Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/070450

Buffer overflow period in a batch-arrival queue with autocorrelated arrivals

*Andrzej Chydzinski**

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

Received: 18 Jan. 2013, Revised: 20 Mar. 2013, Accepted: 20 Mar. 2013
Published online: 1 Jul. 2013

Abstract: In this paper a study on the buffer overflow period in a finite-buffer queue is presented. A special attention is paid to the autocorrelation and batch arrivals, which are supposed to mimic the properties of real arrival processes. Using the batch Markovian arrival process, formulas for the duration of the first, subsequent and stationary overflow periods, as well as formulas for the distribution of the number of consecutive losses during the overflow period are shown. Moreover, analytical results are illustrated via numerical examples. In particular, the influence of the autocorrelation and batch arrivals on the duration of the overflow period is demonstrated.

Keywords: buffer overflow period, remaining service time, batch Markovian arrival process

1 Introduction

The performance of a finite-buffer queueing system depends on losses, i.e. jobs that were rejected the service due to the buffer overflow. The most important characteristics of the loss process include the stationary loss ratio (i.e. the long-run fraction of lost jobs), the number of losses in a finite time interval and the duration of the buffer overflow period.

The buffer overflow period is the time interval from the arrival of a job which overflows the buffer until the first departure of a job from the system (making available one position in the buffer). It is equal to the remaining service time of the job being served upon arrival of the job which overflows the buffer.

The importance of the buffer overflow period is connected with the fact that during this period all arriving jobs are lost, one after another. Therefore, the distribution of the number of consecutive losses depends on the distribution of the duration of the overflow period. On the other hand, in many queueing applications, especially in packet networking (e.g. the Internet), the distribution of the number of consecutive losses plays an important role. Roughly speaking, it makes a difference whether the packets are being lost in long series or in short series with accepted packets between them.

Therefore, the goal of this paper is finding the distribution of the length of the buffer overflow as well as the distribution of the number of jobs lost in the overflow period. In contrast to the previous papers on this subject, two important assumptions on the arrival process are imposed herein:

(a)the interarrival times are correlated,
(b)the arrivals can occur in groups (batches).

To the best of the author's knowledge, there are no published studies on the buffer overflow period in queueing systems that meet requirements (a) and (b). Several papers devoted to the buffer overflow period in the single-arrival queues can be found, [1]-[7], but the systems studied there do not meet (b). There is also a paper on the buffer overflow period in batch arrival queues, [8], but with uncorrelated arrival process. Taking into account both of the aforementioned requirements is crucial in the performance evaluation of queues of packets appearing in the packet networking. In particular, it is well-known that the streams of packets are usually strongly autocorrelated, so the Poisson processes (or other renewal processes) do not reflect properly their characteristics, [9,10]. This explains the need for an autocorrelated arrival process model. The need for the batch structure of arrivals is caused by the popularity of the TCP[1] protocol, which is responsible for the major part of the Internet traffic. In TCP, the packets are sent to the network in groups, according to a variable called *window*,

[1] Transmission Control Protocol

* Corresponding author e-mail: andrzej.chydzinski@polsl.pl

which denotes the number of packets that can be injected into the network without confirmation. (The reader further interested in modeling the TCP behaviour using batch arrivals is referred to [11]).

As will be demonstrated in Section 6, the batch structure of arrivals has a deep impact on the buffer overflow period. Namely, the dependence of the overflow period on the buffer size may have a very irregular form – changing the buffer size only by 1 position may result in a significant change in the overflow period. This effect does not appear in the single-arrival systems studied in [1]-[7].

As for the arrival process, the batch Markovian arrival process (BMAP, [12,13]) is the most suitable for our needs. It meets requirements (a), (b), and can be quite easily parameterized to reflect the parameters of real arrival processes (packet streams and others).

The remaining part of the paper is organized as follows. In Section 2, all details of the queueing model are given. In Section 3, the first buffer overflow period is studied. In particular, Theorem 1 of this section presents the distribution of the length of the first buffer overflow period depending on the initial queue size and the state of the modulating process. In Section 4, the subsequent and stationary overflows are studied using an embedded Markov chain, whose transition matrix is given in Theorem 2. In Section 5, formulas for the number of consecutive losses during the first, subsequent and stationary overflow period are proven. Then, in Section 6, numerical examples are presented. They include examples of the dependence of the duration of the overflow period on the buffer size, the initial queue length, the system load and the autocorrelation structure. The final conclusions are gathered in Section 7.

2 Model

The batch Markovian arrival process (BMAP, [12,13]), is constructed by using a two-dimensional Markov process $(N(t),J(t))$, where variable $N(t)$ is the total number of job arrivals in $(0,t)$, while variable $J(t)$ is the auxiliary state of the modulating Markov process. The Markov process $(N(t),J(t))$ assumes values in state space $\{(i,j) : i \geq 0, 1 \leq j \leq m\}$ and has the following infinitesimal generator:

$$Q = \begin{bmatrix} D_0 & D_1 & D_2 & D_3 & \cdots \\ & D_0 & D_1 & D_2 & \cdots \\ & & D_0 & D_1 & \cdots \\ & & & \ddots & \ddots \end{bmatrix},$$

where D_k, $k \geq 0$ are $m \times m$ matrices, D_k is nonnegative for $k \geq 1$, while D_0 has nonnegative off-diagonal elements and negative diagonal elements. Moreover,

$$D = \sum_{k=0}^{\infty} D_k$$

is an irreducible infinitesimal generator and it must not be equal to D_0.

