Tail distribution of the buffer occupancy in a discrete-time queue fed by general on/off sources

Sabine Wittevrongel and Herwig Bruneel
SMACS\textsuperscript{1} Research Group, Laboratory for Communications Engineering, University of Ghent
Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium.

Abstract

A statistical multiplexer with a finite number of independent, not necessarily identical, bursty traffic sources is considered in this paper. Each source stochastically alternates between on-periods, during which it generates one cell per slot, and off-periods, during which no cells are generated. The on-periods and the off-periods are assumed to be independent, but are allowed to have general distributions. The discrete-time queueing model we thus obtain is analyzed using a generating-functions approach resulting in an accurate geometric approximation for the tail distribution of the buffer contents based on the use of the dominant pole of its probability generating function. The results of the study are applied to investigate the influence of the nature of the distributions of the on-periods and the off-periods of the traffic sources on the multiplexer performance.

Keywords: discrete-time queueing theory, ATM, multiplexer performance, general on/off sources, generating-functions approach

1. Introduction

In recent years, discrete-time queueing models have received increasing attention due to their applicability in the analysis of ATM (asynchronous transfer mode) networks, where all information is transmitted in the form of fixed-length packets called cells. ATM networks are capable of providing a wide variety of telecommunication services with different traffic characteristics such as video, voice and high-speed data communications. For this reason, appropriate traffic models need to be considered, which allow to assess the impact of different traffic types with different characteristics on the performance of ATM multiplexers and ATM switches.

\textsuperscript{1}SMACS: Stochastic Modeling and Analysis of Communication Systems
In this paper, we study the performance of a statistical multiplexer, which is fed by a finite number of general heterogeneous on/off sources. Each source alternates between on-periods, during which it generates one cell per slot, and off-periods, during which no cells are generated. The alternating on-periods and off-periods are assumed to be independent, but may have general distributions. This traffic model enables us to characterize the cell streams of different types of ATM sources. It is obvious that the cell arrival process on each link includes a two-fold correlation. A first reason for this correlation is that having an on-period (off-period) of length $n$ implies that a source becoming on (off) stays on (off) during $n$ consecutive slots. A second reason is the finite number of sources $N$, which implies that the number of sources becoming on (off) during a given slot depends on the number of sources already on (off) in the previous slot. By means of a generating-functions approach, an accurate approximation - of geometric form $B \cdot \gamma^S$ - is obtained for the tail distribution of the buffer contents (i.e. the probability that the buffer contents exceed a certain threshold $S$, for sufficiently large values of $S$), which is generally considered to be one of the most important performance measures of ATM multiplexers.

Statistical multiplexers with on/off sources have been studied before, see e.g. [1]-[9], under various assumptions for the distributions of the on-periods and the off-periods. In [1]-[7], the off-periods of the sources are assumed to be purely geometrically distributed, and the on-periods are either geometrically distributed [1]-[4], or geometrically distributed multiples of fixed-length intervals [5], or distributed according to a mixture of 2 geometric distributions [6], or may have a general distribution [7]. The present paper can thus be viewed as an extension of [7], in the sense that the off-periods are now also allowed to follow a general distribution. This makes it possible to investigate the impact of the nature of the distributions of both the on-periods and the off-periods on the queueing performance of the multiplexer. The main difficulty in considering general instead of geometric off-periods lies in the occurrence of a number of boundary probabilities, which in general are difficult to calculate. In this paper, we therefore present an approximation technique, which avoids the calculation of all the unknown boundary probabilities, and at the same time leads to accurate results. We believe that the presented technique constitutes a good compromise between accuracy and computational complexity. General on-periods and general off-periods have also been considered in [8]-[9]. However, in [8], the system is analyzed by numerically solving a set of balance equations, and hence the analysis is limited by the huge state space and the computational complexity of the algorithms. In [9], an analytic geometric approximation is derived for the tail distribution of the buffer contents. However, the exact numerical calculation of the geometric decay rate $\gamma$ is severely limited by the traffic characteristics, and furthermore, the coefficient $B$ is simply approximated by the multiplexer load. Our approximation on the other hand is shown to be much more accurate, and our results are more explicit.

The outline of the paper is as follows. The modeling assumptions regarding the statistical multiplexer and the traffic sources are stated in Section 2. In Section 3, a set of state variables is defined and the system equations are established. A functional equation for the pgf of the system state vector is obtained in Section 4. Section 5 concentrates on the derivation of the tail distribution of the buffer contents from the functional equation. A number of numerical examples are presented in Section 6, in order to verify the accuracy of the analytic approximation, and to investigate the influence of the distributions of the on-periods and the off-periods on the multiplexer performance. Finally, some conclusions are drawn in Section 7.
2. Queueing model

In this paper, we consider a discrete-time queueing model for an ATM statistical multiplexer. The modeling assumptions are as follows:

