Correlation in ATM Traffic Streams – Some Results*

Bjarne E. Helvik† Per Hokstad† Norvald Stol†

The correlation/covariance in the cell stream from a single ATM source or a group of sources is an important characteristic in addition to the more commonly regarded characteristics. Knowledge of the correlation in the cell stream is necessary for dimensioning, policing, resource allocation and valuation of service characteristics. This paper provides results related to the correlation between the arrivals of cells from single and multiple sources, necessary for engineering of ATM systems. The serial correlation is introduced and discussed and its relation to the index of dispersion is shown. A state oriented source with general, dependent phase type distributed sojourn times at the active-passive (burst) level is introduced. At the cell level, cells may be generated according to a number of rules depending on the state of the source. The correlation in the cell stream from the above type of source is determined. The results are also extended to account for connections and disconnections. The single source correlations are used to obtain exact and approximate results for the superpositioned cell stream from a number of different sources, e.g. the volume of cells in an interval and excess duration.

1 Introduction

The introduction of a broadband multiservice network, B-ISDN, based on the asynchronous transfer mode (ATM) is very likely. This network will handle a very wide range of traffic types. Some of these have known characteristics like telephony and low/medium speed data traffic, while others have unknown or at best anticipated characteristics like variable bitrate coded video. For all traffic types, it is extremely important to be able to characterize the cell stream stemming from both single sources and mixes of a number of sources in a way that allows an appropriate dimensioning of the network and the network elements, and the construction of mechanisms for policing, resource allocation, etc.

The importance of characteristics like the mean cell rate, the peak cell rate and instantaneous variance of the cell rate is commonly recognized. The correlation of the cell stream, however, is an important property which so far has received little attention. The correlation reflects the dynamic properties of the cell stream. A cell stream with rapid fluctuations of the cell rate may for instance have the same mean, peak and variance as a stream with slow variations. The correlation structure of the two streams will however be different. The importance of this for dimensioning, policing, resource allocation and valuation of service characteristics is obvious [1, 11, 13].

The limited attention paid to the correlation as a major characteristic of a cell stream may be due to the Palm-Khinchin Theorem [9, Chapter 5]. Misinterpretation of this theorem may lead to the assumption that a cell stream stemming from a number of sources may be regarded as a Poisson process. This has also led to too simple models for a composite source, like the Markov Modulated Poisson Process (MMPP) [13]. The first to point out that the traffic stemming from a large number of independent sources may be heavily correlated were Srinam and Whitt [14]. A discussion of the importance of taking the correlation into account is given by Ramaswami [13]. In both the above papers a simple source model was used, only superposition of traffic from homogeneous sources and traffic uncoloured by multiplexing and splitting in the network was studied. To investigate the effect of more complex and heterogeneous sources and colouring, a simulation study was performed [7]. This study showed that the overall correlation structure was maintained. The results give no reason to assume that the cell arrival process may be approximated by a Poisson process. Recently, some results are presented for queue lengths arising when the offered traffic stems from a number of heterogeneous correlated sources [11].

The present paper addresses the serial correlation in ATM cell streams. The results presented provide improved insight into the correlation by results useful for traffic engineering. More specific, after a brief introduction to correlation and related measures of traffic streams in Section 2, a quite general state oriented model of a single source is introduced in Section 3. The sources have general dependent phase type distributed sojourn times at the active-passive (burst) level. At the cell level cells may be generated according to a number of rules depending on the state of the source. In Section 4, the correlation in the cell stream from the above class of sources is determined. The results are also extended to account for connections and disconnections. Next, in Section 5, it is shown how the correlation of traffic streams stemming from a large number of heterogeneous sources may be determined. Approximations for such traffic streams are also derived, based on the assumption of normality. The distribution of the volume of cells arriving in a period is derived, and so is the duration of periods where the number of cells arriving exceeds a certain capacity.

*The work reported in this paper are done under contracts with Norwegian Telecom, Research Department.
†ELAB-RUNIT, N-7034 Trondheim, Norway
†SINTEF Safety and Reliability, N-7034 Trondheim, Norway
2 Correlation

The 'correlation structure' of a point process can be described either by correlations of the inter-event times or by the serial correlation of the number of events within various intervals. In general there seems to be no simple expression for the relation between these two sets of correlations. Using the serial correlations as a measure for dependence in the traffic seems advantageous when there are various sources of traffic. When these serial correlations are obtained for each source of traffic, also the corresponding result for the total traffic is rather directly obtained (see Section 5). In the present paper we restrict to consider the serial correlations.

2.1 Serial Correlation

For a stationary point process, let \( N(t) \) be the number of 'events' (i.e. generated cells) within an interval of length \( t \). The unit of time equals one slot, and we actually restrict to consider time instants being a multiple of one slot. Further, the time axis is divided into intervals of constant length \( \ell \), see Figure 1. Thus, one interval is made up of \( \ell \) slots.