Let $\mathbb{1}$ denote the column vector of 1's, $\mathbb{0}$ denote the square matrix of zeroes and π denotes the stationary vector for D, i.e.:

$$\pi D = \mathbb{0}, \quad \pi \mathbb{1} = 1.$$

Using π we can compute the total arrival rate of the BMAP:

$$\Lambda = \pi \sum_{k=1}^{\infty} k D_k \mathbb{1}, \tag{1}$$

and the rate of arrivals of batches:

$$\Lambda_g = \pi(-D_0)\mathbb{1}. \tag{2}$$

The BMAP can be also defined using a constructive definition, which can be easily translated into any programming language and used for simulation purposes. Namely, we initialize the modulating process, $J(t)$, using some state, say i (i.e. $J(0) = i$). The sojourn time in that state is exponentially distributed with parameter μ_i, where

$$\mu_i = -(D_0)_{ii}.$$

At the end of this sojourn time there occurs a transition to another state and/or the arrival of a batch of jobs. In particular, with probability $p_i(j,k)$ there occurs a transition to state k with arrival of a batch of size j, where:

$$p_i(0,i) = 0,$$

$$p_i(0,k) = \frac{1}{\mu_i}(D_0)_{ik}, \qquad 1 \leq i,k \leq m, \quad k \neq i,$$

$$p_i(j,k) = \frac{1}{\mu_i}(D_j)_{ik}, \qquad 1 \leq i,k \leq m, \quad j \geq 1.$$

The probabilities $p_i(k,j)$ will be often used in the sequel, as well as the matrices

$$Y_k = [p_i(k,j)]_{i,j}. \tag{3}$$

An important characteristic of the BMAP is its counting function, $P(n,t)$, which is an $m \times m$ matrix defined as

$$P(n,t) = \left[P_{i,j}(n,t)\right]_{i,j},$$

$$P_{i,j}(n,t) = \mathbb{P}\{N(t) = n, J(t) = j | N(0) = 0, J(0) = i\},$$

where $\mathbb{P}$ stands for the probability.

Other important characteristics of the BMAP are matrices A_k, describing the distribution of the number of arrivals and change of the state of the modulating process during a random interval distributed according to a distribution function F, namely

$$A_k = \left[\int_0^{\infty} P_{i,j}(k,t)dF(t)\right]_{i,j}. \tag{4}$$

Now, having defined the arrival process we can present the whole queueing model. Namely, we will deal with the classic single-server queueing system with finite buffer, whose arrival process is a BMAP. The service time is distributed according to a distribution function $F(\cdot)$, which is not further specified. The buffer size (system capacity) is finite and equal to b, including the service position. If a job upon its arrival finds the buffer full, it is lost and never returns. We assume also that the time origin corresponds to a departure epoch. In the standard Kendall's notation, such queueing system is denoted by $BMAP/G/1/b$.

The queue size at time t will be denoted by $X(t)$. We adopt a convention that $X(t)$ includes the service position. Thus $X(t)$ is, in other words, the total number of jobs present in the system at time t.

3 First overflow

We will start with analysis of the first overflow period using the method of [14]. This method exploits special recurrent sequences to solve the system of equations occurring in the analysis of the model. The method can be used in solving classic queueing models with different Markovian arrival processes (e.g. [14,15]) as well as some special models (e.g. [16]), provided that the buffer size is finite.

The length of the first buffer overflow period will be denoted as β_1 and is defined as

$$\beta_1 = \zeta - \tau,$$

where τ is the first time when the buffer gets full and ζ is the first job departure moment after τ. Naturally, distributions of random variables ζ, τ and β_1 depend on the initial queue size, $X(0)$, and the initial state of the modulating process, $J(0)$.

Let $\phi_{n,i}(t)$ be the tail of the distribution of β_1, i.e.:

$$\phi_{n,i}(t) = \mathbb{P}\{\beta_1 > t | X(0) = n, J(0) = i\}$$

and

$$\phi_n(t) = [\phi_{n,1}(t), \ldots, \phi_{n,m}(t)]^T.$$

Now we will prove the following theorem.

Theorem 1 *Distribution of the length of the first buffer overflow period in the $BMAP/G/1/b$ queue has the following form:*

$$\phi_n(t) = -\sum_{k=1}^{b-n} R_{b-n-k} g_k(t) + \sum_{k=0}^{b-n} R_{b-n-k} A_k \left(\sum_{k=0}^{b} R_{b-k} A_k - \sum_{k=1}^{b} \sum_{l=0}^{k} Y_{b-k} R_{k-l} A_l \right)^{-1} \cdot \left(\sum_{k=1}^{b} R_{b-k} g_k(t) - \sum_{k=1}^{b} \sum_{l=1}^{k} Y_{b-k} R_{k-l} g_l(t) + s_b(t) \right), \quad (5)$$

where

$$R_0 = \mathbb{O}, \quad R_1 = A_0^{-1},$$

$$R_{k+1} = R_1 (R_k - \sum_{i=0}^{k} A_{i+1} R_{k-i}), \quad k \geq 1, \quad (6)$$

$$g_k(t) = (1 - F(t)) \cdot \mathbb{1} - \sum_{i=0}^{k-1} \int_t^{\infty} P(i, u-t) dF(u) \cdot \mathbb{1}, \quad (7)$$

$$s_b(t) = (1 - F(t)) \sum_{k=b}^{\infty} Y_k \cdot \mathbb{1}, \quad (8)$$

A_k, Y_k are defined in (4), (3), respectively.

