Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2012, Article ID 326830, 17 pages
doi:10.1155/2012/326830

Research Article

Transient and Stationary Losses in a Finite-Buffer Queue with Batch Arrivals

Andrzej Chydzinski and Blazej Adamczyk

Institute of Informatics, Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland

Correspondence should be addressed to Andrzej Chydzinski, andrzej.chydzinski@polsl.pl

Received 10 July 2012; Accepted 12 November 2012

Academic Editor: Joao B. R. Do Val

We present an analysis of the number of losses, caused by the buffer overflows, in a finite-buffer queue with batch arrivals and autocorrelated interarrival times. Using the batch Markovian arrival process, the formulas for the average number of losses in a finite time interval and the stationary loss ratio are shown. In addition, several numerical examples are presented, including illustrations of the dependence of the number of losses on the average batch size, buffer size, system load, autocorrelation structure, and time.

1. Introduction

In a finite-buffer queueing system (i.e., a system with the finite waiting room), we should expect losses. Namely, jobs (customers) that arrive at the system when the buffer is full are rejected and lost. Naturally, in most applications of queueing systems, the losses are unwanted. This is especially true in telecommunications. A great part of today's telecommunication systems are based on packet-switched networks, where packets losses occur at buffers in network nodes. For instance, as many as 17% of packets are lost globally due to the buffer overflows in the Internet (observed on November 15th, 2011, see [1]). Therefore, an enormous amount of data has to be retransmitted.

As we know, several characteristics of a finite-buffer queueing system can influence the number of losses. These are, for instance,

(1) the load of the system (traffic intensity),

(2) the buffer size,

(3) the variance of the service times,

(4) the variance of the interarrival times,

(5) the batch arrivals,

(6) the correlation of the interarrival times.

The dependence of the number of losses on (1) and (2) is quite obvious. The dependence on (3)–(6) follows, for instance, from the results presented in [2–5], respectively. However, the queueing models considered in these papers do not take into account all of the aforementioned factors at the same time.

The purpose of this paper is to find formulas describing the loss process in a finite-buffer queueing model that enables fitting all of characteristics (1)–(6). What is more, we want to describe the loss process both in the transient and stationary case and provide closed, easy to use, formulas for the loss process characteristics.

To the best of the authors' knowledge, there are no previously published papers that fulfill these requirements. In particular, the influence of the system load, buffer size, and service time variance on the number of losses is studied in several classic queueing theory textbooks. However, the classic models (like M/G/1/N or G/M/1/N) and the classic methodology do not take into account the autocorrelation in the arrival process and the batch arrivals. Other studies [2, 6–8] do take into account the batch arrivals, but without the autocorrelation structure. Finally, some recent papers that incorporate the autocorrelation structure either do not consider the batch arrivals (like [5, 9–11]) or do not deal with the transient case (like [12]).

As for the arrival process model, we have chosen the batch Markovian arrival process (BMAP) [13]. This is due to the following reasons. Firstly, the BMAP process allows us to model not only the batch arrivals, the variance of interarrival times, and the correlation of interarrival times but also many other subtle characteristics of the arrival process (e.g., the correlation between the local intensity of arrivals of batches with the size of an arriving batch). Secondly, these modeling capabilities can be used in practice, due to the availability of a number of parameter fitting procedures for BMAPs [14–16]. Finally, the BMAP has a unique advantage of combining great complexity and modeling capabilities with the analytical tractability (for more information on BMAPs see [17] and the references given there).

As for the methodology, we will exploit the framework from [5], which has been previously used for simpler Markovian processes (e.g., [18]). The main difference herein is batch structure of arrival process (not present in [5]) which causes some important complications of the method and the results.

The methodology of [5] is used herein due to its ability to solve the transient case (in addition to the stationary one). Other known methods for finding loss characteristics in BMAP queues (e.g., [9, 12]) are devoted to the stationary case only. It is an open question whether they can be extended to cover the transient case as well. Certainly, such extensions would not be trivial.

The remaining part of the paper is structured in the following way. In Section 2, we first give the definition of the arrival process, as well as a few useful formulas for its basic characteristics. Secondly, we present a formal description of the queueing model and the nomenclature used in the paper. The main part of the paper, Section 3, then follows. Namely, it starts with the definition of the main characteristic of interest, which is the average number of losses in interval $(0, t]$. Then, this characteristic is derived by using the Laplace transform technique. Next, the transient intensity of the loss process is derived and some comments on how to use the obtained results in practice are presented. Finally, using the previous results, the stationary loss ratio is computed. In Section 4, a set of numerical results based on four different BMAPs are presented. In particular, the dependence of the loss ratio on the autocorrelation structure, on the batch size distribution, and on the buffer size in the steady-state, as well as the transient intensity of the loss process, are investigated. In Section 5, the remarks concluding the paper are gathered.

2. The Arrival Process and the Queueing Model

Let I denote the identity matrix; let $\mathbf{0}$ be a square matrix of zeroes and $\mathbf{1}$ the column vector of 1's.

The batch Markovian arrival process (BMAP) is defined as a 2-dimensional Markov process $(N(t), J(t))$ on the state space $\{(i, j) : i \ge 0,\ 1 \le j \le m\}$ with an infinitesimal generator Q in the form

$$Q = \begin{bmatrix} D_0 & D_1 & D_2 & D_3 & \cdot & \cdot \\ & D_0 & D_1 & D_2 & \cdot & \cdot \\ & & D_0 & D_1 & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \end{bmatrix}, \tag{2.1}$$

where D_k, $k \ge 0$ are $m \times m$ matrices. D_k, $k \ge 1$ are nonnegative, D_0 has nonnegative off-diagonal elements and negative diagonal elements and $D = \sum_{k=0}^{\infty} D_k$ is an irreducible infinitesimal generator. It is assumed also that $D \ne D_0$.

