Analysis Of A Versatile Multi-Class Delay-Loss System With A Superimposed Markovian Arrival Process

UDO R. KRIEGER Technologiezentrum, Deutsche Telekom AG, Am Kavalieresand 3, D-64295 Darmstadt, Germany *

VALERI NAOUMOV Russian Academy of Sciences, IPIT, ISDN Research Laboratory and Computer Center, Peoples' Friendship University, Moscow, Russia

DIETMAR WAGNER Department of Computer Science, Technische Hochschule Darmstadt, Darmstadt, Germany

ABSTRACT

We analyze the behavior of a multi-class single-server delay-loss system $\sum MAP_i/PH/1/m$ with a superposition of independent MAPs as arrival stream and phase-type distributed service times. Considering the underlying finite Markov chain with its quasi-birth-and-death structure with two boundary sets, we derive a new representation of its steady-state vector by a linear combination of two matrix-geometric terms. Furthermore, we state efficient procedures to calculate the performance characteristics of this delay-loss system.

Key words:
multi-class delay-loss system, finite QBD process, matrix-geometric method

*Further Affiliation: Department of Computer Science, J.W. Goethe-University, Frankfurt, Germany
1 Introduction

In the late sixties Wallace [30] and Evans [7] studied a generalization of a simple birth-and-death process, called quasi-birth-and-death (QBD) process. The corresponding homogeneous continuous-time Markov chain (CTMC) \( \{Z(t), t \geq 0\} \) is two-dimensional vector process \( Z(t) = (X(t), E(t)) \) on a state space \( Z = \mathbb{N}_0 \times \{1, \ldots, l\} \). Here \( X(t) \in \mathbb{N}_0 \) is the process of primary importance with countable state space and \( E(t) \in \{1, \ldots, l\} \) characterizes its random environment with a finite number \( l \) of distinguishable states. \( E(t) \) can be a vector process, too. The resulting generator matrix has a block tridiagonal form with repeating block rows. Neuts [24, Th. 3.1.1, p. 82] and others have shown that the associated unique steady-state distribution vector has a matrix-geometric form if \( Z(t) \) is ergodic (cf. [1, Th. X.5.8, p. 242]). Important single-class single-server models with infinite capacities such as M/PH/1/∞, PH/M/n/∞, M/PH/1/∞, PH/PH/1/∞, MAP/PH/1/∞ and MAP/PH/n/∞ arising from the modeling of telecommunication and computer networks, e.g. buffers in a packet switch, or manufacturing systems are typical QBD processes (cf. [32], [24, Chap. 3], [25]).

At present, however, most studies are devoted to multi-class single-server models with finite capacities. They are motivated by teletraffic theory since different communication services are integrated by advanced high speed packet-switched networks supporting the Asynchronous Transfer Mode (ATM) as basic packet-switching technique (cf. [12], [26]). Applying a continuous-time modeling approach to this context, both a statistical multiplexer and a logical output buffer with FIFO scheduling in a single switching module of an ATM switching fabric can been described by a multi-class single-server delay-loss system \( \sum_i \text{MAP}_i / E / 1 / m \) with finite capacity (cf. [26, p. 159], [4], [32], [28], [14]). The arrival process stems from a superposition of independent general Markovian arrival processes (MAPs) approximating the streams of fixed size packets of each input line in continuous time and is itself a MAP again (cf. [5], [18], [28], [10]). We can use a phase-type (PH) distribution to capture a general service-time distribution since this class is dense within the family of all probability distributions on \((0, \infty)\) (cf. [1, Th. 6.2, p. 76]). Furthermore, it can be adapted to measurement data by appropriate estimation techniques such as the maximum likelihood method. Similar arguments can be applied to buffer modeling in a manufacturing system.

The advanced performance analysis of a finite FIFO buffer is our main objective. We are interested in the effect of traffic loading on the performance characteristics such as the stream-dependent loss probabilities as well as the waiting and sojourn times of different customer streams. First, we show that the steady-state distribution \( \pi \) of the corresponding CTMC with its quasi-birth-and-death structure can be calculated by an improved matrix-geometric technique. It is derived from Naoumov’s general results obtained for generator matrices with block tridiagonal structures (cf. [21, Chap. 3]). We show that \( \pi \) can be represented by a linear
combination of exactly two matrix-geometric terms and develop a tractable numerical method for the performance analysis of a model with a few input streams.