P r o o f of Theorem 1. Assuming that the system is non-empty at $t = 0$, we can write the following system of equations:

$$\mathbb{P}\{ \beta_1 > t \, | X(0) = n, J(0) = i\} = \sum_{j=1}^{m} \sum_{k=0}^{b-n-1} \int_0^{\infty} \mathbb{P}\{\beta_1 > t | X(0) = n+k-1, J(0) = j\} \cdot P_{i,j}(k,u) dF(u) + \int_t^{\infty} \big(1 - \sum_{j=1}^{m} \sum_{k=0}^{b-n-1} P_{i,j}(k, u-t)\big) dF(u), \quad 0 < n < b, 1 \leq i \leq m. \quad (9)$$

System (9) follows from the total probability law, used with respect to the first departure moment. In particular, the first component of (9) reflects the situation where there is no buffer overflow before the first job departure time, u. The second component reflects the situation, where the buffer overflow occurs before time u. In fact, in order to have $\beta_1 > t$, the buffer overflow must occur in time interval $(0, u-t)$ with $u > t$. The probability that the buffer gets overflowed in interval $(0, u-t)$ is equal to $1 - \sum_{j=1}^{m} \sum_{k=0}^{b-n-1} P_{i,j}(k, u-t)$.

Assuming that the system is empty at $t = 0$, we can write the following system:

$$\mathbb{P}\{\beta_1 > t \;|\; X(0) = 0, J(0) = i\} = \sum_{j=1}^{m} \sum_{k=0}^{b-1} p_i(k,j) \mathbb{P}\{\beta_1 > t | X(0) = k, J(0) = j\} + \sum_{j=1}^{m} \sum_{k=b}^{\infty} p_i(k,j)(1 - F(t)), \quad 1 \leq i \leq m. \quad (10)$$

System (10) follows from the total probability law used with respect to the size of the first arriving batch. In particular, the first component corresponds to the case where the first arriving batch does not overflow the buffer, while the second component corresponds to the case where the first batch overflows the buffer.

Using matrix notation to (9) yields:

$$\phi_n(t) = \sum_{k=0}^{b-n-1} A_k \phi_{n+k-1}(t) + g_{b-n}(t). \quad (11)$$

Similarly, (10) yields:

$$\phi_0(t) = \sum_{k=0}^{b-1} Y_k \phi_k(t) + s_b(t). \quad (12)$$

Using substitution $\varphi_n(t) = \phi_{b-n}(t)$ we get:

$$\sum_{k=-1}^{n-1} A_{k+1}\varphi_{n-k}(t) - \varphi_n(t) = \psi_n(t), \qquad 0 < n < b, \quad (13)$$

$$\psi_n(t) = A_n\varphi_1(t) - g_n(t),$$

and

$$\varphi_b(t) = \sum_{k=0}^{b-1} Y_k \varphi_{b-k}(t) + s_b(t) = \sum_{k=1}^{b} Y_{b-k}\varphi_k(t) + s_b(t). \quad (14)$$

Using Lemma 3.2.1 of [14] we can obtain the general solution of system (13) in the form:

$$\varphi_n(t) = R_n c(t) + \sum_{k=1}^{n} R_{n-k}\psi_k(t), \qquad n \geq 1, \quad (15)$$

where $c(t)$ is a column vector which does not depend on n. Substitution $n = 1$ into (15) yields $c(t) = A_0\varphi_1(t)$, which then gives

$$\varphi_n(t) = \sum_{k=0}^{n} R_{n-k}A_k\varphi_1(t) - \sum_{k=1}^{n} R_{n-k}g_k(t). \quad (16)$$

Now, substituting $n = b$ in (16) and exploiting (14) we obtain the equation

$$\begin{aligned}\Sigma_{k=0}^{b} R_{b-k}A_k\varphi_1(t) - \sum_{k=1}^{b} R_{b-k}g_k(t) \\ = \sum_{k=1}^{b} Y_{b-k}\left(\sum_{l=0}^{k} R_{k-l}A_l\varphi_1(t) - \sum_{l=1}^{k} R_{k-l}g_l(t)\right) + s_b(t),\end{aligned} \quad (17)$$

which allows us finding $\varphi_1(t)$:

$$\begin{aligned}\varphi_1(t) = \left(\sum_{k=0}^{b} R_{b-k}A_k - \sum_{k=1}^{b}\sum_{l=0}^{k} Y_{b-k}R_{k-l}A_l\right)^{-1} \\ \cdot \left(\sum_{k=1}^{b} R_{b-k}g_k(t) - \sum_{k=1}^{b}\sum_{l=1}^{k} Y_{b-k}R_{k-l}g_l(t) + s_b(t)\right).\end{aligned} \quad (18)$$

Backward substitution $\phi_n(t) = \varphi_{b-n}(t)$ in (16) finishes the proof.

