

Deriving the tail distribution of the buffer contents in a statistical multiplexer with general heterogeneous on/off sources

S. Wittevrongel and H. Bruneel

SMACS Research Group, University of Ghent

Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium

Phone : +32-9-264 89 01, Fax : +32-9-264 42 95

E-mail : sw@telin.rug.ac.be

Abstract

A statistical multiplexer with a finite number of independent heterogeneous 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. 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 sources on the multiplexer performance.

Keywords

ATM, multiplexer performance, general on/off sources, analytic study

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.

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. 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 exceeds 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. In addition, heavy-traffic approximations for the dominant pole of the probability generating function (pgf) of the buffer contents, which plays a major role in the geometric tail approximation, are derived.

Statistical multiplexers with on/off sources have been studied before, under various assumptions for the distributions of the on-periods and the off-periods. Both the on-periods and the off-periods of the sources are assumed to be purely geometrically distributed in (Janakiraman, Pagurek, Neilson, 1980), (Hirano, Watanabe, 1989), (Viterbi, 1986) and (Bruneel, 1988). The on-periods are geometrically distributed multiples of fixed-length intervals in (Xiong, Bruneel, 1992), are distributed according to a mixture of 2 geometric distributions in (Steyaert, Bruneel, 1995), and may have a general distribution in (Wittevrongel, Bruneel, 1995). The present paper can thus be viewed as an extension of (Wittevrongel, Bruneel, 1995), 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 (Elsayed, 1994) and (Sohraby, 1993). However, in (Elsayed, 1994), 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 (Sohraby, 1993), an analytic geometric approximation is derived for the tail distribution of the buffer contents. However, the exact numerical calculation of the geometric decay rate γ 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.

2 MODELING ASSUMPTIONS

A discrete-time queueing model for an ATM statistical multiplexer is considered. 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 and ends at slot boundaries.
- 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 $N = N_1 + \dots + N_T$.
- Each 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 , the probability mass functions (pmf's) of the on-periods and the off-periods are denoted as $a_t(n)$ and $b_t(n)$, and the corresponding pgf's as $A_t(z)$ and $B_t(z)$.

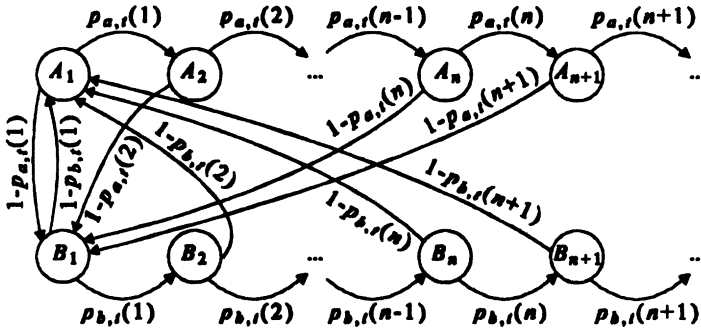


Figure 1 Markov chain model of a type t source.

- 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 ρ 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 Figure 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,

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

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 Figure 1, we then have

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

$$h_{n,k}(t) = \sum_{i=1}^{h_{n-1,k-1}(t)} d_{n-1,i}(t) , \quad n > 1 ; \quad (2)$$

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

$$h_{1,k}(t) = N_t - \sum_{n=1}^{\infty} g_{n,k}(t) - \sum_{n=2}^{\infty} h_{n,k}(t) . \quad (4)$$

The reasoning behind equation (1) 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 $c_{n-1,i}(t)$ in equation (1) 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 $p_{a,t}(n)$, as can be seen from Figure 1. A similar reasoning holds for equations (2)-(4). 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 pgf's

$$C_{n,t}(z) = 1 - p_{a,t}(n) + p_{a,t}(n) z \quad \text{and} \quad D_{n,t}(z) = 1 - p_{b,t}(n) + p_{b,t}(n) z . \quad (5)$$

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 independent, it is easily seen that the joint pgf $N(x_1, \dots, x_T, y_1, \dots, y_T)$ of the $g_n(t)$'s and the $h_n(t)$'s can be expressed as

$$\begin{aligned} N(x_1, \dots, x_T, y_1, \dots, y_T) &\triangleq E \left[\prod_{t=1}^T \left(\prod_{n=1}^{\infty} x_{n,t}^{g_n(t)} \right) \left(\prod_{n=2}^{\infty} y_{n,t}^{h_n(t)} \right) \right] \\ &= \prod_{t=1}^T \left(v_{b,t}(1) + \sum_{n=1}^{\infty} v_{a,t}(n) x_{n,t} + \sum_{n=2}^{\infty} v_{b,t}(n) y_{n,t} \right)^{N_t} , \quad (6) \end{aligned}$$