![Figure 1: Correlation between subsequent intervals](image)

Further, let

\[
X_i(\ell) = N(i \cdot \ell) - N((i-1) \cdot \ell)
\]

the number of generated cells in the \( i \)-th interval of length \( \ell \),

\[
h_j(\ell) = \frac{\text{Cov}(X_i(\ell), X_i+j(\ell))}{\text{Var}(X_i(\ell))}
\]

the serial correlation between \( X_i(\ell) \) and \( X_{i+j}(\ell) \),

Finally, it is noted that \( h_k(\ell) \) is obtained from \( h_j = h_j(1) \) by the following expression (see Section 5.1), i.e.

\[
h_k(\ell) = \sum_{j=1}^{\ell-1} [j - (k-1)\ell] \cdot h_j + \sum_{j=\ell+1}^{k\ell-1} [(k+1)\ell - j] \cdot h_j
\]

\[
\ell + 2 \sum_{j=1}^{\ell-1} (\ell - j) \cdot h_j
\]

It is seen that a triangular weight is put on the single cell correlation around the ones with distance \( k \ell \).

2.2 Index of Dispersion

The number of generated cells in the interval \([0, t)\) is denoted \( N(t) \). The index of dispersion for this process, \( D(t) \), is defined by (see e.g. [2])

\[
D(t) = \frac{\text{Var}(N(t))}{E(N(t))}
\]

This provides an alternative way of describing the dependence in the traffic. For any process without multiple arrivals, \( D(0) = 1 \). Further, for any stationary process \( E(N(t)) \) is directly proportional to \( t \), i.e. \( E(N(t)) = \mu t \). If in addition the number of arrivals in successive intervals are independent (given a Poisson process) also \( \text{Var}(N(t)) = \mu t \), and \( D(t) = 1 \). However, if the number of arrivals in successive intervals are positively correlated, \( \text{Var}(N(t)) \) obviously becomes \( > \mu t \), and \( D(t) \) is an increasing function. Basic properties of the index of dispersion are discussed for instance in [2, 10, 14].

In [12] it is focused on the fact that \( D(t) \) provides information on the correlation structure of the process. This paper provides interesting results on how to identify the range of timescales for which two processes are 'similar'. Also the discussion of the index of dispersion in the present paper will focus on the close connection to the correlation structure of the process. It will be seen that \( D(t) \) incorporates the total information about the correlations existing between the number of arrivals in nonoverlapping intervals. Thus, essentially \( D(t) \) represents an alternative and compact way of presenting this information about the correlation structure.

2.3 Relation between serial correlation and Index of Dispersion

The present section provides a relation between the index of dispersion, \( D(t) \), for time instants \( t = k \cdot \ell \) \( (k = 1, 2, \ldots) \), and the serial correlation of the arrival process, i.e. \( h_j(\ell) \). Thus, observe that the time axis is now discretized, considering time instants being a multiple of \( \ell \). Main results only, are presented here; see [5] for details. We start from the obvious relation,

\[
N(k \ell) = X_1(\ell) + X_2(\ell) + \cdots + X_k(\ell)
\]

and calculate the variance of this sum. Observing that \( \text{Var}(X_i(\ell)) = D(\ell) \cdot E(X_i(\ell)) \) and that \( E(N(k \ell)) = k \cdot E(X_i(\ell)) \), it follows by rather straightforward calculations that

\[
D(k \ell) = D(\ell) \cdot [1 + \frac{2}{k} \sum_{j=1}^{k-1} (k-j) h_j]\]

\[
k = 2, 3, \ldots
\]

Thus the index of dispersion at a time \( t = k \cdot \ell \) \( (k = 2, 3, \ldots) \) is given by the index of dispersion at time \( \ell \) and a weighted sum of all correlation coefficients, \( h_1(\ell), h_2(\ell), \ldots h_{k-1}(\ell) \).

Typically, the \( h_j(\ell) \) will decrease gradually to 0. It is not simple to draw precise conclusions on \( h_j(\ell) \) from the behaviour of \( D(k \ell) \).

In particular, the point in time where the correlations \( h_j(\ell) \) reach 0 is not easily seen from \( D(\ell) \). However, the \( h_j(\ell) \) can actually be obtained from the index of dispersion \( D(\ell) \) by the following relation (see [5] for details):

\[
h_k(\ell) = \frac{(k+1) \cdot D((k+1)\ell) - 2k \cdot D(k \ell) + (k-1) \cdot D((k-1)\ell)}{2D(\ell)}
\]

Thus, a 'one-to-one' relation between the index of dispersion and the correlation structure has been demonstrated: When the \( h_j(\ell) \), \( j < k \) are given (with fixed value of \( \ell \)), then \( D(\ell) \) can be calculated in the points \( t = j \ell \), \( j \leq k \), using equations (2). Note that this particular set of \( h_j(\ell) \) will provide \( D(\ell) \) for \( \ell \) being a multiple of \( \ell \) only. Further, knowledge of \( D(\tau) \) for \( \tau \leq t \) gives full
information of all the correlations \( h_j(t) \) satisfying \( j \ell < t \). Thus, as also pointed out by [12], \( D(r), \tau \leq t \) provides the correlation structure related to intervals within a distance of \( t \).