In some applications it might be convenient to use the density function instead of the tail. It is easy to obtain using Theorem 1. Namely, density of the length of the first overflow period is:

$$\begin{aligned}h_n(t) = -\phi_n'(t) = \sum_{k=1}^{b-n} R_{b-n-k}g_k'(t) - \sum_{k=0}^{b-n} R_{b-n-k}A_k \\ \cdot \left(\sum_{k=0}^{b} R_{b-k}A_k - \sum_{k=1}^{b}\sum_{l=0}^{k} Y_{b-k}R_{k-l}A_l\right)^{-1} \\ \cdot \left(\sum_{k=1}^{b} R_{b-k}g_k'(t) - \sum_{k=1}^{b}\sum_{l=1}^{k} Y_{b-k}R_{k-l}g_l'(t) + s_b'(t)\right).\end{aligned} \quad (19)$$

4 Subsequent and stationary overflows

In this section we will analyze the length of the k-th overflow period, i.e. β_k, for arbitrary k, as well as the limiting distribution when $k \to \infty$.

Let ζ_k denote the end of the k-th overflow period and let α_k be the state of the modulating process at this time, i.e.:

$$\alpha_0 = J(0), \qquad \alpha_k = J(\zeta_k), \quad k \geq 1.$$

As we may note, the distribution of β_k depends on the state of the system (i.e. the queue length and the state of the modulating process) at time ζ_{k-1}. However, the queue size at time ζ_k is equal to $b-1$ for every $k \geq 1$. Therefore, the distribution of β_k depends in fact on the state of modulating process at time ζ_{k-1} only. Moreover, the distribution of β_k for some k is the same as the distribution of β_1 assuming the initial queue size of $b-1$ and proper initial state of the modulation process. Namely, for $k \geq 2$ we have

$$\mathbb{P}\{\beta_k > t|\alpha_{k-1} = i\} = \mathbb{P}\{\beta_1 > t|X(0) = b-1, J(0) = i\}.$$

In this way, knowing the distribution of α_{k-1}, we can find the distribution of β_k by means of Theorem 1. Therefore, finding the distribution of α_k will be our primary goal in this section.

We will start with the distribution of α_1, which depends on $X(0)$ and $J(0)$. It will be presented in a matrix form:

$$S_n = \left[\mathbb{P}\{\alpha_1 = l|X(0) = n, J(0) = i\}\right]_{i,l}, \qquad 0 \leq n \leq b, \quad 1 \leq i,l \leq m.$$

Theorem 2 *Distribution of the state of the modulating process at the end of the first buffer overflow period in the BMAP/G/1/b queue has the form:*

$$\begin{aligned}S_n = -\sum_{k=1}^{b-n} R_{b-n-k}U_k + \sum_{k=0}^{b-n} R_{b-n-k}A_k \\ \cdot \left(\sum_{k=0}^{b} R_{b-k}A_k - \sum_{k=1}^{b}\sum_{l=0}^{k} Y_{b-k}R_{k-l}A_l\right)^{-1} \\ \cdot \left(\sum_{k=1}^{b} R_{b-k}U_k - \sum_{k=1}^{b}\sum_{l=1}^{k} Y_{b-k}R_{k-l}U_l + \sum_{k=b}^{\infty} Y_k U_0\right),\end{aligned} \quad (20)$$

where

$$U_k = \sum_{i=k}^{\infty} A_i. \tag{21}$$

P r o o f of Theorem 2. The proof is similar to the proof of Theorem 1. Namely, for the initially non-empty system we obtain

$$\begin{aligned}\mathbb{P}\{\ \alpha_1 = l\ |X(0) = n, J(0) = i\} \\ = \sum_{j=1}^{m}\sum_{k=0}^{b-n-1}\int_0^{\infty}\mathbb{P}\{\alpha_1 = l|X(0) = n+k-1, J(0) = j\} \\ \cdot P_{i,j}(k,u)dF(u) \\ + \sum_{k=n-b}^{\infty}\int_0^{\infty}P_{i,l}(k,u)dF(u), \\ 0 < n < b, 1 \le i \le m,\end{aligned} \tag{22}$$

while for initially empty system

$$\begin{aligned}\mathbb{P}\{\ \alpha_1 = l\ |X(0) = 0, J(0) = i\} \\ = \sum_{j=1}^{m}\sum_{k=0}^{b-1}p_i(k,j)\mathbb{P}\{\alpha_1 = l|X(0) = 0, J(0) = j\} \\ + \sum_{j=1}^{m}\sum_{k=b}^{\infty}p_i(k,j)\int_0^{\infty}\mathbb{P}\{J(u) = l|J(0) = j\}dF(u).\end{aligned} \tag{23}$$

Using the matrix notation, (22) and (23) yield:

$$S_n = \sum_{k=0}^{b-n-1} A_k S_{n+k-1} + U_{b-n}, \qquad 0 < n < b, \tag{24}$$

$$S_0 = \sum_{k=0}^{b-1} Y_k S_k + \sum_{k=b}^{\infty} Y_k U_0. \tag{25}$$

Proceeding in the same way as in the proof of Theorem 1 we can easily finish the proof.