- The multiplexer has $N$ input links, one output link and an infinite-capacity buffer for the storage of cells.
- Time is divided into fixed-length intervals called slots, such that the multiplexer can transmit exactly one cell from its buffer during each slot. Cells may arrive in the buffer at any time point during a slot. However, the transmission of a cell always starts at the beginning of a slot and ends at the end of this slot. Hence, a cell can never leave the buffer before the end of the slot following its slot of arrival. In other words, when the queueing system is empty at the start of a slot, no cell can leave the buffer at the end of that slot, even if some cells have arrived during the slot.
- There is a finite number $N$ of independent, not necessarily identical, bursty traffic sources. There are $T$ different traffic types. Denoting by $N_t$, $1 \leq t \leq T$, the number of sources of traffic type $t$, we have $N = N_1 + \ldots + N_T$.
- Each traffic source stochastically alternates between on-periods and off-periods. During an on-period, a source generates exactly one cell per slot, whereas no cells are generated during an off-period of a source. The (lengths of the ) on-periods and the off-periods of a source are modeled as independent random variables with general distributions. For a source of traffic type $t$, we denote by $a_t(n)$ and $b_t(n)$ the probability mass functions (pmf’s) of the lengths of the on-periods and the off-periods, respectively, i.e.

\[
a_t(n) = \text{Prob}\{\text{on-period of type } t \text{ source } = n \text{ slots}\};
\]

\[
b_t(n) = \text{Prob}\{\text{off-period of type } t \text{ source } = n \text{ slots}\},
\]

whereas the corresponding pgf’s are denoted as $A_t(z)$ and $B_t(z)$.

- We assume that the queueing system can reach a steady state. This means that the equilibrium condition, being the condition that the mean number of cell arrivals $\rho$ during an arbitrary slot is strictly less than 1, is assumed to be fulfilled.

3. System equations

The goal of this section is to introduce a Markovian state description for the queueing system described above. In order to do so, we first take a closer look at the cell arrival process. As mentioned before, each source alternates between on-periods and off-periods. A source is now called in state $A_n$, $n \geq 1$, if it is in the $n$th slot of an on-period. Similarly, a source is called in state $B_n$, $n \geq 1$, if it is in the $n$th slot of an off-period. Note that exactly two transitions are possible from each state: transition to the same period, but one slot further or transition to the first slot of the other period. Clearly, each source can be characterized by an infinite-dimensional Markov chain,
with states $A_n$ and $B_n$ ($n \geq 1$), and transition probabilities as depicted in Fig. 1. Here $p_{a,t}(n)$ denotes the probability of having an on-period of at least $n+1$ slots, given that this on-period consists of at least $n$ slots. Analogously, $p_{b,t}(n)$ is the probability of having an off-period of at least $n+1$ slots, given that this off-period consists of at least $n$ slots. That is,

$$p_{a,t}(n) = \left(1 - \sum_{i=1}^{n} a_t(i)\right)\left(1 - \sum_{i=1}^{n-1} a_t(i)\right)^{-1}; \quad (1)$$

$$p_{b,t}(n) = \left(1 - \sum_{i=1}^{n} b_t(i)\right)\left(1 - \sum_{i=1}^{n-1} b_t(i)\right)^{-1}. \quad (2)$$

Let us define the random variables $g_{n,k}(t)$ and $h_{n,k}(t)$ ($n \geq 1$, $1 \leq t \leq T$) as the numbers of sources of traffic type $t$ in states $A_n$ and $B_n$ respectively, during slot $k$. From the state transition diagram of Fig. 1, $g_{n,k}(t)$ and $h_{n,k}(t)$ can then be expressed as

$$g_{n,k}(t) = \sum_{i=1}^{g_{n-1,k-1}(t)} c_{n-1,i}(t) , \quad n > 1 ; \quad (3)$$

$$h_{n,k}(t) = \sum_{i=1}^{h_{n-1,k-1}(t)} d_{n-1,i}(t) , \quad n > 1 ; \quad (4)$$

$$g_{1,k}(t) = \sum_{n=1}^{\infty} \sum_{i=1}^{h_{n,k-1}(t)} (1-d_{n,i}(t)) ; \quad (5)$$

$$h_{1,k}(t) = N_t - \sum_{n=1}^{\infty} g_{n,k}(t) - \sum_{n=2}^{\infty} h_{n,k}(t) . \quad (6)$$
The reasoning behind equation (3) is that \( g_{n,k}(t) \) contains one unity for each source of type \( t \) which was in state \( A_{n-1} \) during slot \( k-1 \) and which changes to state \( A_n \) in slot \( k \). The random variable \( c_{n-1,k}(t) \) in equation (3) takes on the values 0 and 1, and equals 1 if and only if the \( i \)th source of type \( t \) in state \( A_{n-1} \) during slot \( k-1 \) remains in an on-period during the next slot. This happens with probability \( p_{a,i}(n) \), as can be seen from Figure 1. A similar reasoning holds for equations (4)-(6). Hence, we have that \( \{c_{n,i}(t), i \geq 1\} \) and \( \{d_{n,i}(t), i \geq 1\} \) \( (n \geq 1, 1 \leq t \leq T) \) are independent sets of i.i.d. Bernoulli random variables with p.g.f’s

\[
C_{n,j}(z) = 1 - p_{a,j}(n) + p_{a,j}(n) z
\]

and

\[
D_{n,j}(z) = 1 - p_{b,j}(n) + p_{b,j}(n) z ,
\]

respectively.