The paper is organized as follows. In Section 2, we describe the features of the versatile delay-loss system $\sum_i \text{MAP}_i/\text{PH}/1/m$ and determine the generator matrix of the underlying Markov chain. In Section 3, we present a new representation of its steady-state distribution by matrix-geometric terms. In Section 4, we derive formulas to calculate the performance measures associated with the different streams of customers. In the Conclusion, we summarize the findings of the paper and provide an outlook on further studies.

2 The multi-class delay-loss system

In this section we derive the generator matrix of the continuous-time Markov chain determining the stochastic behavior of the delay-loss system $\sum_i \text{MAP}/\text{PH}/1/m$. Refining Hajek's [9] well-known matrix-geometric solution technique for finite QBD processes, its steady-state vector is computed in the next section. For this purpose, we assume the reader to be familiar with the theory of homogeneous discrete- and continuous-time Markov chains (abbreviated by DTMC and CTMC) with finite state spaces (cf. [11, Chap. 7, 8], [13]) as well as the theory of nonnegative matrices and M-matrices (cf. [2]). Furthermore, we adopt the notation of Berman and Plemmons [2, Chap. 2, p. 26] regarding vector and matrix orderings.

A matrix $A$ is called ML-matrix if $-A$ is an M-matrix. Thus, the generator matrix $Q$ of a CTMC is an ML-matrix with zero row sums. The negative transpose $A = -Q^t$ with the additional property of zero column sums is called a Q-matrix (cf. [27]). Let $\rho(A)$ denote the spectral radius of $A$ and $\sigma(A)$ its spectrum. Subsequently, let $e$ be the vector with all ones, $\otimes$ the Kronecker product and $\oplus$ the Kronecker sum of two matrices. $I_l$ is the identity matrix of order $l$.

2.1 Mathematical description of the system

In our study we consider a versatile multi-class single-server model of the type $\sum_i \text{MAP}_i/\text{PH}/1/m$ with the following properties:

- The arrival stream is a superposition $\text{MAP}_1 + \ldots + \text{MAP}_x$ of $x$ independent general Markovian arrival processes (MAPs), hence itself a MAP. Each MAP $i$ is described by an irreducible generator matrix $Q^{(i)} = D_0^{(i)} + D_1^{(i)} \in \mathbb{R}^{s_i \times s_i}$ of the CTMC $Y_i$ controlling it. It includes a nonnegative matrix $D_0^{(i)} > 0$ of the rates associated with the transitions of arrivals of customers and a regular ML-matrix $D_1^{(i)}$ of those rates associated with internal phase
shifts without arrivals (cf. [25, p. 269], [18]). This means that \(-D_0^{(i)}\) is a regular M-matrix with positive diagonal elements and \(Q^{(i)} \cdot e = 0\), \(D_0^{(i)} \cdot e = -D_1^{(i)} \cdot e < 0\) hold (cf. [2, Chap. 6], [18]).

Let \(Y(t) = \sum_{i=1}^{s} Y_i, t \geq 0\), record the phase of the resulting superimposed MAP at time \(t \geq 0\), i.e. the state of the corresponding irreducible CTMC \(Y(t)\) with right-continuous trajectories on the state space \(\{1, \ldots, s\}\), \(s = \Pi_{i=1}^{s} s_i\), and with the irreducible generator matrix \(Q = Q^{(1)} \oplus \ldots \oplus Q^{(s)} \in \mathbb{R}^{s \times s}\). Thus, the superimposed process is a MAP with the rate matrix \(D_1 = D_1^{(1)} \oplus \ldots \oplus D_1^{(s)} \in \mathbb{R}^{s \times s}\) and the ML-matrix \(D_0 = D_0^{(1)} \oplus \ldots \oplus D_0^{(s)} \in \mathbb{R}^{s \times s}\) (see [18]).

- The service times of the customers are independent, identically distributed random variables \(\{B_n, n \geq 1\}\) governed by a phase-type distribution \(F\) of order \(k\) with an irreducible representation \((\beta, T)\) (cf. [24, p. 52]). We assume \(\beta^t e = 1\). Furthermore, the service times are assumed to be independent of the arrival process.

- The service facility consists of a single server.