3 Single source model

The model of a single source is state oriented and is split into two levels. The first of these is called active-passive level (also called the burst level), with sojourn times in the range of 100 ms. At this level the source is described by a finite state model. For each of these states, cells are generated according to a rule which may be different for the various states. These rules constitute the cell level models. A cell duration is about 3 \( \mu s \). Detailed models for these levels are introduced in Sections 3.1 and 3.2 respectively. This model is a generalization of the previously presented models [5, 4]. A third level, with longer sojourn times than the active-passive level, representing either the alternating sides in a conversation, or the connection and disconnection of a virtual circuit may also be included and accounted for. This is discussed further in Section 4.6.

In the model and the subsequent derivations, the time, denoted \( t \), is regarded as continuous at the active-passive level, while it is regarded as discrete at the cell level with one cell duration as unit. This is done for computation efficiency and has negligible effect since sojourn times at the active-passive level are extremely long compared to a cell-length.

3.1 Active-passive level

At the active-passive level the source is modelled as a stochastic process with \( M \) states, \( \{1, 2, \cdots M\} \). Given that the source is in a specific state, \( i \), a rule for the generation of cells is known, see Section 3.2. The \( M \) states might be divided into two subsets \( A \) and \( P \). For states belonging to the set

\[
A = \{1, 2, \cdots M_1\}
\]

the source generates cells, but when the source is in any of the states

\[
P = \{M_1 + 1, \cdots M\}
\]

there is no generation. Thus \( A \) and \( P \) are the sets of 'active' and 'passive' states respectively.

A Markov model is applied. The transition rate from state \( i \) to state \( j \) is denoted \( w_{ij} \), see Figure 2. A generalization to a phase type distribution for the sojourn times of the generation rules is also treated, see Section 3.3.

3.2 Cell level

There is a cell level model associated with each state at the active-passive level. In the passive states no cells are generated. A cell is generated in the first slot of an active state. Thereafter cells are generated according to a specific rule as long as the source is in this state. Denote this state \( i \). The \( \nu \)th cell interarrival time is denoted \( L_{i, \nu} \), where \( L_{i, 0} = \infty \) for a passive state and \( L_{i, 0} = 0 \) for an active state. Two cases are regarded:

![Figure 2: Transitions between the states of the source process](image)

**Determistic** The cell interarrivals follow a deterministic sequence. We consider two subcases:

- **Infinite** The sequence has no repetitive pattern. It must however be stationary in the sense that the various interarrivals must not be steadily increasing or decreasing and all values are in the same range. Hence, it is required that the sum of \( m \) subsequent interarrival times starting from any point in the sequence, tends to the same average value, \( \bar{L}_i \), with the order of \( \sqrt{1/m} \), i.e.

\[
\frac{\sum_{k=1}^{m} L_{i, \nu + k}}{m} \sim \frac{1}{\sqrt{m}}, \quad \forall \nu \geq 0
\]

- **Periodic** The sequence has a pattern which repeats itself every \( K_i \)th cell, i.e. \( L_{i, \nu + K_i} = L_{i, \nu} \), \( \forall \nu \geq 1 \). The length of the pattern is \( L_i = \sum_{\nu=1}^{K_i-1} L_{i, \nu} \) and the mean interarrival length \( \bar{L}_i = L_i/K_i \). If \( L_i \) becomes similar to or longer than the average state sojourn time on the active-passive level, the "stationarity" requirement above must be obeyed. The periodic case is illustrated in Figure 3.

**Renewal** The cell interarrivals are i.i.d. from a distribution \( r_{i, \nu} = \mathbb{P}(L_{i, \nu} = 1) \), \( \forall \nu \geq 1 \) on the slotted time axis. As for the deterministic processes the mean interarrival time is denoted \( \bar{L}_i = E(L_{i, \nu}) \).

![Figure 3: Example of a periodic deterministic cell level rule](image)

3.3 Generalization of sojourn times

In the formulation of the model presented in Section 3.1, there are \( M = M_1 \) passive states with exponential sojourn times. Since
these $M - M_1$ states are identical with respect to generation of cells (there is actually no cell generation), the introduction of more than one such state is merely used to specify more general (phase type) sojourn time distributions, the actual distribution being dependent on the first passive state.

Similarly, we may specify a general phase type distribution for the duration of a specific rule for cell generation. Now, divide $A$ into disjoint sets, $A_1, A_2, \ldots$, and assume that all the states within a specific set $A_\eta (\eta = 1, 2, \ldots)$ have the same rule concerning generation of cells. Thus, each set of states define a phase type distribution for the (uninterrupted) duration of a specific rule for cell generation.