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

$$v_{a,t}(n) = v_{a,t}(1) \left(\sum_{i=n}^{\infty} a_t(i) \right), \quad v_{b,t}(n) = v_{b,t}(1) \left(\sum_{i=n}^{\infty} b_t(i) \right), \quad n \geq 1; \quad (7)$$

$$v_{a,t}(1) = v_{b,t}(1) = \frac{1}{A'_t(1) + B'_t(1)}. \quad (8)$$

From (6), it is easy to show that the total average multiplexer load ρ is given by

$$\rho = \sum_{t=1}^T N_t \sigma_t = \sum_{t=1}^T N_t \frac{A'_t(1)}{A'_t(1) + B'_t(1)}, \quad (9)$$

where σ_t is the average load contributed by a type t source and $A'_t(1)$ and $B'_t(1)$ are the mean on-periods and the mean off-periods of a source of type t .

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 :

$$s_{k+1} = (s_k - 1)^+ + e_k, \quad (10)$$

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

$$e_k = \sum_{t=1}^T \sum_{n=1}^{\infty} g_{n,k}(t). \quad (11)$$

From the above equations, it is clear that the vector $(g_{1,k-1}, \dots, g_{T,k-1}, h_{1,k-1}, \dots, h_{T,k-1}, s_k)$, where $g_{t,k-1} = (g_{1,k-1}(t), g_{2,k-1}(t), \dots)$ and $h_{t,k-1} = (h_{2,k-1}(t), h_{3,k-1}(t), \dots)$, 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 $(g_{1,k-1}, \dots, g_{T,k-1}, h_{1,k-1}, \dots, h_{T,k-1}, s_k)$:

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

With this definition and equations (10) and (11), $P_{k+1}(x_1, \dots, x_T, y_1, \dots, y_T, z)$ can be obtained as

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

Based on equations (1)-(3) and averaging over the distributions of the $c_{n,t}(t)$'s and the $d_{n,t}(t)$'s, defined in equation (5), it follows that

$$P_{k+1}(x_1, \dots, x_T, y_1, \dots, y_T, z) = E \left[\left(\prod_{t=1}^T \prod_{n=1}^{\infty} C_{n,t}(x_{n+1,t} z)^{g_{n,k-1}(t)} \right) \left(\prod_{t=1}^T \prod_{n=1}^{\infty} \left(x_{1,t} z D_{n,t} \left(\frac{y_{n+1,t}}{x_{1,t} z} \right) \right)^{h_{n,k-1}(t)} \right) z^{(s_k-1)^+} \right].$$

Upon substitution of equation (4) this can be rewritten as

$$P_{k+1}(x_1, \dots, x_T, y_1, \dots, y_T, z) = \left(\prod_{t=1}^T \left[x_{1,t} z D_{1,t} \left(\frac{y_{2,t}}{x_{1,t} z} \right) \right]^{N_t} \right) \cdot E \left[\left(\prod_{t=1}^T \prod_{n=1}^{\infty} G_{n,t}(x_t, y_t, z)^{g_{n,k-1}(t)} \right) \left(\prod_{t=1}^T \prod_{n=2}^{\infty} H_{n,t}(x_t, y_t, z)^{h_{n,k-1}(t)} \right) z^{(s_k-1)^+} \right], \quad (12)$$

where

$$G_{n,t}(x_t, y_t, z) \triangleq C_{n,t}(x_{n+1,t} z) / x_{1,t} z D_{1,t} \left(\frac{y_{2,t}}{x_{1,t} z} \right), \quad n \geq 1, \quad 1 \leq t \leq T; \quad (13)$$

$$H_{n,t}(x_t, y_t, z) \triangleq D_{n,t} \left(\frac{y_{n+1,t}}{x_{1,t} z} \right) / D_{1,t} \left(\frac{y_{2,t}}{x_{1,t} z} \right), \quad n \geq 2, \quad 1 \leq t \leq T. \quad (14)$$