In this two-dimensional process, $N(t)$ denotes the number of arrivals in $(0, t]$, while $J(t)$ denotes the state of the one-dimensional modulating Markov process at time t. The intensity matrix for the modulating process is equal to D. Its stationary distribution will be denoted by π, where $\pi D = (0, \ldots, 0)$, $\pi\mathbf{1} = 1$.

The evolution of the BMAP process can be also described in the following manner. Given the modulating process J is in some phase i; the sojourn time in that phase is exponentially distributed with parameter λ_i, where

$$\lambda_i = -(D_0)_{ii}. \tag{2.2}$$

At the end of that sojourn time there occurs a transition to another phase and (or) an arrival of a batch. In particular, with probability $p_i(j, k)$ there will be a transition to phase k with a batch arrival of size j, where

$$\begin{gathered} p_i(0, i) = 0, \quad 1 \le i \le m, \\ p_i(0, k) = \frac{1}{\lambda_i}(D_0)_{ik}, \quad 1 \le i, k \le m,\ k \ne i, \\ p_i(j, k) = \frac{1}{\lambda_i}(D_j)_{ik}, \quad 1 \le i, k \le m,\ j \ge 1. \end{gathered} \tag{2.3}$$

Now we will give a few useful characteristics, the BMAP (see [12, 13]). First, the total arrival rate (including batch sizes) can be calculated as

$$\Lambda = \pi \sum_{k=1}^{\infty} k D_k \mathbf{1}. \tag{2.4}$$

The arrival rate of batches (i.e., excluding batch sizes) can be computed as

$$\Lambda_g = \pi(-D_0)\mathbf{1}. \tag{2.5}$$

The variance of the interarrival times is equal to

$$\mathrm{Var} = -\frac{2}{\Lambda_g}\pi D_0^{-1}\mathbf{1} - \frac{1}{\Lambda_g^2}. \tag{2.6}$$

The autocorrelation at lag k of the sequence of interarrival times is

$$\mathrm{Corr}(k) = pD_0^{-1}C\left(C^{k-1} - \mathbf{1}p\right)D_0^{-1}C\frac{\mathbf{1}}{\mathrm{Var}}, \tag{2.7}$$

where

$$C = -D_0^{-1}(D - D_0), \tag{2.8}$$

and p is the stationary vector for $C - I$, namely, $p(C - I) = (0, \ldots, 0)$, $p\mathbf{1} = 1$.

Finally, the counting function for the BMAP, which is defined as

$$P_{i,j}(n,t) = \mathbb{P}\left(N(t) = n,\ J(t) = j \mid N(0) = 0,\ J(0) = i\right), \tag{2.9}$$

has the following generating function:

$$P^*(z,t) = \sum_{n=0}^{\infty} P(n,t)z^n = e^{D(z)t}, \quad D(z) = \sum_{k=0}^{\infty} z^k D_k, \quad |z| \le 1. \tag{2.10}$$

This finishes the description of the arrival process.

As for the queueing model, we deal herein with the simple single-server queueing system of finite capacity. Namely, the arrival process is the batch Markovian arrival process described above, the service time distribution is given by a distribution function $F(t)$ (which may assume any form), and the service discipline is FIFO (FCFS). What is important is that the system capacity is finite and equal to N. This means that the total number of jobs in the system must not exceed N, including the service position. Jobs arriving when the system is full are lost and never return, as, usually, we assume that the service times are mutually independent and that they do not depend on the arrival process. Finally, we assume that $t = 0$ corresponds to a departure epoch.

3. Transient and Stationary Losses

In the sequel, $X(t)$ denotes the queue size at time t (including service position, if occupied), $L(t)$ denotes the number of jobs lost in time interval $(0,t]$, and $\Delta_{n,i}(t)$ denotes its average value assuming $X(0)=n$ and $J(0)=i$, that is,

$$\Delta_{n,i}(t)=\mathbb{E}(L(t)\mid X(0)=n,\ J(0)=i). \tag{3.1}$$

First of all, we want to find a formula for the Laplace transform of $\Delta_{n,i}(t)$:

$$\delta_{n,i}(s)=\int_0^{\infty}e^{-st}\Delta_{n,i}(t)dt. \tag{3.2}$$

For that purpose, we will use two systems of integral equation for $\Delta_{n,i}(t)$.

Namely, assuming that the queue is not empty at $t=0$ and using the law of total probability with respect to the first service completion moment we obtain the following set of integral equations:

$$\begin{aligned}\Delta_{n,i}(t)={}&\sum_{j=1}^{m}\sum_{k=0}^{N-n-1}\int_0^t\Delta_{n+k-1,j}(t-u)P_{i,j}(k,u)dF(u)\\&+\sum_{j=1}^{m}\sum_{k=N-n}^{\infty}\int_0^t\left(k-N+n+\Delta_{N-1,j}(t-u)\right)P_{i,j}(k,u)dF(u)\\&+(1-F(t))\sum_{j=1}^{m}\sum_{k=N-n}^{\infty}(k-N+n)P_{i,j}(k,t),\quad n=1,\ldots,N;\ i=1,\ldots,m.\end{aligned} \tag{3.3}$$

System (3.3) can be explained by naming all the mutually exclusive events used in (3.3). In particular, the first summand after the equality sign corresponds to the event where the first service completion time, u, occurs before t, and there are no losses by the time u. The second summand after the equality sign corresponds to the event where the first service completion time, u, occurs before t, and there are some losses by the time u. The third summand corresponds to the event where the first service completion time is after t.