Note that by this generalization of the model, we may introduce dependencies between the sojourn times of the different sets $P, A_\eta (\eta = 1, 2, \ldots)$. See Figure 4 for a simple example.

![Figure 4: Example of a source model having alternating non-exponential and positively correlated active and passive periods.](image)

### 4 Derivation of the serial correlation

Consider the model and notation is as defined in Section 3. In the present section the objective is to obtain an expression for the correlation between the traffic in slots of various distances. It is assumed that the process has reached stationary conditions. Let the indicator variables, $I_t$, be defined by

$$I_t = \begin{cases} 1 & \text{if a cell is generated in the slot at time } t \\ 0 & \text{otherwise} \end{cases}$$

Consider an arbitrary time instant (slot) $t$. The probability that a cell is generated at this instant is given by (4), and thus the variance of $I_t$ is given by (5).

$$Q = E(I_t)$$
$$Var(I_t) = Q \cdot (1 - Q)$$

Next, to determine the covariances of the $I_t$’s, we need the probability

$$q_t = P \{ \text{generate a cell in slot } \tau + t \mid \text{cell generated in slot } \tau \}$$

Using (4) it now follows that covariances of the $I_t$’s are given by

$$H_t = Cov(I_t, I_{t+1}) = Q \cdot (q_t - Q)$$

Further, the correlations equal

$$h_t = \frac{Cov(I_t, I_{t+1})}{Var(I_t)} = \frac{q_t - Q}{1 - Q}$$

Thus second order properties of the process will be given directly from $q_t$. These will be obtained in sections 4.3 and 4.4. Of course $q_0 = 1$. Further, as $t \to \infty$, the probability $q_t$ of generating a cell in slot $\tau + t$ becomes independent of the given condition that there is a cell in slot $\tau$. Thus $q_t \to Q$ as $t \to \infty$ and so $h_t \to 0$.

Experiences from previous simulation studies and modelling efforts, e.g. [7, 5], indicate that it is the correlation caused by cells generated within the same state occupancy which gives the dominating contribution to the overall correlation. Hence, we derive exact results for the quantities related to this state occupancies and approximate results for the contributions to the correlation due to subsequent state occupancies. Symbols relating to the initial state occupancy are marked with an asterisk, $*$.

#### 4.1 Active-passive level

Now consider the calculation of the probability that the source is in a state $i$ at time $t$, given that the source generated a cell at time $t = 0$. Let $k$ denote the initial state occupied at time $t = 0$, and let $S(t)$ denote the state occupied at $t$. Until the first transition of state occurs, it will be known that the generation of cells follows the original rule of the initial state $k$. Therefore, it should be specified whether the source is still in the initial state, or whether at least one transition has occurred. Thus, the following state probabilities are defined

$$\pi_{ki}(t) = P \{ S(t) = i, \exists \tau \in [0, \ldots, t] \mid S(\tau) \neq k \mid S(0) = k \}$$

It is observed that the total probability of being in state $k$ at time $t$ equals $\pi_{kk}(t) + \pi_{ki}(t)$. The total probability of being in any state $i \neq k$ equals $\pi_{ki}(t)$. Obviously

$$\pi_{ki}(t) = e^{-\omega_{ki} t}$$

By standard Markov analysis, it is found that the remaining state probabilities are obtained from the equations

$$\pi'_{ki}(t) = -\omega_{ii} \pi_{ki}(t) + \sum_{j \neq i} \omega_{ij} \pi_{kj}(t) + \omega_{ki} \pi_{ki}^*(t), \quad i = 1, \ldots, M$$

The equations (9) are now combined with the result (8) and the initial condition $\pi_{ki}(0) = 0, \forall k, i$ and may be solved to give the state probabilities $\pi_{ki}(t)$.

#### 4.2 Cell level

Denote the event that a cell is generated at $t$ by $C(t)$. We then need the following

$$b_{i,t} = P \{ C(t) \mid C(0), \exists \tau \in [0, \ldots, t] \mid S(\tau) \neq i, S(t) = i \}$$

First, an approximation for $b_{i,t}$ is given, and thereafter expressions for $b_{*,t}$ in the various cases.
Since it is assumed that all phase information relative to the cell at \( t = 0 \) is lost in subsequent states, and that \( T_i \) is very short compared to the sojourn time of state \( i \), the probability of finding a cell at time \( t \) is proportional to the average cell generation rate. Hence,
\[
b_i,t \approx \frac{1}{\overline{T}_i} \tag{10}
\]
where \( \overline{T}_i \) are defined for the various cases in Section 3.2. If \( \overline{T}_i \) becomes comparable to the sojourn time, (10) is easily modified to give the correct average rate [6].