In order to compute the distribution of α_k, $k \ge 2$, it is sufficient to note that $\{\alpha_k\}_{k=0}^{\infty}$ constitutes a discrete-time Markov chain, whose transition matrix is equal to S_n in the first step and S_{b-1} in every next step. Therefore, if v is the vector of the initial distribution of the modulating process, the distribution of α_k is $vS_n(S_{b-1})^{k-2}$, where n is the initial queue size. As a consequence, the distribution of β_k is

$$\mathbb{P}\{\beta_k > t\} = vS_n(S_{b-1})^{k-2}\phi_{b-1}(t), \tag{26}$$

where $\phi_{b-1}(t)$ is given in (5) and n is the initial queue size.

Using the transition matrix S_{b-1} we can also obtain the limiting distribution of β_k as $k \to \infty$. Namely, if w is the stationary vector for S_{b-1}, i.e.

$$wS_{b-1} = w, \quad w\mathbb{1} = 1, \tag{27}$$

then we have

$$\lim_{k\to\infty}\mathbb{P}\{\beta_k > t\} = w\phi_{b-1}(t). \tag{28}$$

Consequently, the density of the length of the overflow period in the steady state is

$$h(t) = -w\phi'_{b-1}(t). \tag{29}$$

5 Number of losses during overflow periods

Let γ_k denote the number of loses in the k-th buffer overflow period and

$$q_{n,i}(l) = \mathbb{P}\{\gamma_1 = l|X(0) = n, J(0) = i\}, \quad l = 0,1,2,\ldots,$$

$$q_n(l) = [q_{n,1}(l),\ldots,q_{n,m}(l)]^T.$$

The following theorem is true.

Theorem 3*Distribution of the number of loses in the first buffer overflow period in the BMAP/G/1/b queue has the form:*

$$\begin{aligned}q_n(l) = -\sum_{k=1}^{b-n} R_{b-n-k}A_{k+l}\mathbb{1} \\ + \sum_{k=0}^{b-n} R_{b-n-k}A_k\left(\sum_{k=0}^{b} R_{b-k}A_k - \sum_{k=1}^{b}\sum_{j=0}^{k} Y_{b-k}R_{k-j}A_j\right)^{-1} \\ \cdot\left(\sum_{k=1}^{b} R_{b-k}A_{k+l}\mathbb{1} - \sum_{k=1}^{b}\sum_{j=1}^{k} Y_{b-k}R_{k-j}A_{j+l}\mathbb{1}\right. \\ \left. + \sum_{k=b}^{b+l} Y_k A_{b+l-k}\mathbb{1}\right).\end{aligned} \tag{30}$$

P r o o f of Theorem 3. For the initially non-empty system we obtain now

$$\begin{aligned}\mathbb{P}\{\ \gamma_1 = l\ |X(0) = n, J(0) = i\} \\ = \sum_{j=1}^{m}\sum_{k=0}^{b-n-1}\int_0^{\infty}\mathbb{P}\{\gamma_1 = l|X(0) = n+k-1, J(0) = j\} \\ \cdot P_{i,j}(k,u)dF(u) \\ + \sum_{j=1}^{m}\int_0^{\infty}P_{i,j}(b+l-n,u)dF(u). \\ 0 < n < b, \ 1 \le i \le m.\end{aligned}$$

For initially empty system we get

$$\begin{aligned}\mathbb{P}\{\ \gamma_1 = l\ |X(0) = 0, J(0) = i\} \\ = \sum_{j=1}^{m}\sum_{k=0}^{b-1}p_i(k,j)\mathbb{P}\{\gamma_1 = l|X(0) = k, J(0) = j\} \\ + \sum_{j=1}^{m}\sum_{k=b}^{b+l}p_i(k,j)\int_0^{\infty}\sum_{r=1}^{m}P_{j,r}(b+l-k,u)dF(u), \\ 1 \le i \le m.\end{aligned} \tag{31}$$

Therefore, we have

$$q_n(l) = \sum_{k=0}^{b-n-1} A_k q_{n+k-1}(l) + A_{b+l-n}\mathbb{1}, \quad 0 < n < b, \quad (32)$$

$$q_0(l) = \sum_{k=0}^{b-1} Y_k q_k(l) + \sum_{k=b}^{b+l} Y_k A_{b+l-k}\mathbb{1}. \quad (33)$$

Proceeding in the same way as in the proof of Theorem 1 we can finish the proof.

It is easily seen that the distribution of the number of losses in the k-th overflow period has the form:

$$\mathbb{P}\{\gamma_k = l\} = vS_n(S_{b-1})^{k-2}q_{b-1}(l), \quad (34)$$

where n is the initial queue size, while the number of jobs lost in the overflow period in steady state is:

$$\lim_{k\to\infty} \mathbb{P}\{\gamma_k = l\} = wq_{b-1}(l). \quad (35)$$

6 Examples

Let us note first that the values of the BMAP counting function, $P(n,t)$, as well as the integrals appearing in (4) and (7), can be effectively computed using the uniformization technique (see [12] for more details). Therefore we can effectively obtain numerical results from the presented theorems.