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, $P_k(x_1, \dots, x_T, y_1, \dots, y_T, z)$ becomes independent of k . From equation (12), it is not so difficult to show that the steady-state pgf $P(x_1, \dots, x_T, y_1, \dots, y_T, z)$ satisfies the following functional equation :

$$z P(x_1, \dots, x_T, y_1, \dots, y_T, z) = \left(\prod_{t=1}^T \left[x_{1,t} z D_{1,t} \left(\frac{y_{2,t}}{x_{1,t} z} \right) \right]^{N_t} \right) \cdot \left\{ P(G_1(x_1, y_1, z), \dots, G_T(x_T, y_T, z), H_1(x_1, y_1, z), \dots, H_T(x_T, y_T, z), z) + (z-1) p_0 Q(H_1(x_1, y_1, z), \dots, H_T(x_T, y_T, z)) \right\} , \quad (15)$$

where

$$G_t(x_t, y_t, z) \triangleq (G_{1,t}(x_t, y_t, z), G_{2,t}(x_t, y_t, z), \dots) ; \quad (16)$$

$$H_t(x_t, y_t, z) \triangleq (H_{2,t}(x_t, y_t, z), H_{3,t}(x_t, y_t, z), \dots) , \quad (17)$$

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 (15) is defined as

$$Q(y_1, \dots, y_T) \triangleq \sum_{j_1, \dots, j_T} q(j_1, \dots, j_T) \left(\prod_{t=1}^T \prod_{n=2}^{\infty} y_{n,t}^{j_{n,t}} \right) , \quad (18)$$

where

$$q(j_1, \dots, j_T) \triangleq \lim_{k \rightarrow \infty} \text{Prob} \left[h_{n,k-1}(t) = j_{n,t} \ (n \geq 2, 1 \leq t \leq T) \mid s_k = 0 \right] . \quad (19)$$

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 $q(j_1, \dots, j_T)$ and hence, for the function $Q(y_1, \dots, y_T)$.

Since the pgf $S(z)$ of the steady-state buffer contents s equals $P(1, \dots, 1, z)$, in principle, the functional equation fully describes the buffer behavior. Unfortunately, in general, it is not an easy task to obtain from (15) 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 (15) are identical, i.e. such that

$$x_{n,t} = G_{n,t}(x_t, y_t, z) \quad \text{and} \quad y_{n,t} = H_{n,t}(x_t, y_t, z) . \quad (20)$$

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 (20), it is possible to show that $\chi_{n,t}(z)$ and $\xi_{n,t}(z)$ satisfy the following set of equations :

$$z \left(\sum_{i=n}^{\infty} a_i(i) \right) \chi_{n,t}(z) = \left(\frac{F_t(z)}{z} \right)^{n-1} A_t \left(\frac{z}{F_t(z)} \right) - \sum_{i=1}^{n-1} a_i(i) \left(\frac{F_t(z)}{z} \right)^{n-1-i}, \quad n \geq 1; \quad (21)$$

$$\left(\sum_{i=n}^{\infty} b_i(i) \right) \xi_{n,t}(z) = \chi_{1,t}(z) z \left\{ F_t(z)^{n-1} B_t \left(\frac{1}{F_t(z)} \right) - \sum_{i=1}^{n-1} b_i(i) F_t(z)^{n-1-i} \right\}, \quad n \geq 2, \quad (22)$$

where

$$F_t(z) \triangleq b_t(1) \chi_{1,t}(z) z + (1 - b_t(1)) \xi_{2,t}(z). \quad (23)$$

By choosing $x_{n,t} = \chi_{n,t}(z)$ and $y_{n,t} = \xi_{n,t}(z)$ in (20), we then get

$$P(\chi_1(z), \dots, \chi_T(z), \xi_1(z), \dots, \xi_T(z), z) = \frac{(z-1) p_0 F(z) \tilde{Q}(z)}{z - F(z)}, \quad (24)$$

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

$$F(z) \triangleq \prod_{t=1}^T \left[\chi_{1,t}(z) z D_{1,t} \left(\frac{\xi_{2,t}(z)}{\chi_{1,t}(z) z} \right) \right]^{N_t} = \prod_{t=1}^T F_t(z)^{N_t}; \quad (25)$$

$$\tilde{Q}(z) \triangleq Q(\xi_1(z), \dots, \xi_T(z)). \quad (26)$$

and the unknown constant p_0 can be calculated from the normalisation condition $P(1, \dots, 1) = 1$ as $p_0 = 1 - \rho$. Comparing the results obtained here with those obtained in (Wittevrongel, Bruneel, 1995), 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

$$\text{Prob}[s > S] \cong -\frac{\theta}{z_0 - 1} z_0^{-s-1} \quad (27)$$

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 θ 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 θ .