Let us now denote by \( g_n(t) \) and \( h_n(t) \) the steady-state versions of \( g_{n,k}(t) \) and \( h_{n,k}(t) \). Since all the traffic sources are assumed to be independent, it is easily seen that the joint p.g.f \( N(x_1, \ldots, x_T, y_1, \ldots, y_T) \) of the \( g_n(t) \)’s and the \( h_n(t) \)’s can be expressed as

\[
N(x_1, \ldots, x_T, y_1, \ldots, y_T) = \mathbb{E} \left[ \prod_{t=1}^{T} \left( \prod_{n=1}^{\infty} x_{n,j} g_{n,j}(t) \right) \left( \prod_{n=2}^{\infty} y_{n,j} h_{n,j}(t) \right) \right] \\
= \prod_{t=1}^{T} \left( v_{b,t}(1) + \sum_{n=1}^{\infty} v_{a,t}(n) x_{n,t} + \sum_{n=2}^{\infty} v_{b,t}(n) y_{n,t} \right)^{N_t} ,
\]

where the vectors \( x_t = (x_{1,t}, x_{2,t}, \ldots) \) and \( y_t = (y_{2,t}, y_{3,t}, \ldots) \), and \( v_{a,t}(n) \) and \( v_{b,t}(n) \) are the probabilities that a source of type \( t \) is in state \( A_n \) and \( B_n \) respectively, during an arbitrary slot in the steady state. From the set of balance equations for the Markov chain in Fig. 1, together with the normalization equation, the probabilities \( v_{a,t}(n) \) and \( v_{b,t}(n) \) can be calculated as

\[
v_{a,t}(n) = v_{a,t}(1) \left( \sum_{i=n}^{\infty} a_t(i) \right) , \quad n \geq 1 ;
\]

\[
v_{b,t}(n) = v_{b,t}(1) \left( \sum_{i=n}^{\infty} b_t(i) \right) , \quad n \geq 1 ;
\]

\[
v_{a,t}(1) = v_{b,t}(1) = \frac{1}{A_t(1) + B_t(1)} .
\]
From equation (9), it is easy to show that the total average multiplexer load \( \rho \) is given by

\[
\rho = \sum_{t=1}^{T} N_t \sigma_t = \sum_{t=1}^{T} N_t \frac{A_t'(1)}{A_t'(1) + B_t'(1)},
\]

(13)

where \( \sigma_t \) is the average load contributed by a source of type \( t \) and \( A_t'(1) \) and \( B_t'(1) \) are the mean lengths of the on-periods and the off-periods of a type \( t \) source, respectively.

Next, let \( s_k \) represent the buffer contents at the beginning of slot \( k \), i.e. the total number of cells stored in the multiplexer buffer at the start of slot \( k \), including the cell that will be transmitted during slot \( k \), if any. The evolution of the buffer contents is governed by the following system equation :

\[
s_{k+1} = (s_k - 1)^+ + e_k,
\]

(14)

where \((\cdot)^+ = \max(0, \cdot)\) and \( e_k \) denotes the total number of cell arrivals during slot \( k \), which can be further expressed as

\[
e_k = \sum_{t=1}^{T} \sum_{n=1}^{\infty} g_{n,t}(t).
\]

(15)

From the above equations, it is clear that the vector \((g_{1,k-1}, \ldots, g_{T,k-1}, h_{1,k-1}, \ldots, h_{T,k-1}, s_k)\), where \( g_{t,k-1} = (g_{1,k-1}(t), g_{2,k-1}(t), \ldots) \) and \( h_{t,k-1} = (h_{2,k-1}(t), h_{3,k-1}(t), \ldots) \), constitutes an infinite-dimensional Markovian state description of the queueing system under study at the beginning of slot \( k \).

4. Functional equation

Let us define the joint pgf of the state vector \((g_{1,k-1}, \ldots, g_{T,k-1}, h_{1,k-1}, \ldots, h_{T,k-1}, s_k)\) :

\[
P_k(x_1, \ldots, x_T, y_1, \ldots, y_T, z) = E \left[ \prod_{t=1}^{T} \prod_{n=1}^{\infty} x_{n,t}^{g_{n,t-1}(t)} \prod_{t=1}^{T} \prod_{n=2}^{\infty} y_{n,t}^{h_{n,t-1}(t)} z^{s_k} \right].
\]

(16)

With this definition, \( P_{k+1}(x_1, \ldots, x_T, y_1, \ldots, y_T, z) \) can be obtained by using equations (14) and (15) as

\[
P_{k+1}(x_1, \ldots, x_T, y_1, \ldots, y_T, z) = E \left[ \prod_{t=1}^{T} \prod_{n=1}^{\infty} (x_{n,t})^{g_{n,t}(t)} \prod_{t=1}^{T} \prod_{n=2}^{\infty} y_{n,t}^{h_{n,t}(t)} z^{(s_k - 1)^+} \right].
\]
Based on equations (3)-(5) and averaging over the distributions of the $c_{n,i}(t)$’s and the $d_{n,i}(t)$’s, defined in equations (7) and (8), it follows that

$$P_{k+1}(\mathbf{x}_1, \ldots, \mathbf{x}_T, \mathbf{y}_1, \ldots, \mathbf{y}_T, z)$$

$$= E \left[ \prod_{t=1}^{T} \prod_{n=1}^{\infty} C_{n,t} \left( x_{n+1,t} \right) g_{n,k-1}(t) \prod_{t=1}^{T} \prod_{n=1}^{\infty} x_{1,t} z D_{n,t} \left( \frac{y_{n+1,t}}{x_{1,t} z} \right)^{h_{n,k-1}(t)} \right] z^{(s_k-1)^+} ,$$