For numerical purposes, we use herein the following BMAP (see also [17]):

$$D_0 = \begin{bmatrix} -5.69920 & 0.244077 & 0.0244077 \\ 0.00244077 & -0.569920 & 0.0244077 \\ 0.000244077 & 0.00244077 & -0.0569920 \end{bmatrix},$$

$$D_1 = \begin{bmatrix} 0.00813590 & 0.0650872 & 0.650872 \\ 0.0650872 & 0.0000813590 & 0.00724095 \\ 0.00634600 & 0.000813590 & 0.0000813590 \end{bmatrix},$$

$$D_4 = \begin{bmatrix} 0.00813590 & 0.0650872 & 0.650872 \\ 0.0650872 & 0.0000813590 & 0.00724095 \\ 0.00634600 & 0.000813590 & 0.0000813590 \end{bmatrix},$$

$$D_{10} = \begin{bmatrix} 0.0447475 & 0.357980 & 3.57980 \\ 0.357980 & 0.000447475 & 0.0398252 \\ 0.0349030 & 0.00447475 & 0.000447475 \end{bmatrix}.$$

Thus the batches of sizes 1, 4 and 10 are possible. The total arrival rate is

$$\Lambda = 1, \quad (36)$$

the rate of arrivals of batches is

$$\Lambda_g = 0.125, \quad (37)$$

Table 1: Autocorrelation function for BMAP used in numerical examples.

lag	1	2	3	4	5
correlation	- 0.264	0.214	-0.176	0.144	-0.1183
lag	6	7	8	9	10
correlation	0.097	-0.079	0.065	-0.053	0.04

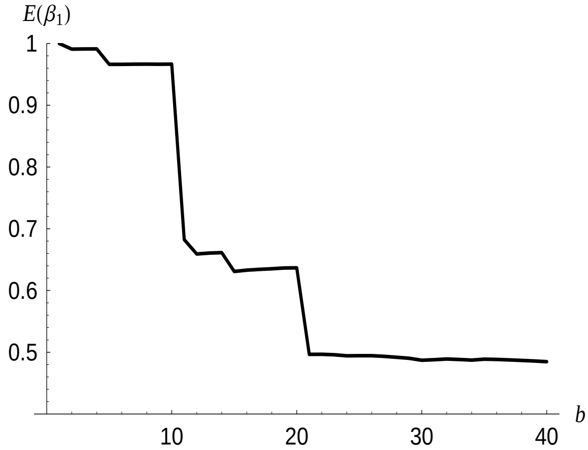

Fig. 1: The expected value of β_1 versus the buffer size for $X(0) = 0$ and $J(0) = 1$.

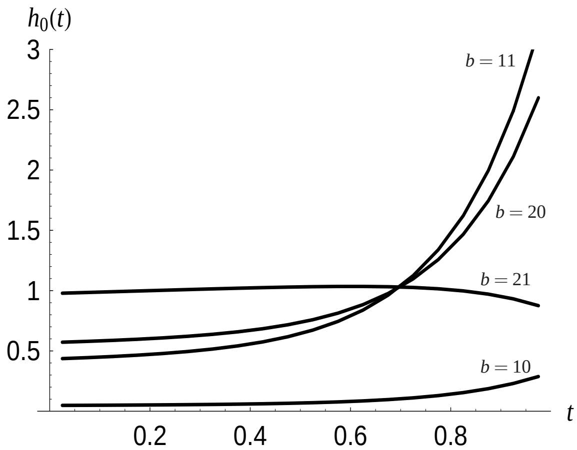

Fig. 2: The density of β_1 for four different buffer sizes, 10, 11, 20, 21, and initial system parameters $X(0) = 0$, $J(0) = 1$.

and the average batch size is 8. Constant service time, $d = 1$, is assumed, so the offered load is $\rho = \Lambda d = 1$. It is also worth noting that the considered BMAP has a non-trivial autocorrelation function, with alternating signs. The correlations for lags 1-10 are presented in Table 1.

In Fig. 1 the average duration of the first overflow period as a function of the buffer size is depicted. The system is initially empty and the initial state of the modulating process is 1. As we can see, the shape is irregular – there are jumps around the buffer sizes of 10 and 20. Naturally, it is connected with multiples of the largest batch size. This effect is further demonstrated in Fig. 2, in which the density function of the length of the first overflow period for buffer sizes of 10, 11, 20 and 21 is depicted. We can see large differences between $b = 10$ and $b = 11$ cases, as well as between $b = 20$ and $b = 21$. In fact, the shape of the density for $b = 11$ is closer to the case $b = 20$ than to the case $b = 10$. Note also that for $b = 10$ the density does not integrate to 1, which means that there is an atom in the probability distribution at $t = 1$. The existence of this atom is easy to explain by the fact, that for $b = 10$ one large batch may overflow the buffer.