5.1 The dominant pole z_0

As in (Xiong, Bruneel, 1992), (Steyaert, Bruneel, 1995) and (Wittevrongel, Bruneel, 1995), it can be argued that the dominant pole z_0 of $S(z)$ is also the pole with the smallest modulus of $P(\chi_1(z), \dots, \chi_T(z), \xi_1(z), \dots, \xi_T(z), z)$. Hence, z_0 is determined by the equation $z - F(z) = 0$, or

$$z - \prod_{t=1}^T F_t(z)^{N_t} = 0 \quad (28)$$

From equations (21)-(23), it furthermore follows that

$$A_t \left(\frac{z}{F_t(z)} \right) B_t \left(\frac{1}{F_t(z)} \right) = 1 \quad , \quad 1 \leq t \leq T \quad (29)$$

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 (28)-(29), we also note that the unknown boundary probabilities $q(j_1, \dots, j_T)$ have no influence on z_0 , which means that the geometric decay rate $\gamma = 1/z_0$ can be calculated exactly.

Next, we describe a method to obtain approximations for z_0 in the case of heavy traffic, i.e. when the total load ρ is high. It is expected that as ρ approaches to one, z_0 will also be close to one. If we now consider z_0 as being a function of ρ , we can expand the function $z_0(\rho)$ around $\rho = 1$. By keeping terms up to $(\rho-1)^n$ in this expansion and neglecting higher-order terms, we then get :

$$z_0 \equiv z_1(n) = 1 + \sum_{k=1}^n \frac{(\rho-1)^k}{k!} f_k \quad . \quad (30)$$

$$\text{where } f_n \triangleq \left. \frac{d^n z_0(\rho)}{d\rho^n} \right|_{\rho=1} .$$

Furthermore, $z_0(\rho)$ behaves as $1/\rho$ as ρ approaches to 1. By considering the Taylor expansion of $z_0(\rho)$ around $\rho = 1$, we obtain a second approximation for z_0 :

$$z_0 \equiv z_2(n) = 1 + \sum_{k=1}^{n-1} \frac{(\rho-1)^k}{k!} f_k + \frac{(\rho-1)^n}{n!} \frac{f_n}{\rho} \quad . \quad (31)$$

From (28) and (29), it follows that $z_0(\rho)$ satisfies the following set of equations :

$$z(\rho) - R_1(\rho)^{N_1(\rho)} \left(\prod_{i=2}^T R_i(\rho)^{N_i} \right) = 0 ; \quad A_t \left(\frac{z(\rho)}{R_t(\rho)} \right) B_t \left(\frac{1}{R_t(\rho)} \right) = 1 , \quad 1 \leq t \leq T ,$$

$$\text{where } N_1(\rho) = \frac{1}{\sigma_1} \left\{ \rho - \sum_{i=2}^T N_i \sigma_i \right\} \quad \text{and} \quad R_t(\rho) = F_t(z(\rho)) . \quad \text{We then obtain}$$

expressions for f_n by evaluating the consecutive derivatives of the above equations with respect to ρ for $\rho = 1$, taking into account that $z_0(1) = 1$. In Figure 2, we have plotted the exact value of z_0 in terms of the total load ρ , together with a number of heavy-traffic approximations, for $N = 8$ homogeneous on/off sources, with negative binomial distributions for both the on-periods and the off-periods, i.e. $A(z) = (1-\alpha)^2 z / (1-\alpha z)^2$ and $B(z) = (1-\beta)^2 z / (1-\beta z)^2$. The parameter K equals $K = A'(1) (1-\rho/N) = B'(1) \rho/N$. The figure shows that the obtained approximations are quite close to the exact results, especially for high values of ρ , but even for lower loads, accurate approximations can be obtained by keeping a sufficient number of terms in (30) and (31). These approximations are excellent starting values when calculating the exact value of z_0 .