- The capacity \(m\) of the system comprises the server and \(m - 1 > 0\) waiting positions.

- The service discipline is 'delay-loss with FIFO'. If the server is idle, an arriving customer occupies it for a random service period. If it is busy at the arrival instant, it joins the waiting line. If no waiting positions are available, it is lost and has no further impact on the system.

The generic service times \(B\) of the customers are determined by a CTMC \(Y^H(t), t \geq 0\) with state space \(\{0, 1, \ldots, k\}\), transient states \(E = \{1, \ldots, k\}\) and an absorbing state \(0\) (cf. [24], [1, p. 74]). Supposing that \(Y^H\) started in a transient state of \(E\) according to a probability vector \(\beta\), \(F\) evaluates the distribution of the random period until absorption in \(0\). The generator matrix of this absorbing CTMC \(Y^H\) has the form \(G = \begin{pmatrix} 0 & 0 \\ T^0 & T \end{pmatrix}\) with a regular M-matrix \(-T \in \mathbb{R}^{k \times k}\) and a vector \(0 < T^0 = (T_{10}, \ldots, T_{k0})^t \in \mathbb{R}^k\) satisfying \(Te + T^0 = 0\).

Hence, \(F(u) = \mathbb{P}\{B \leq u\} = 1 - \beta^t \cdot \exp(T u) \cdot e, u \geq 0\) holds.

In the following, we assume that both the generator matrix \(Q = D_0^H + D_1^H\) of the arrival process and the generator matrix \(S = T + T^0 \beta^t\) of the service renewal process are irreducible.

From a mathematical point of view, the behavior of the delay-loss model \(\sum_{i} MAP_i/PH/1/m\) is described by a vector-valued stochastic process \(Z(t) = (R(t), H(t), Y(t)), t \geq 0\), on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). \(R(t) = r \in \{0, \ldots, m\}\) denotes the number of customers in the system at time \(t\). \(H(t) = h \in \{1, \ldots, k\}\)
is the phase of the customer just served at \( t \). If the server is idle, i.e. \( R(t) = 0 \), the component \( H(t) = h \) denotes the service phase of the customer arriving next. \( Y(t) = y \in \{1, \ldots, s\} \) is the phase of the CTMC controlling the superimposed MAP. The vector process \( \{Z(t), t \geq 0\} \) is a CTMC on the finite state space \( S = \{ z = (r, h, y) \in \mathbb{N}_0^3 : 0 \leq r \leq m, 1 \leq y \leq s, 1 \leq h \leq k \} \).

### 2.2 Construction of the generator matrix

In the following we denote the vector of the stationary distribution of \( Z(t) \) by \( \pi = (\pi_r)_{r \in S} \), \( \pi_r = \pi(r, h, y) = \lim_{t \to \infty} \mathbb{P}\{R(t) = r, H(t) = h, Y(t) = y\} \). Important steady-state performance characteristics of the system \( \sum_{j=1}^{\infty} \text{MAP/PH/1/m} \) such as the time or call congestion and the actual waiting-time distribution of a customer of a specific stream are defined in terms of \( \pi \).

To construct the generator matrix \( A \) of \( Z(t) \), we first divide the state space into macrostates \([r] = \{(r, h, y) \mid \forall h \in \{1, \ldots, k\}, y \in \{1, \ldots, s\} : (r, h, y) \in S\}, 0 \leq r \leq m\), called levels or \( R \)-lumps (cf. [24, p. 5], [14]). Using the convenient ordering of states, these levels are ordered lexicographically, the last component \( Y(t) = y \) as well. The microstates \([(r, h)] = \{(r, h, y) \mid \forall y \in \{1, \ldots, s\} : (r, h, y) \in S\} \) within a level are called \( H \)-lumps. We order all \( H \)-lumps within each level \([r]\) lexicographically, too. Each level \([r]\), \( r \geq 0 \), comprises \( l = ks \) states; hence, the generator matrix \( A \in \mathbb{R}^{d \times d} \) has the order \( d = [m + 1]ks \). Subsequently, we use \( \pi_r = \pi_{[r]}, 0 \leq r \leq m \), as components of the partitioned steady-state vector \( \pi = (\pi_r)_{r=0, \ldots, m} \) on each \( R \)-lump.