In a deterministic infinite sequence, the probability of choosing the \( n \)'th cell at time \( 0 \), \( c_{i,n} \), is approximately proportional to the probability density function of the backward recurrence time at the active-passive level. Since backward and forward recurrence time distributions are equal, we may use \( \pi^*_i(t) \) from (8). From this starting point, the \( L_{i,n+\nu} \) must sum up to \( t \) in order to experience a cell at this time. Hence,
\[
b_{i,0} = 1 \quad \tag{11}
\]
\[
b_{i,t} = \sum_{\nu=0}^{\infty} c_{i,n} \cdot I \left( \exists j \in \{1, \ldots, t\} \text{ so that } \sum_{\nu=0}^{j-1} L_{i,n+\nu} = t \right) \tag{12}
\]
where \( C \) is a normalizing constant, and \( I(\cdot) \) is the indicator function. If the sequence has no regularities, it is reasonable to use the approximation \( b_{i,t} \approx 1/\overline{T}_i \).

For the deterministic periodic sequence we assume that \( L_i \) is short compared to the sojourn time of state \( i \), and, as a result of this, the cell at \( 0 \) is chosen approximately uniform within the pattern with probability \( 1/K_i \). Hence, the probabilities are obtained by finding the average relative positions in the pattern.
\[
b_{i,t} = \frac{1}{K_i} \sum_{n=0}^{K_i-1} I \left( \exists j \in \{n, \ldots, K_i + n - 1\} \text{ so that } \sum_{\nu=0}^{j-1} L_{i,n+\nu} = t \right) \tag{13}
\]
where the mod operation shifts the actual time \( t \) within the first \( L_i - 1 \) cell lengths from time zero, and \( I(\cdot) \) denotes the indicator function as above. See Figure 5 for an illustration. For short sojourn times, (13) are easily enhanced as in the deterministic infinite case [6].

---

**Figure 5:** Example showing one term in the equation determining \( b_{i,t}^* \) for the deterministic periodic rule.

When cells are generated according to a renewal process, \( b_{i,t}^* \) is equal to the renewal density of an ordinary process. From [3] we have
\[
b_{i,t}^* = \sum_{n=1}^{t} [r_{i,n}]^n \tag{14}
\]
where \( n \) denotes \( n \)-fold convolution. The above equation may become computational demanding if an analytical solution is not found. Note however that the number of terms in the sum will be reduced by a factor \( z \) if we require that \( L_i,n \geq z \). Furthermore, \( b_{i,t}^* \) will tend to its stationary value \( 1/\overline{T}_i \) with a speed depending on the standard deviation of the distribution of \( L_i,n \). For sufficiently large \( t \), and no periodicity in \( r_{i,n} \), the approximation \( b_{i,t}^* \approx 1/\overline{T}_i \) suffices. E.g. for a cell process with a coefficient of variation of 1, this approximation is reasonable when \( t > 3\overline{T}_i \).

### 4.3 Probability of cell generation

Given that the process at time \( t = 0 \) is in state \( k \), the probability of generation of a cell at time \( t \), \( P(C(t) \mid S(t) = k, C(t)) \), equals
\[
q_{k,t} = b_{k,t}^* \pi^*_k(t) + \sum_{i=1}^{M} b_{i,t}^* \pi^*_i(t), \quad k = 1, 2, \ldots, M \tag{15}
\]
Actually \( q_{k,t} = 0 \) for \( k > M_1 \). Now consider the initial conditions for the transient process under investigation. The state at time \( t = 0 \) actually equals the state at an 'arbitrarily' chosen instant of the stationary source process, imposing the restriction that a cell is generated at this instant. In order to find the state probability distribution at time \( t = 0 \), introduce
\[
\pi_i = \lim_{t \to \infty} \pi^*_{k(t)} \quad b_i = \lim_{t \to \infty} b_{i,t}^* \quad Q_i = \lim_{t \to \infty} \pi^*_{i(t)} \quad Q = \sum_{i=1}^{M} Q_i \quad \alpha_k = Q_i/Q_k, \quad i = 1, 2, \ldots, M \tag{16}
\]
Here \( Q_i \) is the probability that the stationary process is in state \( k \) at an arbitrary instant in time and that a cell is generated. Observe that \( Q_i = 0 \) for \( i > M_1 \) (i.e. for the passive states) and that the cell generation probability defined in (4) is now determined by the above relations. By normalization, we obtain the probabilities, \( \alpha_k \) of \( k \) being the initial state at time \( t = 0 \) of the transient process. Hence, \( \alpha_k = P(S(0) = k, C(0)) \). Thus, the overall probability of a cell being generated at time \( t \) equals
\[
q_t = \sum_{k=1}^{M} \alpha_k q_{k,t} = \sum_{k=1}^{M} \alpha_k q_{k,t} \tag{17}
\]
where the \( q_{k,t} \) are given by equation (15).