Assuming that the queue is empty at $t=0$, we can obtain another system of integral equations:

$$\begin{aligned}\Delta_{0,i}(t)={}&\sum_{j=1}^{m}\sum_{k=0}^{N}\int_0^t\Delta_{k,j}(t-u)p_i(k,j)\lambda_i e^{-\lambda_i u}du\\&+\sum_{j=1}^{m}\sum_{k=N+1}^{\infty}\int_0^t\left(k-N+\Delta_{N,j}(t-u)\right)p_i(k,j)\lambda_i e^{-\lambda_i u}du,\quad i=1,\ldots,m.\end{aligned} \tag{3.4}$$

Now, the first summand after the equality sign in (3.4) corresponds to the event where the arrival of the first batch to an empty queue occurs in time u, $u<t$, and the size of this arriving batch does not exceed N. Therefore, there are no losses connected with the arrival

of the first batch. The second summand corresponds to the event where the arrival of the first batch to an empty queue occurs in time u, $u < t$; the size of the arriving batch exceeds the capacity of the system and causes a loss of $k - N$ jobs. Finally, note that the absent third summand, corresponding to the event where the first arrival of a batch occurs after time t, is not necessary—there are no losses in $(0, t]$ in such a case. After simple algebraic manipulations from (3.4) we get

$$\begin{aligned}\Delta_{0,i}(t) &= \sum_{j=1}^{m}\sum_{k=0}^{N}\int_0^t \Delta_{k,j}(t-u)p_i(k,j)\lambda_i e^{-\lambda_i u}du \\ &+ \sum_{j=1}^{m}\sum_{k=N+1}^{\infty}\int_0^t \Delta_{N,j}(t-u)p_i(k,j)\lambda_i e^{-\lambda_i u}du \\ &+ \left(1-e^{-\lambda_i t}\right)\sum_{j=1}^{m}\sum_{k=1}^{\infty}kp_i(N+k,j), \quad i = 1,\ldots,m.\end{aligned} \tag{3.5}$$

Applying the Laplace transform to (3.3) and (3.5) and employing matrix notation we obtain

$$\begin{aligned}\delta_n(s) &= \sum_{k=0}^{N-n-1}A_k(s)\delta_{n+k-1}(s) + \sum_{k=N-n}^{\infty}A_k(s)\delta_{N-1}(s) + c_n(s), \qquad n = 1,\ldots,N, \\ \delta_0(s) &= \sum_{k=0}^{N}Y_k(s)\delta_k(s) + \sum_{k=N+1}^{\infty}Y_k(s)\delta_N(s) + \sum_{k=1}^{\infty}k\frac{Y_{N+k}(s)}{s}\cdot\mathbf{1},\end{aligned} \tag{3.6}$$

where $\delta_n(s)$ and $c_k(s)$ are the following column vectors:

$$\begin{aligned}\delta_n(s) &= \left(\delta_{n,1}(s),\ldots,\delta_{n,m}(s)\right)^T, \\ c_k(s) &= \frac{1}{s}\sum_{i=N-k}^{\infty}(i-N+k)A_i(s)\cdot\mathbf{1} + \sum_{i=N-k}^{\infty}(i-N+k)E_i(s)\cdot\mathbf{1},\end{aligned} \tag{3.7}$$

while $Y_k(s)$, $A_k(s)$, and $E_k(s)$ are the following $m \times m$ matrices:

$$Y_k(s) = \left[\frac{\lambda_i p_i(k,j)}{s+\lambda_i}\right]_{i,j}, \tag{3.8}$$

$$A_k(s) = \left[\int_0^{\infty}e^{-st}P_{i,j}(k,t)dF(t)\right]_{i,j}, \tag{3.9}$$

$$E_k(s) = \left[\int_0^{\infty}e^{-st}P_{i,j}(k,t)(1-F(t))dt\right]_{i,j}. \tag{3.10}$$

Now we will solve the system (3.6). Firstly, by changing the indices numeration into

$$\widetilde{\delta}_n(s) = \delta_{N-n}(s), \tag{3.11}$$

we get

$$\sum_{k=-1}^{n} A_{k+1}(s)\widetilde{\delta}_{n-k}(s) - \widetilde{\delta}_n(s) = \psi_n(s), \quad n = 0,\dots,N-1, \tag{3.12}$$

$$\widetilde{\delta}_N(s) = \sum_{k=0}^{N} Y_{N-k}(s)\widetilde{\delta}_k(s) + \sum_{k=N+1}^{\infty} Y_k(s)\widetilde{\delta}_0(s) + x(s), \tag{3.13}$$

with

$$\begin{gathered}\psi_n(s) = A_{n+1}(s)\widetilde{\delta}_0(s) - \sum_{k=n+1}^{\infty} A_k(s)\widetilde{\delta}_1(s) - c_{N-n}(s),\\ x(s) = \sum_{k=1}^{\infty} k \frac{Y_{N+k}(s)}{s} \cdot \mathbf{1}.\end{gathered} \tag{3.14}$$

Thanks to Lemma 3.2.1 of [19], we know that the general solution of system (3.12) has the form

$$\widetilde{\delta}_n(s) = R_{n+1}(s)C(s) + \sum_{k=0}^{n} R_{n-k}(s)\psi_k(s), \tag{3.15}$$

where

$$\begin{gathered}R_0(s) = \mathbf{0}, \qquad R_1(s) = A_0^{-1}(s),\\ R_{k+1}(s) = A_0^{-1}(s)\left(R_k(s) - \sum_{i=0}^{k} A_{i+1}(s)R_{k-i}(s)\right), \quad k = 1,2\dots,\end{gathered} \tag{3.16}$$

and $C(s)$ is a column vector that does not depend on n.