Ordering the levels lexicographically, the generator matrix

\[
A = \begin{pmatrix}
\tilde{A}_1 & A_0 & 0 & \ldots & \ldots & 0 \\
A_2 & A_1 & A_0 & \ldots & \ldots & 0 \\
0 & A_2 & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & 0 & \ldots & \ldots & A_2 & A_1 & A_0 \\
0 & 0 & \ldots & \ldots & 0 & A_2 & A_m
\end{pmatrix}
\]

possesses the structure of a finite quasi-birth-and-death process with two boundary sets \([0]\) and \([m]\).
In the following, we always assume that $A$ is an irreducible ML-matrix with $A\ e = 0$. Hence, all off-diagonal blocks are nonzero, nonnegative matrices, whereas the diagonal blocks are regular ML-matrices (cf. [2]).

3 Computation of the steady-state distribution

In this section we show that the steady-state vector $\pi = (\pi_r)_{r=0,\ldots,m}$ of $Z(t)$ determined by the generator matrix $A$ in (1) can be represented by a linear combination of two matrix-geometric terms. The components $\pi_r$ are elements of a sequence $(x_0, x_1, \ldots, x_m)$ of nonnegative vectors $x_r \geq 0$ satisfying recursions

$$0 = x_0^r \bar{A}_1 + x_1^r A_2$$
$$0 = x_{r-1}^r A_0 + x_r^r A_1 + x_{r+1}^r A_2 \quad 1 \leq r \leq m - 1$$
$$0 = x_{m-1}^r A_0 + x_m^r \bar{A}_m$$

where $A_0, A_1, A_2$ are of the order $l = ks$.

The nonnegative matrices $A_0, A_2$ of the upper and lower block diagonal of $A$ and the regular ML-matrices $A_1$ of the block diagonal of $A$ have the property that

$$A(1) = A_0 + A_1 + A_2 = (T + T^0 \beta^t) \oplus (D_0 + D_1) = S \oplus Q$$

is an irreducible ML-matrix with zero row sums, i.e.

$$0 = A(1)e = A_2 e + A_1 e + A_0 e.$$ 

$A(1)$ is the irreducible generator matrix of the modulating Markovian environment $E(t) = (H(t), Y(t)), t \geq 0$, of the level process $R(t)$. Furthermore,

$$0 = \bar{A}_1 e + A_0 e$$

holds.

Let $p \in \mathbb{R}^l$ be the steady-state vector of the modulating Markovian environment $E(t)$, i.e. the unique nonnegative solution of

$$p^t \cdot A(1) = 0 \quad p^t \cdot e = 1.$$ 

Then it follows $p^t = r_S^t \otimes r_Q^t = [\beta^t(-T)^{-1}] \otimes r_Q^t \cdot [\beta^t(-T)^{-1} e]^{-1}$ with the steady-state vectors $r_S^t \cdot S = 0, r_S^t \cdot e = 1, r_Q^t = \beta^t(-T)^{-1} / \beta^t(-T)^{-1} e, r_Q^t \cdot Q = 0, r_Q^t \cdot e = 1$.

We define the offered load in terms of the fundamental arrival rate $\xi = r_Q^t D_1 e$ of the MAP and the mean service time

$$1/\mu = \beta^t(-T)^{-1} e$$

by

$$\rho = p^t A_0 e / p^t A_2 e = r_Q^t D_1 e \cdot \beta^t(-T)^{-1} e = \xi / \mu.$$
Regarding the solution of the matrix difference equation (3), Naoumov [21, Chap. 3], [22] has developed a comprehensive theory for finite QBD-structured Markov chains with boundary states based on algebraic results for ML-matrices. It has also been extended to Hessenberg-structured generator matrices. It can be applied to a large variety of Markovian models arising from queueing theory. Using a tailored version of his results [21, Chap. 3.1, Theorems 4, 5, p. 70f; Chap. 3.8, p. 106f] (see also [22], [24, p. 82f]), we derive the new representation (24) of the steady-state vector \( \pi \) by a linear combination of two matrix-geometric terms.

