, it will produce an optimized solution. Basically, the reason is that a system of equations (an equation for each correlation lag) with constraints appears when fitting the autocorrelation function. This section only deals with the formulation of the equivalent model algorithm. In a further work (see also Section 6), this algorithm should be tested for many different types of traffic sources, in order to examine the distance between the optimized solution and the real measurements. Also, it is not the purpose of this work to show how the features of the input sources could be obtained, and if they could be measured or not in real time. This paper is not involved with call admission processes, policy functions or other real-time operations; rather, it deals with the existence of a good (in the sense of optimality) homogeneous representation of the heterogeneous ATM sources, in order to simplify and improve the network performance analysis.

More specifically, the optimization problem can be formulated as

$$\min_{\{p_a(l)\}} \sum_k \left\{ t^{(k)} - \left[\sum_{l=1}^k p_a^*(l) \cdot t^{(k-l)} \right] \right\}^2, \quad (34)$$

where the mean square error has been formulated and $\{p_a^*(l)\}$ is a conforming solution:

$$p_a^*(l) \geq 0; \quad \forall l, \quad (35a)$$

$$\sum_{l=1}^{MBL} p_a^*(l) = 1. \quad (35b)$$

Since the elements $t^{(k)}$ and the correlations are linearly related, an optimal solution for those elements will be an optimal solution for the correlations. The dimension of the optimization problem is *MBL*; therefore, this parameter has to be fixed, despite it has not been included in the set of relevant features of the source. This means that the maximum burst length of the model will be an open parameter in the algorithm, which will explore different values of it and provide, for each value, an optimal solution. Finally, the equivalent model will be the best solution among all the preliminary. This approach is more general than assuming a fixed

value for *MBL*, usually the real *MBL*. Anyway, if *MBL* is considered as a relevant feature, our algorithm reduces the problem to just provide a single solution.

From the expressions of the average rate and burstiness, which constitute part of the input data of the algorithm, it is not difficult to set up a simple relation f between the squared coefficient of variation of the silence period distribution and the mean active period of the on-off model:

$$c_s^2 = f(\bar{a}) = \frac{b + (1 - m)^2}{(1 - m)^2 \bar{a}} - 1. \tag{36}$$

The formulation of the optimization problem for the active period distribution, requires the elements $t^{(k)}$ to be fixed from the correlations. This also requires to set numerical values for the mean active period and the squared coefficient of variation of the silence period - see expression (32) -. On the other hand, the resolution of the optimization problem will usually produce a different value for the mean active period, since this will be newly computed from the resultant active period distribution. Moreover, from expression (36), a new mean active period implies a new squared coefficient of variation. In other words, the optimization problem will be formulated and solved in the core of an iterative procedure, which will stop when the variation of the mean active period from one iteration to another is less than some prefixed precision ϵ .

We can now set up the algorithm for any given value of the maximum burst length *MBL*; of course, the following process should be repeated for each new value of *MBL*:

Initialization step. Set c_s^2 . For instance, $c_s^2 = 0$ (constant silence period distribution). Compute \bar{a} as $\bar{a} = f^{-1}(c_s^2)$.

Step 1. For all k , compute the elements $t^{(k)}$ from the correlations:

$$t^{(k)} = \phi(k) \left[1 + c_s^2 - 1/\bar{a} \right] + 1/\bar{a}$$

Step 2. Solve the optimization problem:

$$\min_{\{p_a^*(l)\}} \sum_k \left\{ t^{(k)} - \left[\sum_{l=1}^k p_a^*(l) \cdot t^{(k-l)} \right] \right\}^2$$

with $p_a^*(l) \geq 0$ and $\sum_{l=1}^{MBL} p_a^*(l) = 1$.

Step 3. Compute the variation on the mean active period:

$$\text{If } \left| \sum_{l=1}^{MBL} l \cdot p_a^*(l) - \bar{a} \right| \leq \epsilon \text{ then STOP.}$$

Else:

$$\text{Set } \bar{a} = \sum_{l=1}^{MBL} l \cdot p_a^*(l). \text{ Compute } c_s^2 \text{ as } c_s^2 = f(\bar{a}).$$

Go to step 1.

6 CONCLUSIONS AND FURTHER RESEARCH

In this paper, we have proposed the idea of a source independent model in contrast to most of the ATM traffic models, which are source-specific. We have focused on the features of the ATM sources that are specially relevant to the network performance viewpoint, namely average rate, burstiness and short and long term autocorrelation behaviour. The idea of constructing some kind of universal model is not only attractive by itself, but also it is quite interesting since it allows to homogenize and simplify the analysis of the aggregate or multiplexed traffic independently of the nature of sources.