Besides the buffer size, the duration of the first overflow period depends on the initial queue size, n, and the initial modulating state, j. Now we will check these dependences. In Fig. 3, the density functions of the length of the first overflow period for initial queue sizes of 0, 10 and 19 are compared while in Fig. 4 the densities of the length of the first overflow period for all possible initial states of the modulating process are compared.

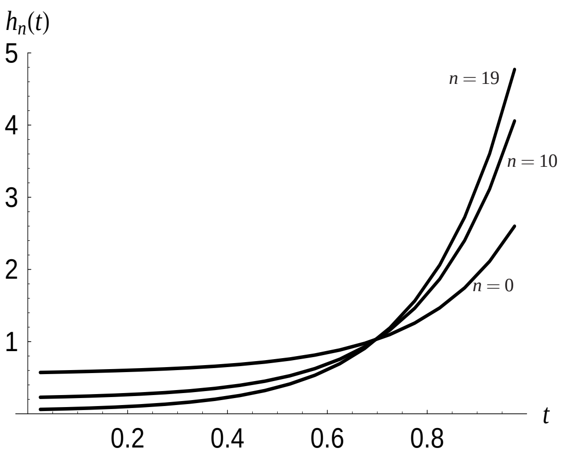

Fig. 3: The density of β_1 for three different initial queue sizes, 0, 10, 19, the buffer size of 20 and $J(0) = 1$.

Now we can proceed to the overflow period in the steady state. First, we need the transition matrix S_{b-1} and its stationary vector, w. Using (20) and (27) we obtain, respectively,

$$S_{19} = \begin{bmatrix} 0.019689 & 0.069005 & 0.911304 \\ 0.129544 & 0.084486 & 0.785969 \\ 0.086155 & 0.109131 & 0.804712 \end{bmatrix},$$

and

$$w = [0.084984, \ 0.103178, \ 0.811837].$$

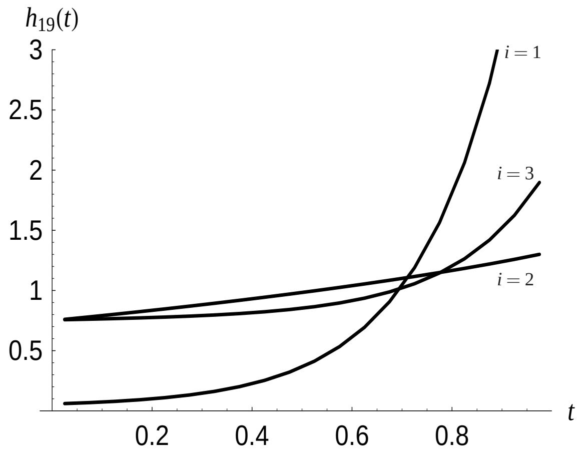

Fig. 4: The density of the length of β_1 for three different initial modulating states, the buffer size of 20 and $X(0) = 19$.

The resulting distribution of the duration of the buffer overflow period in the steady state is depicted in Fig. 5. In addition, the density of the overflow period in the batch Poisson arrivals case is depicted in this figure. Naturally, the same batch sizes, 1, 4, 10, the same average batch size, 8, and the same total arrival rate, 1, as in the BMAP case, were used. Therefore, the difference between these two arrival processes is the autocorrelation, which is equal to zero in the batch Poisson case. As we can see, the autocorrelation itself has a significant influence on the distribution of the overflow period.

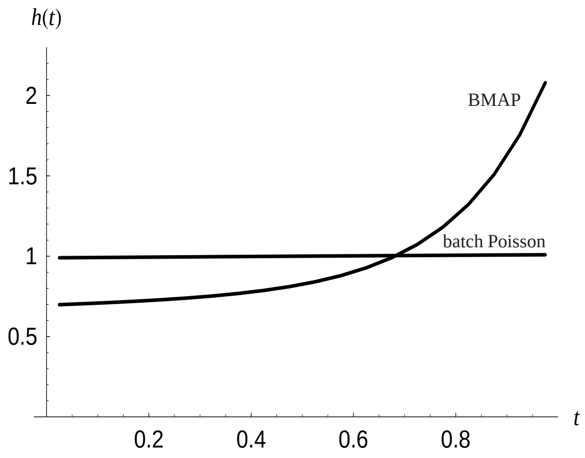

Fig. 5: The density of the length of the overflow period in the steady state for the batch Poisson and BMAP arrivals, $b = 20$.

In the next example, we check the dependence of the duration of the overflow period on the load offered to the system. For this purpose, the considered BMAP is scaled so that its batch and autocorrelation structure is preserved, but the total arrival rate varies from 0.2 to 5. As the service time remains unaltered (and equal to 1), the offered load also varies from 0.1 to 5. The resulting average duration of the overflow period for three different buffer sizes is depicted in Fig. 6. As we can see, the curve is almost flat in every case, which means that the steady-state length of the overflow period depends very little on the offered load. The same effect can be observed for other BMAPs, with different autocorrelation functions and batch size distributions.