Formula (3.12) for $n = 0$ gives

$$\sum_{k=0}^{\infty} A_k(s)\widetilde{\delta}_1(s) - \widetilde{\delta}_0(s) = -c_N(s), \tag{3.17}$$

and, as a consequence,

$$\begin{aligned}\psi_n(s) &= A_{n+1}(s)\widetilde{\delta}_0(s) - \sum_{k=n+1}^{\infty} A_k(s)\overline{A_0}^{-1}(s)\left(\widetilde{\delta}_0(s) - c_N(s)\right) - c_{N-n}(s)\\ &= B_n(s)\widetilde{\delta}_0(s) + \overline{A}_{n+1}(s)\left(\overline{A}_0\right)^{-1}(s)c_N(s) - c_{N-n}(s),\end{aligned} \tag{3.18}$$

with

$$\overline{A}_n(s) = \sum_{k=n}^{\infty} A_k(s), \qquad B_n(s) = A_{n+1}(s) - \overline{A}_{n+1}(s)\left(\overline{A}_0(s)\right)^{-1}. \tag{3.19}$$

On the other hand, formula (3.15) for $n = 0$ gives

$$C(s) = A_0(s)\widetilde{\delta}_0(s). \tag{3.20}$$

Now, putting (3.15), (3.18), and (3.20) into (3.13) we get the following equation for $\widetilde{\delta}_0(s)$:

$$\begin{aligned}
&R_{N+1}(s)A_0(s)\widetilde{\delta}_0(s) + \sum_{k=0}^{N} R_{N-k}(s)\left(B_k(s)\widetilde{\delta}_0(s) + \overline{A}_{k+1}(s)\left(\overline{A}_0(s)\right)^{-1} c_N(s) - c_{N-k}(s)\right) \\
&\quad = \sum_{k=0}^{N} Y_{N-k}(s)\Bigg[R_{k+1}(s)A_0(s)\widetilde{\delta}_0(s) \\
&\qquad\qquad + \sum_{l=0}^{k} R_{k-l}(s)\left(B_l(s)\widetilde{\delta}_0(s) + \overline{A}_{l+1}(s)\left(\overline{A}_0(s)\right)^{-1} c_N(s) - c_{N-l}(s)\right)\Bigg] \\
&\qquad + \sum_{k=N+1}^{\infty} Y_k(s)\widetilde{\delta}_0(s) + x(s).
\end{aligned} \tag{3.21}$$

Solving this equation with respect to $\widetilde{\delta}_0(s)$ we obtain

$$\widetilde{\delta}_0(s) = Q_N^{-1}(s)q_N(s), \tag{3.22}$$

where

$$\begin{aligned}
Q_N(s) &= R_{N+1}(s)A_0(s) + \sum_{k=0}^{N} R_{N-k}(s)B_k(s) - \sum_{k=0}^{N} Y_{N-k}(s)R_{k+1}(s)A_0(s) \\
&\quad - \sum_{k=0}^{N}\sum_{l=0}^{k} Y_{N-k}(s)R_{k-l}(s)B_l(s) - \sum_{k=N+1}^{\infty} Y_k(s), \\
q_N(s) &= \sum_{k=0}^{N}\sum_{l=0}^{k} Y_{N-k}(s)R_{k-l}(s)\left[\overline{A}_{l+1}(s)\left(\overline{A}_0(s)\right)^{-1} c_N(s) - c_{N-l}(s)\right] \\
&\quad - \sum_{k=0}^{N} R_{N-k}(s)\left[\overline{A}_{k+1}(s)\left(\overline{A}_0(s)\right)^{-1} c_N(s) - c_{N-k}(s)\right] + x(s).
\end{aligned} \tag{3.23}$$

Finally, rewriting (3.15) with (3.20) and (3.22) we have proven the following theorem.

Theorem 3.1. *The Laplace transform of the average number of losses in* $(0,t]$ *in a finite-capacity queue with the batch Markovian arrivals is equal to*

$$\begin{aligned}\delta_n(s) &= R_{N-n+1}(s)A_0(s)Q_N^{-1}(s)q_N(s)\\ &\quad+\sum_{k=0}^{N-n}R_{N-n-k}(s)\left[B_k(s)Q_N^{-1}(s)q_N(s)+\overline{A}_{k+1}(s)\left(\overline{A}_0(s)\right)^{-1}c_N(s)-c_{N-k}(s)\right].\end{aligned}\tag{3.24}$$

It should be stressed that (3.24) can be easily used to obtain numerical results. It is connected with the fact that all the matrices and vectors that appear in (3.24) are either simple functions of the BMAP parameters, or simple functions of matrices $A_k(s)$ and $E_k(s)$ defined in (3.9) and (3.10), respectively. Fortunately, matrices (3.9) and (3.10) are well known in the theory of BMAPs and can be computed using, for instance, formulas (65)–(67) from [13]. Finally, for practical purposes we are rather interested in $\Delta_{n,i}(t)$ than in its Laplace transform. To obtain originals from (3.24), one of the many available Laplace inversion formulas can be used. We use and recommend the formula based on the Euler summation. It can be found in [20].

Theorem 3.1 describes the average number of losses in $(0,t]$ interval. We may also be interested in the local intensity of the loss process. It can be obtained simply by differentiating $\Delta_{n,i}(t)$. Namely, denoting the local loss intensity by $K_{n,i}(t)$,

$$K_{n,i}(t)=\frac{d\Delta_{n,i}(t)}{dt},\qquad K_n(t)=\left(K_{n,1}(t),\ldots,K_{n,m}(t)\right)^T,\tag{3.25}$$

its Laplace transform by $\kappa_n(t)$,

$$\kappa_n(s)=\int_0^\infty e^{-st}K_n(t)dt,\tag{3.26}$$

and using the properties of the Laplace transform we obtain the following corollary.

Corollary 3.2. *The Laplace transform of the transient intensity of the loss process in a finite-capacity queue with the batch Markovian arrivals is equal to*

$$\kappa_n(s)=s\delta_n(s),\tag{3.27}$$

where $\delta_n(s)$ *is given in* (3.24).

Naturally, the numerical values of $K_n(t)$ can be obtained in the same way as described below Theorem 3.1.