**Proposition 1**

Let \( A_0, A_2 \in \mathbb{R}^{l \times l} \) be nonnegative, nonzero matrices and \( A_1 \in \mathbb{R}^{l \times l} \) be a regular ML-matrix such that \( A(1) \) in (5) is an irreducible ML-matrix with the property (6). Let \( p \in \mathbb{R}^l \) be the unique nonnegative solution of (8) and \( \rho \) in (10) be the offered load. Set \( A(z) = A_0 + zA_1 + z^2A_2, z \in \mathcal{C}, \) and assume that there is a \( z \in \mathcal{C} \) such that \( \det(A(z)) \neq 0. \) Then following properties hold:

1. There exist minimal nonnegative solutions \( R, S \in \mathbb{R}^{l \times l} \) of
   
   \[
   0 = A_0 + RA_1 + R^2A_2 \\
   0 = S^2A_0 + SA_1 + A_2. 
   \]

2. It holds:
   
   \[
   \rho(R) < 1 \iff \rho^tA_0e < \rho^tA_2e \iff \rho < 1 \\
   \rho(S) < 1 \iff \rho^tA_0e > \rho^tA_2e \iff \rho > 1 
   \]

3. If \( SA_0 + A_1 + RA_2 \) is regular, then \( \rho^tA_0e \neq \rho^tA_2e \) holds.

**Proof:** see [24, Theorem 3.1.1, p. 82], [21, Chap. 3.3., Theorem 2, p. 84, Theorem 3, p. 86], [22] \( \square \)

To compute the matrix \( R > 0, \) several equivalent procedures were developed in the last decade (cf. [24], [15], [16], [33], [17], [3], [22]). They are formulated in terms of the basic nonnegative matrices \( R, G(R), U(R) \in \mathbb{R}^{l \times l} \). Hajek [9] has shown that these matrices are the minimal nonnegative solutions of (11), its dual equation

\[
0 = A_2 + A_1G(R) + A_0G^2(R)
\]

and

\[
U(R) = (-A_1)^{-1} A_0 [I - U(R)]^{-1} (-A_1)^{-1} A_2.
\]
They satisfy the following relations:

\[
R = A_0[-W^{-1}(R)]
\]

\[
G(R) = -W^{-1}(R)A_2 = [I - U(R)]^{-1} (-A_1)^{-1} A_2
\]

\[
U(R) = (-A_1)^{-1} A_0 G(R)
\]

\[
W(R) = A_1 + A_0 G(R) = A_1 + R A_2 = A_1 [I - U(R)]
\]

Let \( R^D = (-A_1)^{-1} R(-A_1) \). These matrices have the following stochastic meaning (cf. [24], [9], [15], [29, p. 267]):

- \( R^D \) is the expected number of visits of the DTMC \( Z_k \) embedded at jump epochs of the level crossing process into level \([n+1], n \geq 1\), at \( z_1 = (n+1,j) \) before the first return to level \([n]\) provided that it started in level \([n]\) at \( z_0 = (n,i) \).

- \( G_{ij}(R) \) is the first entrance probability into level \([n], n \geq 0\), at \( z_1 = (n,j) \) provided that the CTMC \( Z(t) \) started in level \([n+1]\) at \( z_0 = (n+1,i) \).

In the case \( \rho < 1 \), \( G(R) \) is a stochastic matrix (cf. [24, Th. 3.3.2, p. 1001]).

- \( U_{ij}(R) \) is the taboo probability that the DTMC \( Z_k \) eventually returns to level \([n], n \geq 1\), by visiting state \( z_1 = (n,j) \) at some time \( k \geq 1\), under taboo of the levels \([0], \ldots, [n-1]\), i.e. without visiting any \((l,j'), 0 \leq l \leq n \) in between, provided that the DTMC started in level \([n]\) at \( z_0 = (n,i) \).

In the case \( \rho < 1 \), \( U(R) \) is strictly substochastic with \( U(R) e = -A_1^{-1} A_0 e < e \).

Since \( U(R) \) is strictly substochastic in the case \( \rho < 1 \) and \( -W(R) = -A_1 - (-A_1) U(R) \) is a convergent weak regular splitting, \( W(R) \) is a regular ML-matrix (cf. [2, Th. 6.2.3, N46, p. 138]).