### 4.4 Generalization of sojourn times

Now consider the more general model described in Section 3.3. The total set of active states, \( A \), is divided into disjoint subsets, \( A_1, A_2, \ldots \), so that the rule for cell generation is the same for all states within a subset. Let \( A(t) \) denote the subset occupied at \( t \). The initial state at time \( t = 0 \) belongs to the set \( A_0 \). Now the probabilities \( \pi^*_i(t) \) and \( \pi^*_k(t) \) no longer provide the desired information from the process. Presently it is not important whether the initial state \( k \) has been left. The key point is whether the process has left...
the initial subset of states, \( A_n \), occupied at time \( t = 0 \). Thus, for any \( k \in A_n \) the previous definitions of the state probabilities are modified as:

\[
\pi_{ki}(t) = P \{ S(t) = i, \exists \tau \in [0, \cdots , t] A(\tau) \neq A_n \mid S(0) = k \in A_n \}
\]

For a given \( k \in A_n \) these probabilities are obtained from the following state equations

\[
\pi_{ki}(t) = -\omega_{ii} \pi_{ki}(t) + \sum_{j \in A_n} \omega_{ji} \pi_{kj}(t), \ i \in A_n \quad (18)
\]

\[
\pi'_{ki}(t) = -\omega_{ii} \pi_{ki}(t) + \sum_{j \in A_n} \omega_{ji} \pi_{kj}(t)
\]

\[+ I(i \notin A_n) \sum_{j \in A_n} \omega_{ji} \pi_{kj}(t), \ i = 1, 2, \cdots , M \quad (19)
\]

By these two sets of equations, the state probabilities are obtained for any initial state within the set \( A_n \). Here the \( \pi_{ki}(t) \) are first found from equation (18).

Nearly all the cell level results in Section 4.2 holds for the model with generalized holding times, simply replacing index \( j \) with index \( k \). The only exception is the backward (forward) recurrence times in (12), which should be replaced by:

\[
c_{\eta,n} \approx C \cdot \sum_{k \in A_n} \alpha_k \sum_{i \in A_n} -d \frac{d \pi_{ki}(t)}{dt} \bigg|_{t = \sum_{n=0}^{n} L_{n\cdot}} \quad (20)
\]

Where \( C \) is a normalizing constant as before, and \( \pi_{ki}(t) \) is given by (18) and \( \alpha_k \) by (16).

Now the equation (15) for the probability of generation of a cell at time \( t \), given state \( k \) at time 0 is replaced by

\[
q_{kt} = b^n_{kt} \sum_{i \in A_n} \pi_{ki}(t) + \sum_{c \in A_c} b^{c,t}_{ki} \sum_{i \in A_n} \pi_{ki}(t), \ k \in A_n \quad (21)
\]

The asymptotic distribution \( \pi \), is also obtained from (19) by inserting \( \pi'_{ki}(t) = \pi_{ki}(t) = 0 \). This distribution is independent of \( k \), and the corresponding \( \alpha_k \)'s follow from (16). The total probability \( q_t \) of a cell being generated at time \( t \) is found by inserting these \( \alpha_k \) and the result (21) into equation (17).

4.5 Comparison with simulations

The accuracy of the approximation is investigated [6]. It is good as long as there are no deterministic sojourn times, which would require an infinite number of states in our model. Figure 6 shows the results from a simple example source, having a passive state, 1, an ordinary active state, 2, and a burst state, 3. For exemplification, short sojourn times (and cell patterns) are used. Hence, enhanced versions of (10) and (13), taking this into account, is used for the comparison.

4.6 Generalization to three levels

The results in the previous section can be extended to traffic sources which alternate between being 'connected' and 'disconnected' or being the 'talking' and 'listening' part in a dialogue. Assume that the connect periods, \( C_k \), are random variables having means \( E(C_k) = c \). It is assumed that \( c \gg \omega_{ii}^{-1} \). Similarly for the disconnect periods, which have the mean \( d \). \( C^* \) is the recurrence time of the connect period, which is easily obtained from its distribution [3].

Considering one traffic source only, the following approximate results can be derived for this new situation [5]:

\[
E(I_t) = Q^{od} = \frac{c}{c + d} Q 
\]

\[
h_t^{od} \approx q_t \cdot P(C^* \geq t) - Q^{od} 
\]

The probabilities \( q_t \) are as given in (17). The condition for applying this approximation is that there is a negligible contamination between traffic of different connect periods, which follows from the assumption \( c, d \gg \omega_{ii}^{-1} \).

5 The cell stream from many sources

In Section 4 first and second order moments of the cell stream \( I_t \) from a single source was derived. In the present section generalizations and applications of these results to a cell stream generated by a number of sources, are given.