Now, Theorem 3.1 and Corollary 3.2 describe the transient behaviour of the loss process. However, they can be also exploited to obtain stationary characteristics. The most important stationary characteristic is the loss ratio, L, defined as a long-run fraction of jobs that were lost. The loss ratio can be obtained using the following limit

$$L=\lim_{t\to\infty}\frac{K_{n,i}(t)}{\Lambda}.\tag{3.28}$$

Instead of computing $K_{n,i}(t)$, we can obtain this limit directly from (3.24), using the properties of the Laplace transform again. As the limit depends neither on n nor i, we can use, for instance, $n = N$ and $i = 1$. From (3.22) it follows that

$$\delta_N(s) = Q_N^{-1}(s)q_N(s) \cdot (1, 0 \ldots, 0), \tag{3.29}$$

which gives the following corollary.

Corollary 3.3. *The stationary loss ratio in a finite-capacity queue with the batch Markovian arrivals is equal to*

$$L = \lim_{s\to 0+} \frac{s^2 Q_N^{-1}(s)q_N(s) \cdot (1, 0 \ldots, 0)}{\Lambda}. \tag{3.30}$$

Note that (3.30) can be used to obtain quickly the numerical value of L, without applying the transform inversion.

4. Numerical Examples

4.1. Example 1

In the first example we will see how the stationary loss ratio varies with the traffic intensity, autocorrelation structure, and the buffer size. For that purpose, we will consider three arrival processes (all of them have the same average batch size and the total arrival rate, but differ in the autocorrelation structure).

$BMAP_1$: this is in fact a simple batch Poisson process, with batch arrivals of size 1, 4, and 10, $p_1 = 2/30$, $p_4 = 7/30$, $p_{10} = 21/30$, and the rate of batch arrivals of 0.125. It is easy to check that the average batch size is 8, and the total arrival rate is 1. The batch Poisson process is chosen here as an example of BMAP with no autocorrelation, that is, $\mathrm{Corr}(k) \equiv 0$.

$BMAP_2$: It is parameterized by the following matrices:

$$\begin{gathered}
D_0 = \begin{bmatrix} -5.69920 & 0.244077 & 0.0244077 \\ 0.00244077 & -0.569920 & 0.0244077 \\ 0.000244077 & 0.00244077 & -0.0569920 \end{bmatrix}, \\
D_1 = \begin{bmatrix} 0.00813590 & 0.0650872 & 0.650872 \\ 0.0650872 & 0.0000813590 & 0.00724095 \\ 0.00634600 & 0.000813590 & 0.0000813590 \end{bmatrix}, \\
D_4 = \begin{bmatrix} 0.00813590 & 0.0650872 & 0.650872 \\ 0.0650872 & 0.0000813590 & 0.00724095 \\ 0.00634600 & 0.000813590 & 0.0000813590 \end{bmatrix}, \\
D_{10} = \begin{bmatrix} 0.0447475 & 0.357980 & 3.57980 \\ 0.357980 & 0.000447475 & 0.0398252 \\ 0.0349030 & 0.00447475 & 0.000447475 \end{bmatrix}.
\end{gathered} \tag{4.1}$$

Table 1: The loss ratio for different arrival processes and system loads $N = 50$.

Arrival process	$\rho = 0.5$	$\rho = 1$	$\rho = 1.5$
BMAP_1	0.000870	0.087475	0.335441
BMAP_2	0.028960	0.173523	0.361348
BMAP_3	0.092648	0.324134	0.454306

Again, we have batch arrivals of size 1, 4, and 10. The matrices were carefully chosen so that the average batch size is 8 and the total arrival rate is 1 again. However, we have now a correlation between interarrival times. The autocorrelation function for BMAP_2 is depicted in Figure 1. As we can see, this is an example of the autocorrelation with alternating signs.

$BMAP_3$: It is parameterized by the following matrices:

$$D_0 = \begin{bmatrix} -0.0499514 & 0.00399715 & 0.00128940 \\ 0.00528656 & -0.0774334 & 0.00528656 \\ 0.00141834 & 0.00141834 & -0.274511 \end{bmatrix},$$
$$D_1 = \begin{bmatrix} 0.0181806 & 0.00141834 & 0.00270775 \\ 0.00141834 & 0.00399715 & 0.00270775 \\ 0.00270775 & 0.00399715 & 0.00657596 \end{bmatrix},$$
$$D_4 = \begin{bmatrix} 0.00141834 & 0.00270775 & 0.00141834 \\ 0.00141834 & 0.0413899 & 0.00141834 \\ 0.00528656 & 0.00141834 & 0.00270775 \end{bmatrix},$$
$$D_{10} = \begin{bmatrix} 0.00483682 & 0.00714007 & 0.00483682 \\ 0.00253357 & 0.00944332 & 0.00253357 \\ 0.00714007 & 0 & 0.241841 \end{bmatrix}. \quad (4.2)$$

As in the previous processes, the matrices were chosen so that the average batch size is 8 and the total arrival rate is 1. This time we have a strong positive autocorrelation between interarrival times. The autocorrelation function for BMAP_3 is depicted in Figure 2.

As for the service process, we assume that the service time is constant and denoted by d. Therefore, manipulating d we can manipulate the load of the system, that is,

$$\rho = \Lambda d. \quad (4.3)$$

The system capacity $N = 50$ is assumed. Now we can present numerical results.

Firstly, in Table 1 the loss ratio for the three considered BMAPs and three distinct loads of the system is presented. As expected, the highest values of the loss ratio are obtained for high ρ and positively autocorrelated BMAP. A more surprising thing is that even for a very low load (0.5), we can obtain a very high loss ratio (9.2%). Another interesting observation is that the loss ratio for BMAP_2, that is, in the case of alternating autocorrelation, is much higher than in the case of flat autocorrelation (BMAP_1). The detailed dependence of the loss ratio on the system load for the three considered BMAPs is depicted in Figure 3.