which, upon substitution of equation (6) can be rewritten as

$$P_{k+1}(\mathbf{x}_1, \ldots, \mathbf{x}_T, \mathbf{y}_1, \ldots, \mathbf{y}_T, z) = \prod_{t=1}^{T} x_{1,t} z D_{n,t} \left( \frac{y_{2,t}}{x_{1,t} z} \right)^{N_t}$$

$$\cdot E \left[ \prod_{t=1}^{T} \prod_{n=1}^{\infty} G_{n,t} \left( \mathbf{x}_i, \mathbf{y}_i, z \right) g_{n,k-1}(t) \prod_{t=1}^{T} \prod_{n=2}^{\infty} H_{n,t} \left( \mathbf{x}_i, \mathbf{y}_i, z \right)^{h_{n,k-1}(t)} \right] z^{(s_k-1)^+} , \quad (17)$$

where

$$G_{n,t}(\mathbf{x}_i, \mathbf{y}_i, z) \overset{d}{=} \frac{C_{n,t} \left( x_{n+1,t} z \right)}{x_{1,t} z D_{n,t} \left( \frac{y_{2,t}}{x_{1,t} z} \right)} , \quad n \geq 1 , \quad 1 \leq t \leq T ; \quad (18)$$

$$H_{n,t}(\mathbf{x}_i, \mathbf{y}_i, z) \overset{d}{=} \frac{D_{n,t} \left( \frac{y_{n+1,t}}{x_{1,t} z} \right)}{D_{i,t} \left( \frac{y_{2,t}}{x_{1,t} z} \right)} , \quad n \geq 2 , \quad 1 \leq t \leq T . \quad (19)$$

Since a cell can never leave the buffer before the end of the slot right after its arrival slot, $s_k = 0$ implies that $e_{k-1} = 0$ and hence also $g_{n,k-1}(t) = 0$ $(n \geq 1, \ 1 \leq t \leq T)$. With this property, the right-hand side of the above equation can be further expressed in terms of the function $P_k$. In the steady state, $P_k(\mathbf{x}_1, \ldots, \mathbf{x}_T, \mathbf{y}_1, \ldots, \mathbf{y}_T, z)$ becomes independent of $k$. From equation (17), it is not so difficult to show that the steady-state pgf $P(\mathbf{x}_1, \ldots, \mathbf{x}_T, \mathbf{y}_1, \ldots, \mathbf{y}_T, z)$ satisfies the following functional equation:

$$z P(\mathbf{x}_1, \ldots, \mathbf{x}_T, \mathbf{y}_1, \ldots, \mathbf{y}_T, z) = \prod_{t=1}^{T} x_{1,t} z D_{n,t} \left( \frac{y_{2,t}}{x_{1,t} z} \right)^{N_t}$$

$$\cdot \left\{ P(G_1(\mathbf{x}_1, \mathbf{y}_1, z) \ldots \ldots G_T(\mathbf{x}_T, \mathbf{y}_T, z), H_1(\mathbf{x}_1, \mathbf{y}_1, z) \ldots \ldots H_T(\mathbf{x}_T, \mathbf{y}_T, z), z) \right\}$$

$$+ (z-1) p_0 Q(H_1(\mathbf{x}_1, \mathbf{y}_1, z) \ldots \ldots H_T(\mathbf{x}_T, \mathbf{y}_T, z)) \right\} , \quad (20)$$
where

\[ G_t(x_t, y_t, z) \triangleq \left( G_{1,t}(x_t, y_t, z), G_{2,t}(x_t, y_t, z), \ldots \right) ; \]

\[ H_t(x_t, y_t, z) \triangleq \left( H_{2,t}(x_t, y_t, z), H_{3,t}(x_t, y_t, z), \ldots \right) , \]

for \( 1 \leq t \leq T \), and \( p_0 \) denotes the steady-state probability of an empty buffer. Note that \( G_t(x_t, y_t, z) \) and \( H_t(x_t, y_t, z) \) are only functions of \( x_t, y_t \) and \( z \), which is due to the fact that all the \( T \) different traffic types are assumed to be independent. The \( Q \)-function in equation (20) is defined as

\[ Q(y_1, \ldots, y_T) \triangleq \sum_{j_1, \ldots, j_T} q(j_1, \ldots, j_T) \left( \prod_{t=1}^{T} \prod_{n=2}^{\infty} y_{n,t}^{j_n} \right) , \]

where

\[ q(j_1, \ldots, j_T) \triangleq \lim_{k \to \infty} \text{Prob}\left[h_{n,k-1}(t) = j_{n,t} \ (n \geq 2, \ 1 \leq t \leq T) \mid s_k = 0\right] . \tag{24} \]

Because of the indirect relation between \( h_{n,k-1}(t) \) and \( s_k \), it is very difficult to obtain an exact expression for all the unknown boundary probabilities \( q(j_1, \ldots, j_T) \) and hence, for the function \( Q(y_1, \ldots, y_T) \).