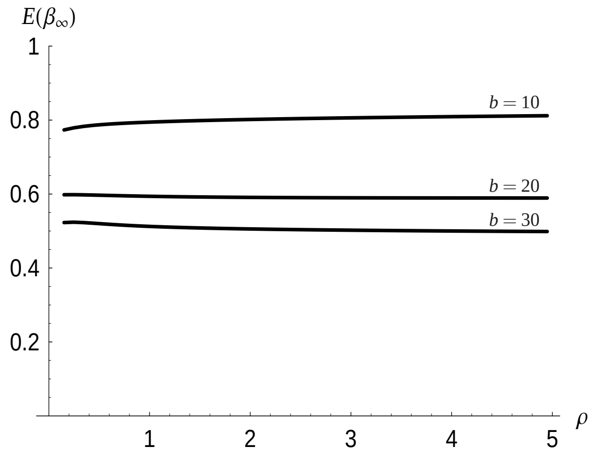

Fig. 6: The average duration of the steady-state overflow period versus the system load for three buffer sizes.

7 Conclusions

We dealt with the buffer overflow period in the queueing system fed with an arrival stream of autocorrelated and batch structure. Formulas for the length of the first, subsequent and stationary overflow periods were proven. Moreover, the formulas for the distribution of the number of consecutive losses during the overflow period were shown. Using these formulas several numerical examples were computed, presenting the dependence of the overflow period on the buffer size, the initial queue size, the system load and the autocorrelation.

Among other things, the examples demonstrated the irregularity of the duration of the overflow period as a function the buffer size, a strong dependence of the overflow period on the autocorrelation structure and a very small dependence of the steady-state overflow period on the system load.

8 Acknowledgement

This material is based upon work supported by the Polish National Science Centre under Grant No. N N516 479240.

References

[1] Asmussen S. Equilibrium properties of the M/G/1 queue, Z. Wah. verw. Geb., 58, pp. 267-281, (1981).
[2] Fakinos D. The expected remaining service time in a single server queue. Operations Research, 30, pp. 1014–1018, (1982).
[3] Gelenbe E., Mang X. and Önvural R. Diffusion based statistical call admission control in ATM. Performance Evaluation, 27&28, pp. 411–436, (1996).
[4] de Boer P. T., Nicola V. F. and van Ommeren J. C. W. The Remaining Service Time Upon Reaching a High Level in M/G/1 Queues. Queueing Systems 39, pp. 55-78, (2001).
[5] Chydzinski A. On the remaining service time upon reaching a target level in M/G/1 queues. Queueing Systems, vol. **47**, issue 1/2, pp. 71–80, (2004).
[6] Chae K. C., Kim K. and Kim N. K. Remarks on the remaining service time upon reaching a target level in the M/G/1 queue. Operations Research Letters, Volume **35**, Issue 3, pp. 308–310, (2007).
[7] Chydzinski A. Buffer Overflow Period in an MAP Queue. Mathematical Problems in Engineering, vol. **2007**, pp. 1–18, (2007).
[8] Chydzinski A. Duration of the buffer overflow period in a batch arrival queue. Performance Evaluation, vol. **63**:4-5, pp. 493–508, (2006).
[9] Paxson V., Floyd S. Wide-area traffic: the failure of Poisson modeling. IEEE/ACM Trans. Netw., Volume **3**, Issue 3, pp. 226–244, (1995).
[10] Leland W. Taqqu M. Willinger W. and Wilson D. On the self-similar nature of ethernet traffic (extended version). IEEE/ACM Transactions on Networking **2**(1): pp. 1–15, (1994).
[11] Garetto M. and Towsley D. An efficient technique to analyze the impact of bursty TCP traffic in wide-area networks. Performance Evaluation 65, pp. 181-202, (2008).
[12] Lucantoni D. M. New results on the single server queue with a batch Markovian arrival process. Commun. Stat., Stochastic Models 7, No.1, pp. 1–46, (1991).
[13] Lucantoni D. M., Choudhury G. L. and Whitt W. The transient BMAP/G/1 queue. Commun. Stat., Stochastic Models 10, No. 1, pp. 145-182, (1994).
[14] Chydzinski A. Time to reach buffer capacity in a BMAP queue, Stochastic Models, vol. **23**, pp. 195–209, (2007).
[15] Chydzinski A. Transient Analysis of the MMPP/G/1/K Queue. Telecommunication Systems, vol. **32**, n. 4, pp. 247–262, (2006).
[16] Chydzinski A. The oscillating queue with finite buffer. Performance Evaluation, vol. **57**(3), pp. 341–355, (2004).
[17] Chydzinski A., Adamczyk B. Transient and stationary losses in a finite-buffer queue with batch arrivals. Mathematical Problems in Engineering, vol. 2012, pp. 1-17, (2012)

Andrzej Chydzinski received his MS (in applied mathematics), PhD (in computer science) and DSc (in computer science) degrees from the Silesian University of Technology, Gliwice, Poland, in 1997, 2002 and 2008, respectively. He is currently a professor in the Institute of Informatics of this university. His academic and professional interests are with computer networking, in particular with performance evaluation of computer networks (via mathematical modelling and simulations), queueing theory and Future Internet design. Prof. Chydzinski authored and co-authored 3 books, about 40 conference papers and about 40 journal articles. He is reviewing articles for several high-quality journals and serving as a Technical Program Committee member for several conferences. Prof. Chydzinski has been a leader of scientific projects founded by Polish state and European Union. He is also an IEEE Senior Member.