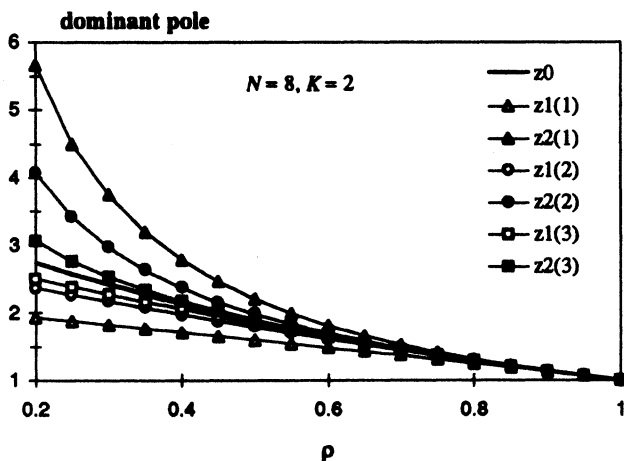


Figure 2 Dominant pole versus the load ρ : exact results and heavy-traffic approximations, for negative binomial on-periods and off-periods.

5.2 The residue θ

Let us consider the situation where the buffer contents just after a given slot is extremely large ($\gg N$). 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 buffer contents. Hence, if j is sufficiently large ($j > J$), we may assume that the 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 $j \rightarrow \infty$, denoted by $\omega(i_1, \dots, i_T, k_1, \dots, k_T)$, with corresponding pgf $\Omega(x_1, \dots, x_T, y_1, \dots, y_T)$.

The residue θ can then be derived, in a similar way as explained in (Xiong, Bruneel, 1992), (Steyaert, Bruneel, 1995) and (Wittevrongel, Bruneel, 1995), as

$$\theta = \frac{z_0 (z_0 - 1) (1 - \rho) \tilde{Q}(z_0)}{[1 - F'(z_0)] \Omega(\chi_1(z_0), \dots, \chi_T(z_0), \xi_1(z_0), \dots, \xi_T(z_0))} \quad (32)$$

Furthermore, the following equation can be established for the pgf $\Omega(x_1, \dots, x_T, y_1, \dots, y_T)$ of the limiting cell arrival process :

$$z_0 \Omega(x_1, \dots, x_T, y_1, \dots, y_T) = \left(\prod_{t=1}^T \left[x_{1,t} z_0 D_{1,t} \left(\frac{y_{2,t}}{x_{1,t} z_0} \right) \right]^{N_t} \right) \cdot \Omega(G_1(x_1, y_1, z_0), \dots, G_T(x_T, y_T, z_0), H_1(x_1, y_1, z_0), \dots, H_T(x_T, y_T, z_0)) \quad (33)$$

As expected intuitively, it is possible to show that the solution $\Omega(x_1, \dots, x_T, y_1, \dots, y_T)$ of (33) has the same form of expression as the pgf $N(x_1, \dots, x_T, y_1, \dots, y_T)$ of the unconditional cell arrival process. Specifically,

$$\Omega(x_1, \dots, x_T, y_1, \dots, y_T) = \prod_{t=1}^T \left(\mu_{b,t}(1) + \sum_{n=1}^{\infty} \mu_{a,t}(n) x_{n,t} + \sum_{n=2}^{\infty} \mu_{b,t}(n) y_{n,t} \right)^{N_t}, \quad (34)$$

where $\mu_{a,t}(n)$ and $\mu_{b,t}(n)$ are the (conditional) probabilities of finding a source of type t in state A_n or B_n , respectively, when the number of cells in the multiplexer buffer is extremely large. Substitution of (34) into (33) 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

$$\mu_{a,t}(1) = \frac{z_0 (z_0 - F_t(z_0)) B_t(1/F_t(z_0)) (F_t(z_0) - 1)}{F_t(z_0)^2 (z_0 - 1) [1 - B_t(1/F_t(z_0))]}; \quad (35)$$

$$\mu_{a,t}(n) = \mu_{a,t}(1) \left(\frac{z_0}{F_t(z_0)} \right)^{n-1} \left(\sum_{i=n}^{\infty} a_t(i) \right), \quad n \geq 2; \quad (36)$$

$$\mu_{b,t}(2) = \frac{(z_0 - F_t(z_0)) [1 - b_t(1)] (F_t(z_0) - 1)}{F_t(z_0)^3 (z_0 - 1) [1 - B_t(1/F_t(z_0))]}; \quad (37)$$

$$\mu_{b,t}(n) = \frac{\mu_{b,t}(2)}{[1 - b_t(1)]} \left(\frac{1}{F_t(z_0)} \right)^{n-2} \left(\sum_{i=n}^{\infty} b_t(i) \right), \quad n \geq 3. \quad (38)$$