The most efficient procedure to compute a solution \( R \) of the matrix-quadratic equation (11) is provided by Naoumov’s algorithm [23]. It is equivalent to the logarithmic reduction algorithm of Ramaswami and Latouche [16] and derived from (13) and (14). The algorithm uses the following recursions for \( k \geq 0 \) (see also [16]):

\[
N_0 = A_1
\]

\[
A_0 = A_0
\]

\[
M_0 = A_2
\]

\[
B_0^{(k)} = -N_k^{-1} L_k
\]

\[
B_2^{(k)} = -N_k^{-1} M_k
\]

\[
N_{k+1} = N_k \left( I - B_0^{(k)} B_2^{(k)} - B_2^{(k)} B_0^{(k)} \right)
\]

\[
A_{k+1} = A_k B_0^{(k)}
\]
\[ M_{k+1} = M_k B_3^{(k)} \]
\[ W(R) = A_1 + \sum_{k=0}^{\infty} \Lambda_k (-N_k^{-1}) M_k \tag{15} \]

They yield the following procedure to calculate the solution \( R \geq 0 \) of a matrix-quadratic equation (11):

**Improved Logarithmic Reduction Algorithm:**

\[
\begin{align*}
N &:= A_1 \\
\Lambda &:= A_0 \\
M &:= A_2 \\
W &:= 0 \\
R &:= A_0 \\
\text{stop} &:= \text{false}
\end{align*}
\]

**repeat**

\[
\begin{align*}
X &:= -N^{-1} \Lambda \\
Y &:= -N^{-1} M \\
Z &:= \Lambda Y \\
\text{stop} &:= (\|Z\| < \varepsilon) \\
W &:= W + Z \\
N &:= N + Z + M X \\
Z &:= \Lambda X \\
\Lambda &:= Z \\
Z &:= M Y \\
M &:= Z
\end{align*}
\]

**until** \( \text{stop} \)

\[
\begin{align*}
W &:= W + A_1 \\
R &:= R (-W)^{-1}
\end{align*}
\]

Here \( \varepsilon \in (0, 1) \) is a prescribed accuracy level. The computational complexity of the basic loop of this algorithm is given by \( O(19/3 l^3) \) flops.

This improved algorithm is an attractive alternative to numerical solution methods for infinite QBD models or their truncated variants such as block LU schemes or the folding algorithm (cf. \[33\], \[17\]). The latter can be derived from a block LU factorization. It is applied recursively to the permuted generator matrix of a finite QBD model arising from an odd-even ordering of the levels of the QBD (cf. \[33\]). The folding algorithm is also equivalent to a cyclic-reduction technique applied to the block tridiagonal Toeplitz-like generator matrix (cf. \[3, \text{Sec. 3}\]).
Table 1: Time complexity of computational schemes for $R$

<table>
<thead>
<tr>
<th></th>
<th>Naoumov</th>
<th>Latouche &amp; Ramaswami</th>
<th>successive substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O(19/3 , I_N^3)$</td>
<td>$O(25/3 , I_L^3)$</td>
<td>$O(7/3 , I_U^3)$</td>
<td></td>
</tr>
<tr>
<td>$I_w = \text{max. number of iteration steps for } w \in {N, L, U}$, $I_U \gg I_N$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In their recent papers, Bini and Meini [3] and Naoumov [23] proved the equivalence between the reduction algorithm of Latouche and Ramaswami [16] and an odd-even reduction approach applied to an infinite block tridiagonal Toeplitz-like generator matrix. A comparison of the time complexities (see Table 1) illustrates that Naoumov’s algorithm is superior to all others, since the logarithmic reduction approach is more efficient than the traditional solution of the fixed-point equation (11) by successive substitution or Newton’s method. Therefore we apply this most efficient algorithm to our context.

In a way similar to Hajek’s approach [9], it is the key idea of Naoumov’s new representation [22, Corollary, see (61)] of the vector $\pi$ to reduce the solution $\pi$ of (2) to (4) to the subset $[0] \cup [m]$ of boundary states. The condition $A e = 0$ on the row sums is preserved by this operation.