Since the pgf \( S(z) \) of the steady-state buffer contents is equal to \( P(1, \ldots, 1, z) \), in principle, the functional equation fully describes the buffer behavior. Unfortunately, in general, it is not an easy task to obtain from (20) an explicit expression for the pgf \( S(z) \). Nevertheless, an approximation for the tail distribution of the buffer contents can be derived from it. In order to do so, we select only those values of \( x_{n,t}, y_{n,t} \) and \( z \) for which the arguments of the \( P \)-functions on both sides of (20) are identical, i.e. such that

\[ x_{n,t} = G_{n,t}(x_t, y_t, z) \quad \text{and} \quad y_{n,t} = H_{n,t}(x_t, y_t, z) . \tag{25} \]

From these relationships, \( x_{n,t} \) and \( y_{n,t} \) can be solved in terms of \( z \). It turns out that for a given value of \( z \), there may be more than one set of solutions. Only one of these sets, which will be denoted by \( \chi_{n,t}(z) \) and \( \xi_{n,t}(z) \), has the additional property that \( x_{n,t} = 1 \) and \( y_{n,t} = 1 \), for \( z = 1 \). From (25), it is possible to show that \( \chi_{n,t}(z) \) and \( \xi_{n,t}(z) \) satisfy the following set of equations:

\[ z \left( \sum_{i=n}^{\infty} a_i(i) \right) \chi_{n,t}(z) = \left( \frac{F_t(z)}{z} \right)^{n-1} A_t \left( \frac{z}{F_t(z)} \right) \left( \sum_{i=1}^{n-1} a_i(i) \left( \frac{F_t(z)}{z} \right)^{n-i} \right) , \quad n \geq 1 ; \tag{26} \]

\[ \left( \sum_{i=n}^{\infty} b_i(i) \right) \xi_{n,t}(z) = \chi_{1,t}(z) z \left\{ F_t(z)^{n-1} B_t \left( \frac{1}{F_t(z)} \right) - \sum_{i=1}^{n-1} b_i(i) F_t(z)^{n-i} \right\} , \quad n \geq 2 , \tag{27} \]
where

\[
F_t(z) = b_t(1) \chi_{1,t}(z) + (1-b_t(1)) \xi_{2,t}(z).
\]  
(28)

Choosing \( x_{n,t} = \chi_{n,t}(z) \) and \( y_{n,t} = \xi_{n,t}(z) \) in (20) then yields

\[
P(\chi_1(z),...,\chi_T(z),\xi_1(z),...,\xi_T(z),z) = \frac{(z-1) p_0 F(z) \tilde{Q}(z)}{z - F(z)},
\]  
(29)

where the functions \( F(z) \) and \( \tilde{Q}(z) \) are defined as

\[
F(z) = \prod_{t=1}^{T} \left[ \chi_{1,t}(z) D_{1,t} \left( \frac{\xi_{2,t}(z)}{\chi_{1,t}(z)} \right) \right]^{N_t} = \prod_{t=1}^{T} F_t(z)^{N_t};
\]  
(30)

\[
\tilde{Q}(z) = Q(\xi_1(z),...,\xi_T(z));
\]  
(31)

and the unknown constant \( p_0 \) can be calculated from the normalisation condition \( P(1, ..., 1) = 1 \), which yields \( p_0 = 1 - p \). Comparing the results obtained here with those obtained in [7], under the assumption of geometrically distributed off-periods, we see that the main difficulty in considering general instead of geometric off-periods lies in the occurrence of a number of boundary probabilities, which in general are difficult to calculate. In the next section, we therefore present an approximation technique, which avoids the calculation of all the unknown boundary probabilities, and at the same time leads to accurate results for the tail distribution of the buffer contents.

5. Tail distribution of the buffer contents

A performance measure of considerable interest is the probability that the buffer contents exceeds a certain threshold \( S \). This quantity is often used in a buffer model with a finite waiting room of size \( S \), to approximate the cell loss ratio, i.e. the fraction of the arriving cells that is lost upon arrival because of buffer overflow. In many cases, it has been shown that the buffer-contents distribution exhibits a geometric tail behavior. That is, for sufficiently large \( S \), the tail distribution of the buffer contents can be approximated as

\[
\text{Prob}[S > S] \approx -\frac{\theta}{z_0 - 1} \left( \frac{1}{z_0} \right)^{S+1}.
\]  
(32)

Here \( z_0 \) is the pole of \( S(z) \) with the smallest modulus, which must necessarily be real and positive in order to ensure that the tail distribution is nonnegative anywhere, and the constant \( \theta \) is the residue of \( S(z) \) in the point \( z = z_0 \). In the rest of this section, we will now first of all describe how
the pole $z_0$ can be calculated. Next, we will derive an accurate closed-form approximation for the residue $\theta$. 
5.1. The dominant pole $z_0$

As in [5]-[7], it can be argued that the dominant pole $z_0$ of $S(z)$ is also the pole with the smallest modulus of $P(x_1(z),\ldots,x_T(z),\xi_1(z),\ldots,\xi_T(z),z)$. Hence, $z_0$ is determined by the equation $z - F(z) = 0$, or

$$z - \prod_{t=1}^{T} F_t(z)^{N_t} = 0 \ .$$ (33)

From equations (26)-(28), it furthermore follows that

$$A_t \left( \frac{z}{F_t(z)} \right) B_t \left( \frac{1}{F_t(z)} \right) = 1 \ , \ 1 \leq t \leq T \ .$$ (34)

The pole $z_0$ can easily be calculated numerically from the above set of equations by using, for instance, a combination of repeated substitutions and the Newton-Raphson algorithm. From equations (33)-(34), we also note that the unknown boundary probabilities $q(j_1,\ldots,j_T)$ have no influence on $z_0$, which means that the geometric decay rate $\gamma = 1/z_0$ can be calculated exactly.