Furthermore, by using (21)-(23), (28) and the above expressions, we obtain

$$\Omega(\chi_1(z_0), \dots, \chi_T(z_0), \xi_1(z_0), \dots, \xi_T(z_0)) = \prod_{t=1}^T \left(\frac{[z_0 - F_t(z_0)] [F_t(z_0) - 1]}{(z_0 - 1) F_t(z_0)^3 \left[1 - B_t \left(\frac{1}{F_t(z_0)} \right) \right]} \right)^{N_t} \cdot \left[z_0 A_t' \left(\frac{z_0}{F_t(z_0)} \right) B_t \left(\frac{1}{F_t(z_0)} \right) + A_t \left(\frac{z_0}{F_t(z_0)} \right) B_t' \left(\frac{1}{F_t(z_0)} \right) \right]^{N_t}. \quad (39)$$

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

$$F'(z_0) = \sum_{t=1}^T \frac{N_t z_0 A_t' \left(\frac{z_0}{F_t(z_0)} \right) B_t \left(\frac{1}{F_t(z_0)} \right)}{\left[z_0 A_t' \left(\frac{z_0}{F_t(z_0)} \right) B_t \left(\frac{1}{F_t(z_0)} \right) + A_t \left(\frac{z_0}{F_t(z_0)} \right) B_t' \left(\frac{1}{F_t(z_0)} \right) \right]} . \quad (40)$$

The only unknown quantity in (32) is $\tilde{Q}(z_0)$, which contains a very large number of unknown boundary probabilities $q(j_1, \dots, j_T)$, defined in (19). 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 $g_{n,k-1}(t) = 0$ ($n \geq 1, 1 \leq t \leq T$), 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

$$q(j_1, \dots, j_T) \quad (41)$$

$$\cong \lim_{k \rightarrow \infty} \text{Prob} \left[h_{n,k-1}(t) = j_{n,t} \ (n \geq 2, 1 \leq t \leq T) \mid g_{n,k-1}(t) = 0 \ (n \geq 1, 1 \leq t \leq T) \right] .$$

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 (18), the function $\tilde{Q}(z)$ can be expressed in terms of the pgf $N(x_1, \dots, x_T, y_1, \dots, y_T)$. As a result, the following approximation is obtained for $\tilde{Q}(z)$:

$$\tilde{Q}(z) \cong \frac{N(0, \dots, 0, \xi_1(z), \dots, \xi_1(z))}{N(0, \dots, 0, 1, \dots, 1)} . \quad (42)$$

A little more algebra finally yields

$$\tilde{Q}(z_0) \cong \prod_{t=1}^T \left\{ \frac{1}{B_t'(1)} \left[1 + A_t \left(\frac{z_0}{F_t(z_0)} \right) \left(F_t(z_0) B_t \left(\frac{1}{F_t(z_0)} \right) - 1 \right) / \left[1 - F_t(z_0) \right] \right] \right\}^{N_t} . \quad (43)$$

Combination of equations (32), (39), (40) and (43) then gives an explicit approximation for the parameter θ in terms of z_0 .

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 in (Sohraby, 1993). The most important differences between our study and Sohraby's are :

- The exact numerical calculation of the geometric decay rate $\gamma = 1/z_0$ by means of the method presented in (Sohraby, 1993) requires knowledge of the algebraic expression of the Perron-Frobenius (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 z_0 can easily be calculated from (28) and (29) by means of a numerical algorithm. In addition, we have presented a method to derive accurate approximations for the dominant pole z_0 , which could be used as starting values, when calculating the exact value of z_0 .
- Our study not only yields the geometric decay rate; it also leads to an accurate closed-form approximation for the coefficient of the geometric form, whereas in (Sohraby, 1993) this coefficient is simply approximated by the multiplexer load, i.e., the tail distribution of the buffer contents is approximated as

$$\text{Prob}[s > S] \cong \rho z_0^{-S} . \quad (44)$$

The approximation (44) 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 on-periods are geometrically distributed, whereas the off-periods are distributed according to a mixture of 2 geometric distributions, i.e.

$$A(z) = \frac{(1-\alpha)z}{1-\alpha z} \quad \text{and} \quad B(z) = \frac{q(1-\beta_1)z}{1-\beta_1 z} + \frac{(1-q)(1-\beta_2)z}{1-\beta_2 z} , \quad \beta_1 > \beta_2 . \quad (45)$$