**Proposition 2**

*The matrix*

$$\hat{A} = \begin{pmatrix} \hat{A}_{00} & \hat{A}_{0m} \\ \hat{A}_{m0} & \hat{A}_{mm} \end{pmatrix}$$

*with*

$$\begin{align*}
\hat{A}_{00} &= \bar{A}_1 + A_0 \cdot G(R) \\
\hat{A}_{0m} &= R^m \cdot A_0 \cdot [I - G(R)] \\
\hat{A}_{m0} &= -S^m \cdot [T \otimes I_s + S \cdot A_0] \\
\hat{A}_{mm} &= W(S) + A_0
\end{align*}$$

satisfies $\hat{A} \cdot e = 0$. Hence, there exists a solution $\bar{x}^t = (x_0^t, x_m^t) \neq 0$ of

$$\bar{x}^t \cdot \hat{A} = 0.$$  \hspace{1cm} (21)

**Proof:** Rewriting the boundary equations in [22, Corollary, see (61)], the relations (17) to (20) are derived.

Multiplying (11) by $e$ and using (6), it follows $R^2 A_2 e - RA_0 e = RA_2 e - A_0 e$. Hence, the conservation laws

$$R^k (RA_2 e - A_0 e) = RA_2 e - A_0 e \quad k \geq 1$$

(22)
hold by induction. Reasoning by analogy, we conclude
\[ S^k(SA_0e - A_2e) = SA_0e - A_2e \quad k \geq 1. \] (23)

Then \( \tilde{A}_{00}e + \tilde{A}_{0m}e = \tilde{A}_1e + RA_2e + R^m [A_0e - RA_2e] = \tilde{A}_1e + RA_2e + A_0e - RA_2e = e \otimes (D_0 + D_1)e = 0 \) follows by (22) and the properties of the generator matrix \( Q = D_0 + D_1 \). Applying (23) and (6), we conclude \( \tilde{A}_{m0}e + \tilde{A}_{mm}e = -S^m [(T \otimes I_s)e + SA_0e] + A_1e + SA_0e + A_0e = S^{m-1} [SA_1e + A_2e - A_2e + SA_0e] = S^m [e \otimes (D_0 + D_1)e] = 0 \). These relations induce the zero row sums and imply the equation (21).

The matrices \( A_0 \cdot G(R) = R \cdot A_2, A_2 \cdot G(S) = S \cdot A_0 \) required in (17) and (19) coincide with the series \( \sum_{k=0}^{\infty} A_k (-N_k^{-1}) M_k \) in (15) that is used to compute \( R \) and \( S \) in (11) and (12), respectively. They are computed by the improved algorithm on the fly in (16).

This solution \( \tilde{x} \neq 0 \) can be used to derive a closed form of the steady-state vector \( \pi \) as linear combination of two matrix-geometric terms.

**Theorem 1**

We consider an irreducible finite QBD process \( Z(t) \). Let \( A \in \mathbb{R}^{l \times l} \) be its irreducible generator matrix with the structure (1), hence, let (2) to (8) be fulfilled and compute \( R, S \in \mathbb{R}^{l \times l} \) as minimal nonnegative solutions of (11) and (12), respectively. Furthermore, let \( \det(A(z)) \neq 0 \) for some \( z \in \mathbb{C} \) and \( SA_0 + A_1 + RA_2 \) be regular. Then following results hold:

1. The vector \( \pi = (\pi_r)_{r=0}^{n-1} \in \mathbb{R}^d \) defined by

\[ \pi_r = x_0^r R^r + x_m^r S^{m-r} \quad 0 \leq r \leq m \] (24)

is a solution of

\[ \pi \cdot A = 0 \] (25)

if and only if \( \tilde{x} = (x_0^t, x_m^t) \neq 0 \) is a solution of (21).

2. The normalized variant of \( \pi \) in (24) such that

\[ \left( x_0^t \sum_{r=0}^{m} R^r + x_m^t \sum_{k=0}^{m} S^k \right) \cdot e = 1 \] (26)

is the unique steady-state vector \( \pi \gg 0 \) of \( Z(t) \).

**Proof:** The result follows immediately from [22, Corollary, see (61)] and Propositions 1 and 2. □
According to the assumptions 2. and 3. of Proposition 1,

\[ \begin{align*}
\text{either } & \quad \rho(R) < 1, \quad 1 \in \sigma(S) \quad \iff \quad \rho < 1 \quad (27) \\
\text{or } & \quad \rho(S) < 1, \quad 1 \in \sigma(R) \quad \iff \quad \rho > 1 \quad (28)
\end{align*} \]

holds (cf. [21, Chap. 3.3, Theorem 5], [22, Prop. 6]). Thus, only one of the sums in equation (26) can be simplified, i.e. either for (27) \( \sum_{r=0}^{m} R^r = (I - R)^{-1} (I - R^{m+1}) \) or for (28) \( \sum_{r=0}^{m} S^r = (I - S)^{-1} (I - S^{m+1}) \) holds (see [21, Chap. 3.3, Theorem 5; Chap. 3.6, p. 102f]). Therefore, one of these partial sums can be replaced in the following sections, depending on the validity of the load condition \( \rho = p' A_0 e / p' A_2 e < 1 \).