Consider the case that there are \( S \) sources of traffic, with different distributions for the sojourn times at the active-passive level, and having different rules for generation of cells. Further, the sources are assumed to be mutually independent. Thus there is no correlation between traffic of different sources. The various variables are illustrated in Figure 1. The results of Section 4 are valid for any specific source \( s \) (\( s = 1, 2, \cdots, S \)). The previous notation is maintained. However, all parameters \( L, Q, H_t, \cdots \) of Section 4 are given a superscript \( (s) \) to identify the various sources. The total number of generated cells in slot \( t \) now equals

\[
Y_t = \sum_{s=1}^{S} I_t^{(s)}
\]
It follows that \( E(Y_t) = \sum_{i=1}^{S} Q^{(0)} \), \( \text{Var}(Y_t) = \sum_{i=1}^{S} Q^{(0)}(1 - Q^{(0)}) \),
\( \text{Cov}(Y_{t'}, Y_{t''}) = \sum_{i=1}^{S} H^{(0)}_i \). Here the actual expressions for \( Q^{(0)} \)
and \( H^{(0)}_i \) are given by (4), (16) and (6) respectively.

5.1 Volume of cells

First consider the traffic from a specific source over an interval of length \( t \), \( N^{(0)}(t) = \sum_{i=1}^{S} I^{(0)}_i \). Here \( E(N^{(0)}(t)) = \sum_{i=1}^{S} E(N^{(0)}_i(t)) \), \( \text{Var}(N^{(0)}(t)) = \sum_{i=1}^{S} \text{Var}(N^{(0)}_i(t)) \). Further, the total traffic from \( S \) independent traffic sources (in an interval of length \( t \)) equals
\[
N(t) = \sum_{s=1}^{S} N^{(0)}(t)
\]
Thus, the mean and variance of \( N(t) \) follows directly \( E(N(t)) = \sum_{s=1}^{S} E(N^{(0)}(t)) \), \( \text{Var}(N(t)) = \sum_{s=1}^{S} \text{Var}(N^{(0)}(t)) \).

Now consider the correlation between the traffic in intervals of length \( t \) equals \( X^{(0)}_k(t) = N^{(0)}(kt) - N^{(0)}((k-1)t) \). The covariances between these variables are denoted
\[
H^{(0)}_k(t) = \text{Cov}(X^{(0)}_k(t), X^{(0)}_{k+l}(t))
\]
By letting \( k = 1 \), we get \( t = k \), and thus the previously introduced covariances equal \( H^{(0)}_k(t) = H^{(0)}_0(t) \). Rather straightforward calculations give, cf. (1),
\[
H^{(0)}_k(t) = \sum_{\ell=(k-1)t+1}^{kt} \sum_{\ell=(k+1)t+1}^{(k+1)t} \left( (t - (k-1)t)H^{(0)}_0 \right) + \sum_{\ell=(k+1)t+1}^{(k+1)t} \left( (k+1)t - t \right) H^{(0)}_0
\]
The \( H^{(0)}_k(t) \) in this formula are the covariances (for source \( s \)) directly given by (6).

The total traffic from all sources equals \( X_k(t) = \sum_{s=1}^{S} X^{(0)}_s(t) = N^{(k)}(t) = N^{((k-1)+1)t} \), and the covariances of these are given by the above result:
\[
H_k(t) = \text{Cov}(X_k(t), X_{k+l}(t)) = \sum_{s=1}^{S} H^{(0)}_k(t)
\]
As \( X_k(t) = N(t) \), mean and variance of \( X_k(t) \) directly follow from the results previous in this subsection. In particular,
\[
\text{Var}(X_k(t)) = H_0(t) = \sum_{s=1}^{S} \left( \ell Q^{(0)}(1 - Q^{(0)}) + 2 \sum_{s=1}^{\ell-1} (\ell - t) H^{(0)}_s \right)
\]
The correlation coefficients are now found from \( h_k(t) = H_k(t) / H_0(t) \). By inserting \( t = 1 \) in the above formulas, they reduce to the results for the \( Y_1 = X_1(t) \) above. By the present results, also the index of dispersion of the arriving traffic is found from the relations given in Section 2, see equation (2).

5.2 Distribution of the traffic volume

The mean and variance of the total traffic, \( N(t) \), over an interval \( t \) was obtained in Section 5.1. Thus, the parameters
\[
\mu = E(N(t))/t, \quad \sigma^2 = \text{Var}(N(t))
\]
are given by these results.

Further, asymptotically, as the number of traffic sources, \( S \), approaches infinity, the traffic over one slot, \( Y_r \), becomes normally distributed. Thus, asymptotically, \((Y_1, Y_2, \cdots, Y_t)\) has a multinormal distribution, and also \( N(t) = \sum_{r=1}^{t} Y_r \) has asymptotically a normal distribution. It should be noted that for a finite \( S \) this will not necessarily give a good approximation for the distribution of the \( Y_r \)'s and \( N(t) \).