Secondly, in Figures 4, 5, and 6 the loss ratio as a function of the system capacity is presented for $\rho = 0.5$, $\rho = 1$, and $\rho = 1.5$, respectively. As we can see, the loss ratio decreases exponentially with the buffer size for $\rho < 1$ and subexponentially for $\rho = 1$. A very interesting

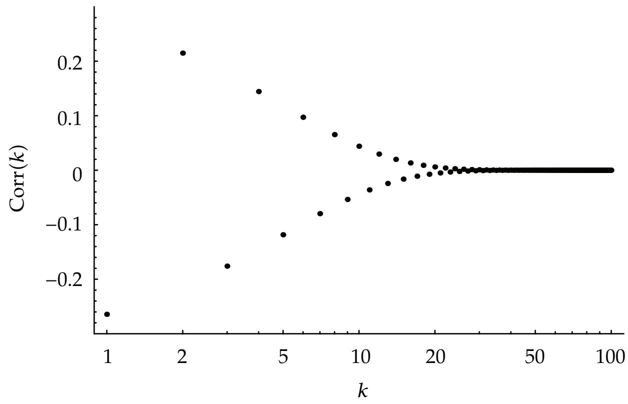

Figure 1: The autocorrelation at lag k of the sequence of interarrival times in BMAP_2.

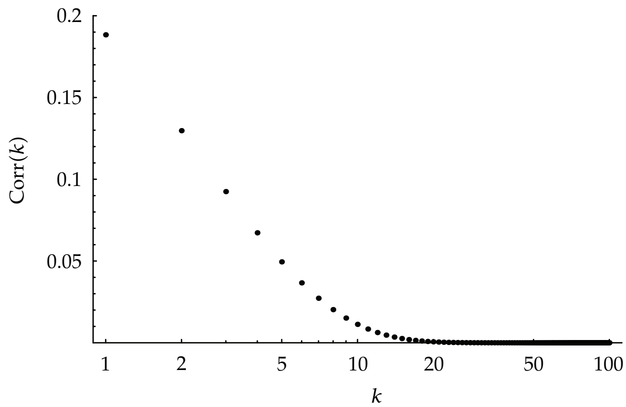

Figure 2: The autocorrelation at lag k of the sequence of interarrival times in BMAP_3.

fact is that the BMAP_2 and BMAP_3 curves cross somewhere in the interval $(10, 20)$. This means that if the system capacity is, for instance, 10, then BMAP_2 causes more losses than BMAP_3. On the other hand, if the system capacity is 30, then BMAP_3 causes more losses than BMAP_2. This counterintuitive behaviour can be explained by computing the variances of the interarrival times for both processes. We obtain $\text{Var}_2 = 208.75$ and $\text{Var}_3 = 151.49$, that is, BMAP_2 has a greater variance than BMAP_3. For small N, the impact of the variance on the loss ratio prevails and we observe more losses in the BMAP_2 case. On the other hand, for large N, the autocorrelation prevails and more losses are caused by BMAP_3.

4.2. Example 2

In the second example we want to observe the dependence of the loss ratio on the average batch size. For this purpose, we consider family $\text{BMAP}_4(k)$ of BMAPs. In this family we have BMAPs with batch arrivals of size k, $2k$, $3k$.

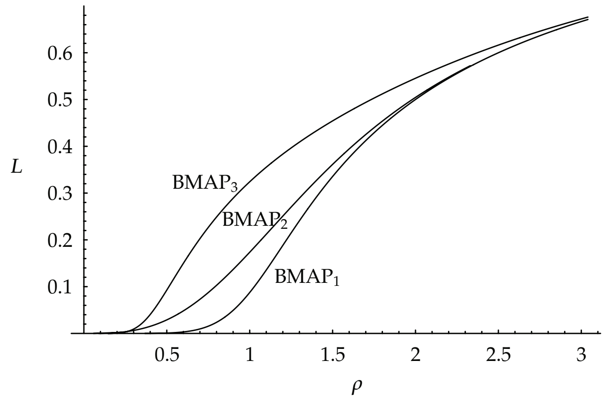

Figure 3: The loss ratio versus the load of the system. $N = 50$.

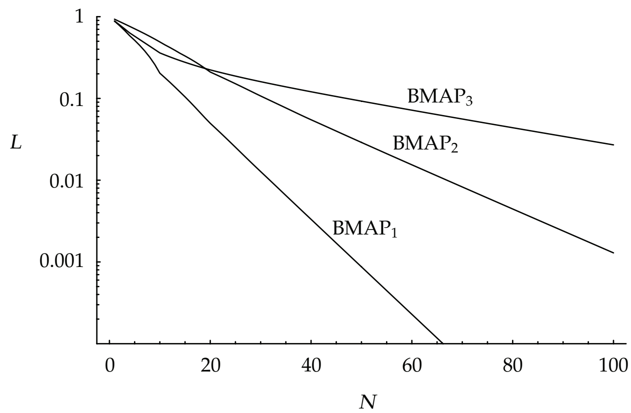

Figure 4: The loss ratio versus the capacity of the system. $\rho = 0.5$.

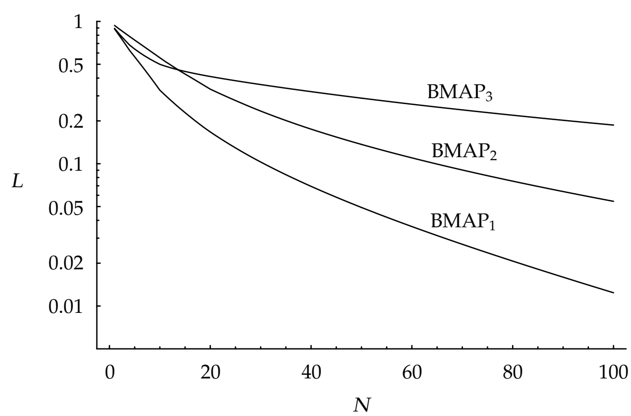

Figure 5: The loss ratio versus the capacity of the system. $\rho = 1$.