As a first attempt, we have considered the on-off models with general distributions for the following reasons:

- Because of their definition, these models seem to be very flexible to capture the characteristics of interest, in contrast to other models, which are more restrictive and exhibit more specific and rigid correlation patterns.
- Another important reason is the possibility of applying some powerful and very well known analytical tools, such as renewal theory and Markov chain theory, to characterize the on-off models with general distributions in an exact way.
- Finally, the statistical multiplexing of several on-off sources with general distributions is a problem that has already been addressed in the literature - see (Elsayed, 1994), (Sohraby, 1993) and (Wittevrongel, Bruneel, 1994).

In this paper, we have developed a new exact closed-form expression of the complete autocorrelation function of the on-off models with general distributions. We have obtained the correlation patterns generated by different distributions of the active period, such as constant, uniform or geometric. Finally, we have set up an algorithm that is intended to produce an equivalent on-off model for an arbitrary ATM source. Specifically, it has been designed to generate an on-off model with the same average rate and burstiness as the original source, and with an optimized autocorrelation function. This is one of the first great-scale proposed algorithms for capturing the whole cell-level behaviour of the ATM sources. Currently, we are in the process of studying experimentally its real capabilities. To perform this task, we are collecting real measurements of ATM traffic at cell level, such as voice, data and video traffic traces. The analysis should take into account timing considerations as well. The extension of the methodology proposed in this paper to some other traffic models is still an open issue. New equivalent-model algorithms could be obtained and compared from different points of view, such as flexibility and time efficiency.

7 REFERENCES

- Cosmas, J. P. et al. (1994) A review of voice, data and video traffic models for ATM. *Europ. Trans. on Telecomm.*, 5, 11-26.

- Cosmas, J. (1994) Stochastic source models and applications for ATM. *Proceedings of the Second Workshop on Performance Modelling and Evaluation of ATM Networks, Bradford (UK)*.
- Elsayed, K. (1994) On the superposition of discrete-time Markov renewal processes and application to statistical multiplexing of bursty traffic sources. *Proceedings of the IEEE GLOBECOM'94, San Francisco*.
- Frost, V.S. and Melamed, B. (1994) Traffic modelling for telecommunications networks. *IEEE Communications Magazine*, 32.
- Galmés, S., Perros, H.G. and Puigjaner, R. (1994) Analysis of burstiness for on-off traffic models with arbitrary distributions. *Proceedings of the Second Workshop on Performance Modelling and Evaluation of ATM Networks, Bradford (UK)*.
- Gühr, O. and Tran-Gia, P. (1992) A layered description of ATM cell traffic streams and correlation analysis. *Proceedings of the IEEE INFOCOM'92*, 2D.4.1-2D.4.8.
- Kuehn, P.J. (1995) Reminder on queueing theory for ATM networks. *First International ATM Traffic Expert Symposium, Basel*.
- Sohraby, K. (1993) On the theory of general on-off sources with applications in high-speed networks. *Proceedings of the IEEE INFOCOM'93, San Francisco*, 401-410.
- Taylor, H.M. and Karlin, S. (1984) *An introduction to stochastic modelling*. Academic Press.
- Wittevrongel, S. and Bruneel, H. (1994) Queue length and delay for statistical multiplexers with variable-length messages. *Proceedings of the IEEE GLOBECOM'94, San Francisco*.

8 BIOGRAPHY

Sebastián Galmés received the degree of Telecommunication Engineer from the Universitat Politècnica de Catalunya (Spain) in 1989. In 1991 he joined the Department of Computer Science of the Universitat de les Illes Balears, where he is currently Assistant Professor. His current teaching and research interests are in the area of performance evaluation of distributed systems and simulation analysis.

Ramon Puigjaner received the degree of Industrial Engineer from the Universitat Politècnica de Catalunya (Spain) in 1964, his Master degree in Aeronautical Sciences from the Ecole Nationale Supérieure de l'Aéronautique de Paris, his Ph. D. degree from the Universitat Politècnica de Catalunya in 1972 and his degree of License in Informatics from the Universidad Politécnica de Madrid. He is member of the IFIP WG 6.3 on Performance of Computer Networks, WG 6.4 on Local and Metropolitan Communication Systems and WG 10.3 on Distributed Systems. His current research interests are the performance evaluation of computer systems.