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

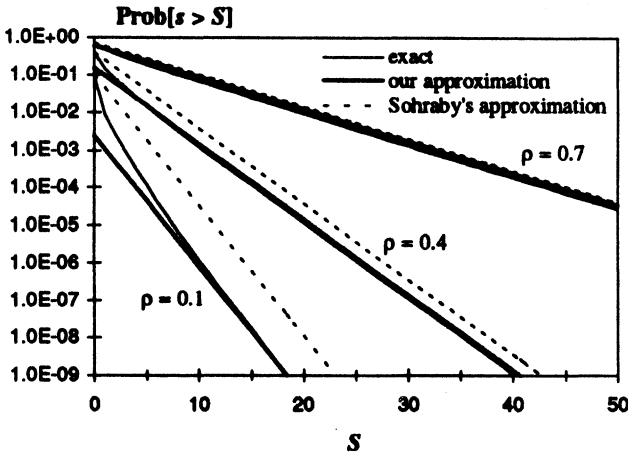


Figure 3 Prob[$s > S$] versus S , for geometric on-periods and mixed geometric off-periods, for $N = 4, K = 2, L = 2$ and $q = 0.3$.

parameters. First of all, the total multiplexer load ρ 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 σ 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 β_1 and β_2 can be derived from

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

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

Hence, the set (N, ρ, K, L, q) also fully describes the cell arrival process.

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

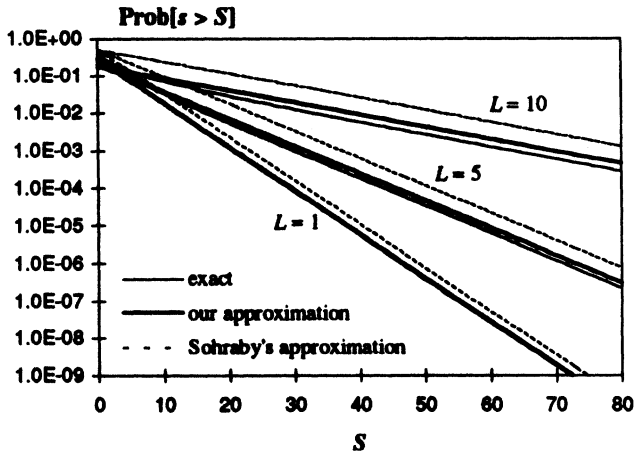


Figure 4 $\text{Prob}[s > S]$ versus S , for geometric on-periods and mixed geometric off-periods, for $N = 4$, $\rho = 0.5$, $K = 3$ and $q = 0.1$.

of balance equations was feasible, leading to the exact complementary pmf of the buffer contents. In Figures 3 and 4, 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 in (Sohrab, 1993) 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. Figure 4 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 underestimate 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 be geometrically distributed with parameters α and β , 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

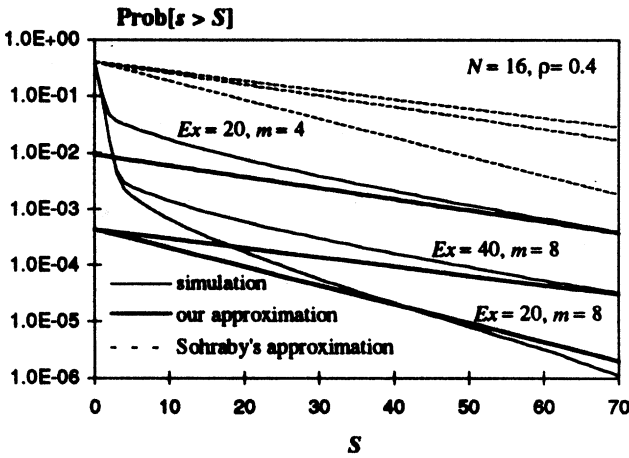


Figure 5 $\text{Prob}[s > S]$ versus S , for fixedly spaced on/off sources.

$$A(z) = z \quad \text{and} \quad B(z) = \alpha z^{m-1} + \frac{(1-\alpha)(1-\beta)z^m}{1-\beta z} \quad (46)$$

In Figures 5 and 6, the geometric tail approximation of the buffer contents is compared with the exact complementary pmf, obtained by simulation, and with the approximation derived in (Sohraby, 1993), 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 $\text{Prob}[s > S]$, while calculating our approximation for the tail distribution requires very little computing time. Figure 5 also shows that for given values of the total load ρ and the average burst length Ex , the performance improves with increasing values of the spacing interval m . Figure 6 illustrates that, for given ρ and m , longer average burst lengths Ex give rise to higher buffer contents.