5.2. The residue $\theta$

Let us consider the situation where the buffer contents just after a given slot is extremely large (>> $N$). It is reasonable to think that the number of cell arrivals during this slot (which cannot be larger than $N$) has almost no impact on the total number of cells in the buffer. If $j$ is sufficiently large ($j > J$), we may therefore assume that the conditional probabilities $\text{Prob}[g_n(t) = i_{n,t} \ (n \geq 1, \ 1 \leq t \leq T), \ h_n(t) = k_{n,t} \ (n \geq 2, \ 1 \leq t \leq T) \ | \ s = j]$ are almost independent of $j$, and approach to some limiting values for $j \to \infty$, denoted by $\omega(i_1,\ldots,i_T,k_1,\ldots,k_T)$, with corresponding joint pgf $\Omega(x_1,\ldots,x_T,y_1,\ldots,y_T)$. This means that for $j > J$,

$$\text{Prob}\left[ g_n(t) = i_{n,t} \ (n \geq 1,1 \leq t \leq T), h_n(t) = k_{n,t} \ (n \geq 2,1 \leq t \leq T) \mid s = j \right] \equiv \omega(i_1,\ldots,i_T,k_1,\ldots,k_T) \ .$$ (35)

The residue $\theta$ can then be derived, in a similar way as explained in [5]-[7], as

$$\theta = \frac{z_0 \left( z_0 - 1 \right) \left( 1 - \rho \right) \bar{Q}(z_0)}{\left[ \left( 1 - F'(z_0) \right) \Omega(x_1(z_0),\ldots,x_T(z_0),\xi_1(z_0),\ldots,\xi_T(z_0)) \right]} \ .$$ (36)
Furthermore, the following equation can be established for the pgf \( \Omega(x_1, \ldots, x_T, y_1, \ldots, y_T) \) of the limiting cell arrival process:

\[
z_0 \Omega(x_1, \ldots, x_T, y_1, \ldots, y_T) = \prod_{t=1}^{T} \left[ x_{1,t} z_0 D_t \left( \frac{y_{2,t}}{x_{1,t} z_0} \right) \right]^{N_t} \cdot \Omega(G_1(x_1, y_1, z_0) \ldots G_T(x_T, y_T, z_0) H_1(x_1, y_1, z_0) \ldots H_T(x_T, y_T, z_0)) \right).
\]

As expected intuitively, it is possible to show that the solution \( \Omega(x_1, \ldots, x_T, y_1, \ldots, y_T) \) of (37) has the same form of expression as the pgf \( N(x_1, \ldots, x_T, y_1, \ldots, y_T) \) of the unconditional cell arrival process. Specifically, \( \Omega(x_1, \ldots, x_T, y_1, \ldots, y_T) \) can be expressed as

\[
\Omega(x_1, \ldots, x_T, y_1, \ldots, y_T) = \prod_{t=1}^{T} \left( \mu_{b,t}(1) + \sum_{n=1}^{\infty} \mu_{a,t}(n) x_{n,t} + \sum_{n=2}^{\infty} \mu_{b,t}(n) y_{n,t} \right)^{N_t},
\]

where \( \mu_{a,t}(n) \) and \( \mu_{b,t}(n) \) are the (conditional) probabilities of finding a source of traffic type \( t \) in state \( A_n \) or \( B_n \), respectively, when the number of cells in the multiplexer buffer is extremely large. Substitution of the expression (38) into equation (37) and identification of the coefficients of \( x_{n,t} \) and \( y_{n,t} \) on both sides of the resulting equation leads to a set of equations from which \( \mu_{a,t}(n) \) and \( \mu_{b,t}(n) \) can be derived explicitly as

\[
\mu_{a,t}(1) = \frac{z_0 \left( z_0 - F_t(z_0) \right) B_t \left( \frac{1}{F_t(z_0)} \right) \left( F_t(z_0) - 1 \right)}{F_t(z_0)^2 \left( z_0 - 1 \right) \left[ 1 - B_t \left( \frac{1}{F_t(z_0)} \right) \right]} \quad ;
\]

\[
\mu_{a,t}(n) = \mu_{a,t}(1) \left( \frac{z_0}{F_t(z_0)} \right)^{-n-1} \left( \sum_{i=n}^{\infty} a_t(i) \right), \quad n \geq 2 \quad ;
\]

\[
\mu_{b,t}(2) = \frac{\left( z_0 - F_t(z_0) \right) \left[ 1 - b_t(1) \right] \left( F_t(z_0) - 1 \right)}{F_t(z_0)^3 \left( z_0 - 1 \right) \left[ 1 - B_t \left( \frac{1}{F_t(z_0)} \right) \right]} ;
\]

\[
\mu_{b,t}(n) = \mu_{b,t}(2) \left[ 1 - b_t(1) \right] \left( \frac{1}{F_t(z_0)} \right)^{n-2} \left( \sum_{i=n}^{\infty} b_t(i) \right), \quad n \geq 3 \quad .
\]
Furthermore, by using (26)-(28), (33) and the above expressions, we obtain

\[
\Omega (\xi_1(z_0), \ldots, \xi_T(z_0) ; \xi_1(z_0), \ldots, \xi_T(z_0)) = \prod_{t=1}^{T} \left[ \frac{[z_0 - F_t(z_0)] [F_t(z_0) - 1]}{(z_0 - 1) F_t(z_0)^3 \left[ 1 - B_t \left( \frac{1}{F_t(z_0)} \right) \right]} \right]^{N_t} \cdot \left[ z_0 A_t \left( \frac{z_0}{F_t(z_0)} \right) B_t \left( \frac{1}{F_t(z_0)} \right) + A_t \left( \frac{z_0}{F_t(z_0)} \right) B_t^* \left( \frac{1}{F_t(z_0)} \right) \right]^{N_t}.
\]