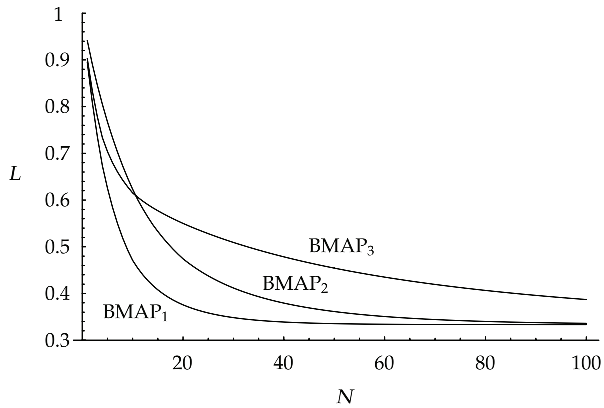

Figure 6: The loss ratio versus the capacity of the system. $\rho = 1.5$.

$BMAP_4(k)$: is parameterized by the following matrices:

$$D_0 = \begin{bmatrix} -0.387361 & 0.00958278 & 0.00309122 \\ 0.0126740 & -0.384279 & 0.0126740 \\ 0.00340034 & 0.00340034 & -0.832092 \end{bmatrix},$$

$$D_k = \begin{bmatrix} 0.250860 & 0.0195707 & 0.0373622 \\ 0.0195707 & 0.0551537 & 0.0373622 \\ 0.0373622 & 0.0551537 & 0.0907367 \end{bmatrix},$$

$$D_{2k} = \begin{bmatrix} 0.00680068 & 0.0129831 & 0.00680068 \\ 0.00680068 & 0.198456 & 0.00680068 \\ 0.0253480 & 0.00680068 & 0.0129831 \end{bmatrix}, \tag{4.4}$$

$$D_{3k} = \begin{bmatrix} 0.0115958 & 0.0171176 & 0.0115958 \\ 0.00607399 & 0.0226394 & 0.00607399 \\ 0.0171176 & 0 & 0.579790 \end{bmatrix}.$$

The average batch size for BMAP$_4(k)$ is $\beta = 2k$; the total arrival rate is k. As we want to maintain the same load, $\rho = 1$, for every BMAP$_4(k)$, we have to scale the service time to $1/k$.

The resulting loss ratio as a function of the average batch size is depicted in Figure 7. The results were computed for two system capacities, $N = 20$ and $N = 50$. For other values of ρ, the shape of this function is similar, except for the fact that it becomes more flat as ρ grows and vice versa.

4.3. Example 3

In the third example we will present the transient characteristics of the loss process. BMAP$_3$ with $\rho = 0.9$ and $N = 50$ will be used. Naturally, in the transient case the loss characteristics depend on the initial queue length, n, and the initial phase of the modulating process, i.

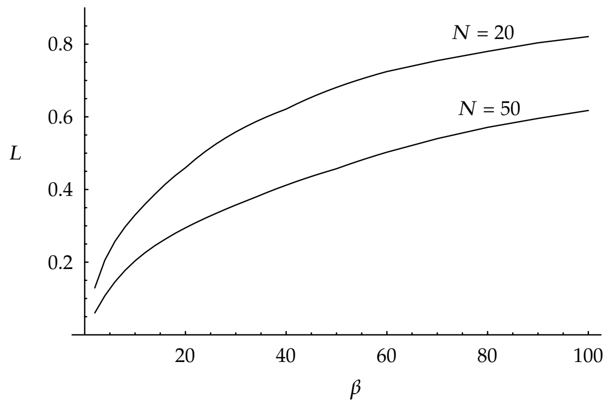

Figure 7: The loss ratio versus the average batch size for two different system capacities.

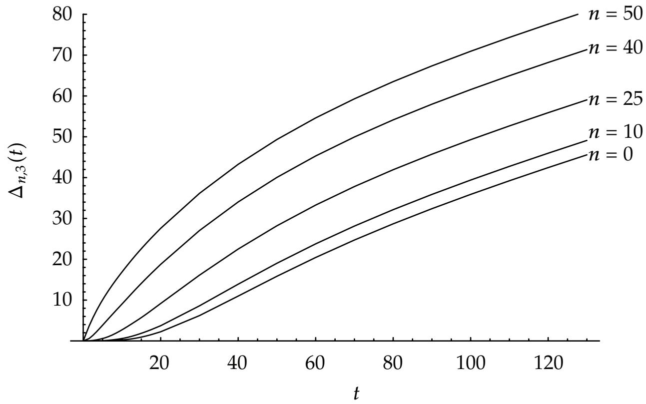

Figure 8: The average number of losses in $(0, t]$ as a function of t for $n = 0, 10, 25, 40$, and 50.

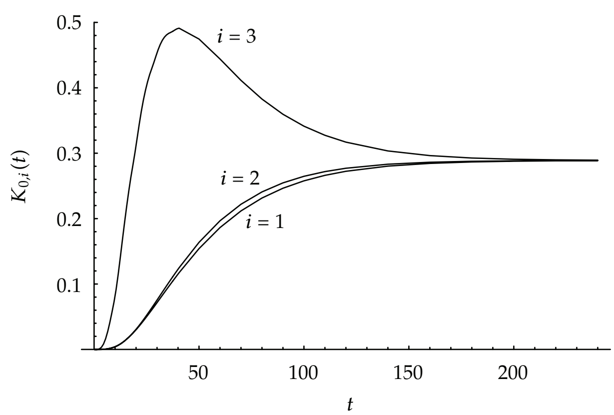

Figure 9: The intensity of the loss process in time for $i = 1$, 2, and 3.

Table 2: The loss ratios obtained from analysis and simulations.