7 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.

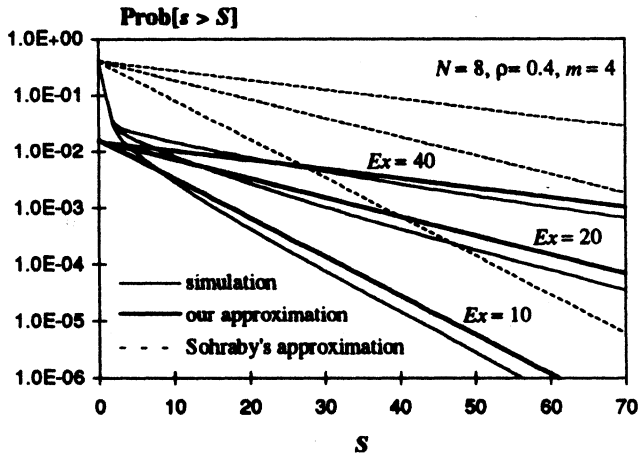


Figure 6 $\text{Prob}[s > S]$ versus S , for fixedly spaced on/off sources.

8 REFERENCES

- Bruneel, H. (1988) Queuing behavior of statistical multiplexers with correlated inputs. *IEEE Transactions on Communications*, **36**, 1339-1341.
- Elsayed, K. (1994) On the superposition of discrete-time Markov renewal processes and application to statistical multiplexing of bursty traffic sources, in *Proceedings of IEEE GLOBECOM '94*, San Francisco, 1113-1117.
- Hirano, M. and Watanabe, N. (1989) Characteristics of a cell multiplexer for bursty ATM traffic, in *Proceedings of IEEE ICC '89*, Boston, 399-403.
- Janakiraman, N.; Pagurek, B. and Neilson, J. E. (1980) Multiplexing low-speed buffered data terminals. *IEEE Transactions on Communications*, **28**, 1838-1843.
- Sohrab, K. (1993) On the theory of general ON-OFF sources with applications in high-speed networks, in *Proceedings of IEEE INFOCOM '93*, San Francisco, 401-410.
- Steyaert, B. and Bruneel, H. (1995) On the performance of multiplexers with three-state bursty sources: analytical results. *IEEE Transactions on Communications*, **43**, 1299-1303.
- Viterbi, A. M. (1986) Approximate analysis of time-synchronous packet networks. *IEEE Journal on Selected Areas in Communications*, **4**, 879-890.
- Wittevrongel, S. and Bruneel, H. (1995) Effect of the on-period distribution on the performance of an ATM multiplexer fed by on/off sources: an analytical study, in *Proceedings of PCN '95*, Istanbul, 33-47.

Xiong, Y. and Bruneel, H. (1992) Performance of statistical multiplexers with finite number of inputs and train arrivals, in *Proceedings of IEEE INFOCOM '92*, Firenze, 2036-2044.

9 BIOGRAPHIES

Sabine WITTEVRONGEL was born in Gent, Belgium, in 1969. She received the M.S. degree in Electrical Engineering from the University of Ghent, Belgium, in 1992. Since September 1992, she has been with the SMACS Research Group, Laboratory for Communications Engineering, University of Ghent, first in the framework of various projects, and since October 1994, as a researcher for the Belgian National Fund for Scientific Research (N.F.W.O.). Her main research interests include discrete-time queueing theory, performance evaluation of ATM networks and the study of traffic control mechanisms.

Herwig BRUNEEL was born in Zottegem, Belgium, in 1954. He received the M.S. degree in Electrical Engineering, the degree of Licentiate in Computer Science, and the Ph.D. degree in Computer Science in 1978, 1979 and 1984 respectively, all from the University of Ghent, Belgium. Since 1979, he has been working as a researcher for the Belgian National Fund for Scientific Research (N.F.W.O.) at the University of Ghent, where he currently leads the SMACS Research Group within the Laboratory for Communications Engineering. He is also a Professor in the Faculty of Applied Sciences at the same university. His main research interests include stochastic modeling of digital communication systems, discrete-time queueing theory, and the study of ARQ protocols. He has published more than 120 papers on these subjects and is coauthor of the book *H. Bruneel and B. G. Kim, "Discrete-Time Models for Communication Systems Including ATM"* (Kluwer Academic Publishers, Boston, 1993).