Also, we get the following explicit expression for \( F'(z_0) \):

\[
F'(z_0) = \sum_{t=1}^{T} \left[ z_0 A_t \left( \frac{z_0}{F_t(z_0)} \right) B_t \left( \frac{1}{F_t(z_0)} \right) + A_t \left( \frac{z_0}{F_t(z_0)} \right) B_t^* \left( \frac{1}{F_t(z_0)} \right) \right]^{N_t}.
\]

The only unknown quantity in (36) is \( \tilde{Q}(z_0) \), which contains a very large number of unknown boundary probabilities \( q(j_1, \ldots, j_T) \), defined in (24). Since in general, it is very difficult to calculate these probabilities exactly, we will derive an approximation for \( \tilde{Q}(z_0) \) in stead. Our approximation is based on the observation that implies that \( g_{n,k-1}(t) = 0 \ (n \geq 1, 1 \leq t \leq T) \), if \( s_k = 0 \). For this reason, one might think that the difference between the cell arrival process observed when \( s_k = 0 \) or when \( g_{n,k-1}(t) = 0 \ (n \geq 1, 1 \leq t \leq T) \) respectively is small. Therefore, we approximate the unknown boundary probabilities as

\[
q(j_1, \ldots, \hat{j_T}) = \lim_{k \to \infty} \text{Prob}\left[ h_{n,k-1}(t) = j_{n,t} \ (n \geq 2, 1 \leq t \leq T) \mid g_{n,k-1}(t) = 0 \ (n \geq 1, 1 \leq t \leq T) \right].
\]

Numerical results presented in Section 6 show that the above approximation leads to accurate results. Since the random variables \( g_{n,k-1}(t) \) and \( h_{n,k-1}(t) \) are both related to the cell arrival process, by introducing the above approximation into equation (23), the function \( \tilde{Q}(z) \) can be expressed in terms of the pgf \( N(\xi_1, \ldots, \xi_T, y_1, \ldots, y_T) \) of the unconditional cell arrival process. As a result, the following approximation is obtained for \( \tilde{Q}(z) \):

\[
\tilde{Q}(z) \approx \frac{N(0,0,\tilde{\xi}_1(\tilde{z}), \ldots, \tilde{\xi}_1(\tilde{z}))}{N(0,0,1,\ldots,1)}.
\]
A little more algebra finally yields

$$
\tilde{\theta}(\alpha) \equiv \prod_{t=1}^{T} \left[ 1 + \frac{1}{B_t(1)} \left( 1 + A_t \left( \frac{\alpha_0}{F_t(\alpha_0)} \right) \left( \frac{F_t(\alpha_0)}{B_t} \left( \frac{1}{\alpha_0} \right) - 1 \right) \right) \right]^{N_t} \, .
$$

(47)

Combination of equations (36), (43), (44) and (47) then gives an explicit approximation for the parameter $\theta$ in terms of $\alpha_0$. It has been observed that in the case of geometrically distributed off-periods, the above derived approximation for $\theta$ corresponds to the exact result obtained in [7].

6. Numerical examples

In this section, we illustrate the obtained results by means of two specific examples. In the first example, geometric on-periods and mixed geometric off-periods are considered, whereas the second example is concerned with fixedly spaced on/off sources. The purpose of these examples is first of all, to verify the accuracy of the obtained results and secondly, to investigate the influence of the distributions of the on-periods and the off-periods on the multiplexer performance.

The model we analyzed in this paper has also been investigated by Sohraby [9]. The most important differences between our study and Sohraby’s are as follows:

- The exact numerical calculation of the geometric decay rate $\gamma = 1 / \alpha_0$ by means of the method presented in [9] requires knowledge of the algebraic expression of the Perron-Frobenious (PF) eigenvalue of each of the individual sources, which except for some special cases, is very difficult to obtain. For our method, on the other hand, the expression of the PF eigenvalue is not necessary, and the dominant pole $\alpha_0$ can easily be calculated from (33) and (34) by means of a numerical algorithm.

- Our study not only yields the exact value of the geometric decay rate; it also leads to an accurate closed-form approximation for the coefficient of the geometric form, whereas in [9] this coefficient is simply approximated by the multiplexer load, i.e. in [9], the tail distribution of the buffer contents is approximated as

$$
\text{Prob}[s > S] \approx \rho \, \alpha_0^{-S} \, .
$$

(48)

The approximation (48) is much less accurate than our’s, as we will show in the sequel.