Arrival process	Analytical results	Simulation results
$BMAP_1$	0.087475	0.087430
$BMAP_2$	0.173523	0.173378
$BMAP_3$	0.324134	0.324022
$BMAP_4(1)$	0.060252	0.060328

In Figure 8, the average number of losses in $(0, t]$ is presented as a function of t in five cases, when the initial queue size is 0, 10, 25, 40, and 50 jobs. In every case initial $i = 3$ was set. For the visual interpretation, it is easier to use the transient intensity of the loss process, $K_{n,i}(t)$. In Figure 9 this intensity is depicted in time for all three values of the initial phase of the modulating process. In every case initial $n = 0$ was set (empty system). Two interesting observations can be made using Figure 9. Firstly, for some initial conditions, the intensity of the loss process may not change monotonically. Here for $i = 3$ we have a maximum at t around 40 s. Secondly, we can tell more or less when the transient period is finished. Namely, after about 200 s, the loss intensity gets very close to the stationary value (which is $L = 0.288951$), no matter what the initial i was.

4.4. Example 4

In order to check the analytical results for possible mistakes, we have also performed a number o simulations and compared the simulation and the analytical results. For that purpose OMNeT++ discrete event simulator [21] was used. All the BMAPs appearing in Examples 1–3 were simulated; the service time was set to 1, the system capacitywas set to 50. In each simulation run, 10^8 jobs passing through the queueing system were simulated. The results are gathered in Table 2. As we can see, the analytical results agree very well with simulations.

5. Conclusions

In this paper we presented transient and stationary characterizations of the loss process in a finite-buffer queue fed by the batch Markovian arrivals. Due to the flexibility of the arrival process, the obtained results enable modeling of losses in many real-life queueing systems, with several properties influencing the number of losses, for example, the autocorrelation function, the batch size distribution, the interarrival time variance, and others.

The analytical results were presented in closed, easy to use formulas and accompanied by sample numerical calculations, demonstrating their applicability.

Acknowledgment

The material is based upon a work supported by the Polish National Science Centre under Grant no. N N516 479240.

References

[1] http://www.internettrafficreport.com/.
[2] H. Takagi, *Queueing Analysis—Finite Systems*, vol. 2, North-Holland Publishing, Amsterdam, The Netherlands, 1993.
[3] N. K. Kim and K. C. Chae, "Transform-free analysis of the $GI/G/1/K$ queue through the decomposed Little's formula," *Computers & Operations Research*, vol. 30, no. 3, pp. 353–365, 2003.
[4] D. R. Manfield and P. Tran-Gia, "Analysis of a finite storage system with batch input arising out of message packetization," *IEEE Transactions on Communications*, vol. 30, no. 3, pp. 456–463, 1982.
[5] A. Chydzinski, R. Wojcicki, and G. Hryn, "On the number of losses in an MMPP queue," in *Next Generation Teletraffic and Wired/Wireless Advanced Networking*, vol. 4712 of *Lecture Notes in Computer Science*, pp. 38–48, 2007.
[6] F. N. Gouweleeuw, "The loss probability in finite-buffer queues with batch arrivals and complete rejection," *Probability in the Engineering and Informational Sciences*, vol. 8, pp. 221–227, 1994.
[7] H. C. Tijms, "Heuristics for finite-buffer queues," *Probability in the Engineering and Informational Sciences*, vol. 6, no. 3, pp. 227–285, 1992.
[8] M. Bratiychuk and A. Chydzinski, "On the loss process in a batch arrival queue," *Applied Mathematical Modelling*, vol. 33, no. 9, pp. 3565–3577, 2009.
[9] A. Baiocchi and B. MElazzi, "Steady-state analysis of the MMPP/G/1/K queue," *IEEE Transactions on Communications*, vol. 41, no. 4, pp. 531–534, 1993.
[10] R. Nagarajan, J. F. Kurose, and D. Towsley, "Approximation techniques for computing packet loss in finite-buffered voice multiplexers," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 3, pp. 368–377, 1991.
[11] N. B. Shroff and M. Schwartz, "Improved loss calculations at an ATM multiplexer," *IEEE/ACM Transactions on Networking*, vol. 6, no. 4, pp. 411–421, 1998.
[12] A. N. Dudin, A. A. Shaban, and V. I. Klimenok, "Analysis of a queue in the BMAP/G/1/N system," *International Journal of Simulation*, vol. 6, no. 1-2, pp. 13–23, 2005.
[13] D. M. Lucantoni, "New results on the single server queue with a batch Markovian arrival process," *Communications in Statistics*, vol. 7, no. 1, pp. 1–46, 1991.
[14] L. Breuer, "An EM algorithm for batch Markovian arrival processes and its comparison to a simpler estimation procedure," *Annals of Operations Research*, vol. 112, pp. 123–138, 2002.
[15] A. Klemm, C. Lindemann, and M. Lohmann, "Modeling IP traffic using the batch Markovian arrival process," *Performance Evaluation*, vol. 54, no. 2, pp. 149–173, 2003.
[16] P. Salvador, A. Pacheco, and R. Valadas, "Modeling IP traffic: Joint characterization of packet arrivals and packet sizes using BMAPs," *Computer Networks*, vol. 44, no. 3, pp. 335–352, 2004.
[17] S. Chakravarthy, "The batch markovian arrival process: a review and future work," in *Advances in Probability and Stochastic Processes*, A. Krishnamoorthy et al., Ed., pp. 21–49, Notable Publications, Neshanic Station, NJ, USA, 2001.
[18] A. Chydzinski, "The oscillating queue with finite buffer," *Performance Evaluation*, vol. 57, no. 3, pp. 341–355, 2004.
[19] A. Chydzinski, "Time to reach buffer capacity in a BMAP queue," *Stochastic Models*, vol. 23, no. 2, pp. 195–209, 2007.
[20] J. Abate, G. L. Choudhury, and W. Whitt, "An introduction to numerical transform inversion and its application to probability models," in *Computational Probability*, W. Grassman, Ed., pp. 257–323, Kluwer, Boston, Mass, USA, 2000.
[21] http://www.omnetpp.org/.