However, the normal distribution might be applied for \( N(t) \) even if it gives a rather bad fit for the \( Y_r \)'s. Under the assumption that \( N(t) \) is approximately normally distributed, the use of (26) will provide useful results on the traffic over an interval of finite length. Now let \( \Phi(.) \) be the distribution function of the standard normal distribution (having mean 0 and variance 1). Then
\[
p_t = P\{N(t) > K_t\} = 1 - \Phi\left(\frac{K_t - \mu}{\sigma_t}\right)
\]
\[
E(N(t)|N(t) > K_t) = \mu + \frac{\sigma_t}{\sigma}\left(1 - \frac{1}{2}\phi(K_t - \mu)\right)
\]
Observe that these expressions for \( p_t \) and \( E(N(t)|N(t) > K_t) \) depend on \( \mu, \sigma_t \) essentially through the (squared) coefficient of variation of \( N(t) \), i.e. \( \gamma^2 = \text{Var}(N(t))/E(N(t)) = \sigma_t^2/(\mu^2) = D(t)/\mu^2 \). For large \( K_t = t - K \), \( \gamma^2 \) will depend on \( K \) and \( \mu \) through 'traffic intensity' \( \rho = K/\mu \). Details are found in [5], which also presents the mean and variance of the 'traffic overflow',
\[
W(t) = \begin{cases} N(t) - K_t & \text{if } N(t) > K_t \\ 0 & \text{otherwise} \end{cases}
\]

5.3 Peak duration

Now consider
\[
T_p = \text{The length of an uninterrupted period with } Y_r > K,
\]
as illustrated in Figure 7. Thus, \( T_p \) is the 'peak duration', being related to the 'overflow' discussed in the previous section. In principle, the distribution of \( T_p \) is given from the joint distribution of the \( Y_r \). Introducing \( f(y_1, y_2, \cdots, y_t) \) as the joint multinormal distribution, we define
\[
\Psi(t) = \sum_{y_1 > K, y_2 > K, \cdots, y_t > K} f(y_1, y_2, \cdots, y_t) dy_1 dy_2 \cdots dy_t
\]
Then the distribution of \( T_p \) equals
\[
P\{T_p \geq t\} = \int_{y_1 \leq K} \cdots \int_{y_t \leq K} \Psi(t) = \frac{\Psi(t) - \Psi(t + 1)}{\Psi(1) - \Psi(2)}
\]
The integral $\Psi(t)$ of dimension $t$ is in general not easy to calculate. Such calculations are discussed in Chapters 35 and 36 of [8]. There it is concluded that the multidimensional integrals of the normal distribution are rather difficult to handle unless the dimension, $t$, is rather small (up to 4 - 5). Therefore, the calculation of the exact distribution of $T_p$ is not pursued in the present paper. However, expressions for mean and variance are easily obtained:

$$E(T_p) = \sum_{t=1}^{\infty} P \{ T_p \geq t \} = \frac{\Psi(1)}{\Psi(1) - \Psi(2)} \quad (28)$$

and as

$$\sum_{t=1}^{\infty} t P \{ T_p \geq t \} = \frac{E(T^2_p) + E(T_p)}{2}$$

it follows that

$$E(T^2_p) = \frac{\Psi(1) + 2 \sum_{t=1}^{\infty} \Psi(t) - 2 \Psi(2)}{\Psi(1) - \Psi(2)} \quad (29)$$

By calculating the first terms in (29) at least a lower limit for $V ar(T_p)$ can be obtained.

Note that the variable $T_p$ defined above does not necessarily give the most interesting definition of 'peak duration'. If there are frequent fluctuations in the traffic, there might also be frequently occurring (but short) $T^*_p$s. Thus, looking at $E(T_p)$ alone might give a misleading picture (the small $E(T^*_p)$’s indicating satisfactory behaviour of the traffic, even in cases where the traffic is actually very high). Thus, both the mean of $T_p$ and the frequency of these intervals should be calculated in order to obtain a proper understanding of the traffic. Secondly, it might be advisable to 'smooth' the traffic before calculating the mean of $T^*_p$. This could be done, by considering the traffic in intervals of length $\ell$, rather than the traffic in single slots. That is, $Y_s$ should be replaced by $X_s(\ell)$ when calculating the peak duration (the previous discussion based on $Y_s$ corresponds to $\ell = 1$). The calculations for $X_s(\ell)$ can be carried out in exactly the same manner, using the results of Section 5.1. When this is done for various values of $\ell$, a rather good picture of the traffic is obtained.

6 Concluding remarks

This paper provides a set of results useful for the study of ATM traffic. It gives improved insight in the traffic process. The relation between the Index of Dispersion and the correlation structure allows an interpretation of this commonly used index. A simple and useful approximation for the serial correlation of the cell stream for a class of sources are derived. This class has three activity levels with different time constants: a level with alternating connected and disconnected periods, an active-passive (burst) level with possibly dependent and general phase type distributed 'state' durations, and a cell level where deterministic and stochastic cell generation rules are associated with each of the above states. Furthermore, results are derived for a number of important properties of a composite traffic stream based on the correlation in the single streams. These results may help in the buffer dimensioning, policing and resource allocation problems in the design of ATM systems.

Ongoing work supplements the above, by considering changes of the correlation of a stream introduced by FIFO buffering [6].

References