6.1. Geometric on-periods and mixed geometric off-periods

Let us assume that the multiplexer is fed by $N$ homogeneous on/off sources. For each source, the lengths of the on-periods are geometrically distributed with parameter $\alpha$, whereas the
off-periods are distributed according to a mixture of 2 geometric distributions, i.e.

\[ A(z) = \frac{(1 - \alpha) z}{1 - \alpha z} \quad \text{and} \quad B(z) = \frac{q (1 - \beta_1) z}{1 - \beta_1 z} + \frac{(1 - q) (1 - \beta_2) z}{1 - \beta_2 z}, \quad \beta_1 > \beta_2. \]  

(49)

The set of parameters \((N, \alpha, q, \beta_1, \beta_2)\) fully characterizes the cell arrival process to the multiplexer. However, we will use here a different, more intuitive set of parameters. First of all, the total multiplexer load \(\rho\) is given by \(\rho = N/[1 + (1 - \alpha)B'(1)]\). Also, as above, we define the burstiness factor \(K\) of a source as the ratio of the mean on-period (or off-period) of the source in the considered traffic model to the corresponding quantity in the case of a Bernoulli arrival process with the same load \(\sigma = \rho / N\), i.e. \(K = (1 - \sigma)/(1 - \alpha) = \sigma B'(1)\). Where the load \(\sigma\) of a source is a measure for the ratio of the mean lengths of the on-periods and the off-periods, clearly the parameter \(K\) describes the absolute lengths of these periods. Furthermore, we define \(L\) as the ratio of the variance of an off-period in the considered model, to the variance of a geometrically distributed off-period with the same mean length. It can be shown that for given values of \(q, L\) and \(B'(1)\), the parameters \(\beta_1\) and \(\beta_2\) can be derived from

\[ \frac{1}{1 - \beta_1} = B'(1) + \frac{(1 - q) (L - 1) B'(1) [B'(1) - 1]}{2q}; \]  

(50)

\[ \beta_2 = 1 - \frac{(1 - q) (1 - \beta_1)}{B'(1) (1 - \beta_1) - q}. \]  

(51)

Hence, it is clear that the set \((N, \rho, K, L, q)\) also fully describes the cell arrival process.

**Fig. 2**  Tail distribution of the buffer contents Prob\([s > S]\) versus \(S\), for geometric on-periods and mixed geometric off-periods, for \(N = 4, K = 2, L = 2\) and \(q = 0.3\).
In order to check the accuracy of the results derived in the previous section, we have also numerically analyzed the queueing system under the assumption of a finite, but large waiting room. For the considered traffic model, provided that the number of sources $N$ was not too large, the numerical solution of the resulting set of balance equations was feasible, leading to the exact complementary pmf of the buffer contents. In Figs. 2 and 3, we have plotted the obtained exact numerical results for $\text{Prob}[s > S]$ in terms of $S$, together with our geometric tail approximation and the approximation obtained by Sohraby in [9], for various traffic parameters. As one can observe, our approximation is really very tight to the actual tail distribution, whereas Sohraby’s results overestimate the tail distribution. Fig. 3 furthermore illustrates the importance of the precise distribution of the off-periods: a simplified model which uses a geometric distribution for the off-periods, i.e. $L = 1$, tends to underestimates the tail distribution of the buffer contents.

6.2. **On/off sources with fixed spacing**

We now consider an ATM statistical multiplexer with $N$ homogeneous fixedly spaced on/off sources. The on/off traffic source model with fixed spacing is described as follows. Each source alternates between “active” and “passive” periods. The number of cells generated in an active period and the length of a passive period are both assumed to be geometrically distributed with parameters $\alpha$ and $\beta$, respectively. Hence, the mean number of cells entering the system during an active period, the so-called “average burst length”, is given by $E x = 1 / (1 - \alpha)$. The interarrival time between two adjacent cells generated in the same active period is fixed, and equals $m$ slots. The pgf’s $A(z)$ and $B(z)$ describing the lengths of the on-periods and the off-periods, as defined in Section 2, can then be obtained as

$$A(z) = z \quad \text{and} \quad B(z) = \alpha z^{m-1} + \frac{(1-\alpha)(1-\beta)z^m}{1-\beta z}.$$ (52)
Fig. 4  Tail distribution of the buffer contents Prob\([s > S]\) versus \(S\), for fixedly spaced on/off sources, for \(N = 16\) and \(\rho = 0.4\).

Fig. 5  Tail distribution of the buffer contents Prob\([s > S]\) versus \(S\), for fixedly spaced on/off sources, for \(N = 8\), \(\rho = 0.4\) and \(m = 4\).

In Figs. 4 and 5, the geometric tail approximation of the buffer contents is compared with the exact complementary pmf, obtained by simulation, and with the approximation derived in [9], for various values of \(Ex\) and \(m\). These figures reveal that our approximation is quite close to the simulation results. It is worth noting here that deriving the simulation results is very time consuming and becomes even unpracticable for low values of Prob\([s > S]\), while calculating our approximation for the tail distribution requires very little computing time. Fig. 4 also shows that for
given values of the total load \( \rho \) and the average burst length \( E_x \), the performance improves with increasing values of the spacing interval \( m \). Fig. 5 illustrates that, for given \( \rho \) and \( m \), longer average burst lengths \( E_x \) give rise to higher buffer contents.

7. Conclusions

In this paper, we have evaluated the performance of a statistical multiplexer fed by \( N \) independent heterogeneous bursty on/off sources, under the assumption of general distributions for the lengths of both the on-periods and the off-periods. A simple approximation - of geometric form - has been obtained for the tail distribution of the buffer contents based on the use of the dominant pole of its pgf. By means of some numerical examples, the accuracy of the obtained results has been verified and the impact of the distributions of the on-periods and the off-periods on the performance has been demonstrated.

Acknowledgement

The authors wish to thank the Belgian National Fund for Scientific Research (N.F.W.O.) for support of this research. Also, they wish to thank their former colleague, Mr. Yijun Xiong, for providing the simulation results presented in Section 6.2.

References