In the latter case, the block matrices of the first row of \( \hat{A} \) have zero row sums, i.e. \( 0 = \hat{A}_{00} e = \hat{A}_{0m} e \), in the other case (28) the block matrices of the last row have zero row sums, i.e. \( 0 = \hat{A}_{m0} e = \hat{A}_{mm} e \).

4 Performance characteristics of the model

Considering the application of the multi-class single-server delay-loss system \( \Sigma_{i} \text{MAP}_i / \text{PH} / 1 / m \text{ FIFO} \) as a model of a finite buffer, its steady-state performance characteristics are required. They comprise the time- and arrival-stationary distributions of the number of customers in the system and the actual waiting- and sojourn-time distributions of a customer as well as the stream-dependent characteristics. For this purpose, we assume that the CTMC \( \{Z(0), t \geq 0\} \) is in steady state and we use its steady-state distribution \( \pi \).

4.1 Time- and arrival-stationary distributions of the number of customers

Given a feasible state \( z = (r, h, y) \in S \), the time-stationary probabilities \( \pi(r, h, y) = \lim_{t \to \infty} P{R(t) = r, H(t) = h, Y(t) = y} = P{R = r, H = h, Y = y} \) can be computed as normalized solution \( \pi \) of the balance equations (25) by the sketched matrix-geometric method for a finite QBD process. The time-stationary distribution \( P \) of the number of customers in the system is given by the marginal probabilities

\[ P_r = \sum_{(r, h, y) \in [r]} \pi(r, h, y) = \pi_r \cdot e = \left( x_0^t \cdot R^r + x_m^t S^{m-r} \right) \cdot e \]

of each level \([r], r \in \{0, \ldots, m\}\). The time congestion is determined by:

\[ P_m = \pi_m \cdot e = \left( x_0^t \cdot R^m + x_m^t \right) \cdot e \quad (29) \]

To calculate the arrival-stationary distribution \( P^{(0)} \) of the number of customers in the system, i.e. the probabilities \( P_r^{(0)} \) that an arriving customer finds \( r \) customers
in the system at its arrival instant, we employ Melamed's approach [20] (see also [8], [6, p. 71]). By the stochastic intensity principle the arrival-stationary probability $P_r^{(0)}$ coincides with the ratio of the stochastic intensity of those arrivals leaving level $[r]$ and the total intensity of all arrivals at the system including those that find the system occupied and are lost. It is given by

$$
P_r^{(0)} = \frac{\pi_r \cdot A_{rr+1} \cdot e}{\sum_{j=0}^{m} \pi_j \cdot A_{jj+1} \cdot e} \quad (30)$$

where we set $A_{mm+1} = A_m = \mathbb{I} \otimes D_1$ and use $A_{jj+1} = A_0$ for $0 \leq j < m$.

Using the equalities $\xi = \sum_{j=0}^{m} \pi_j \cdot A_{jj+1} \cdot e = (\mathbb{I}^2 \{ Y = y \})_{y=1,\ldots,s} \cdot D_1 \cdot e = r_Q^f \cdot D_1 \cdot e$ for the fundamental arrival rate $\xi$ of the MAP (cf. [18], [25, p. 279]), we see that the arrival-stationary distribution of the number of customers in the system is given by:

$$
P_r^{(0)} = \frac{\pi_r \cdot A_0 \cdot e}{\xi} = \frac{\pi_r \cdot (e \otimes D_1 \cdot e)}{r_Q^f \cdot D_1 \cdot e} \quad 0 \leq r \leq m$$

In particular, the call congestion of an arbitrary arriving customer is determined by:

$$
P_m^{(0)} = \frac{\pi_m \cdot (e \otimes D_1 \cdot e)}{\xi} = \frac{(x^m_0 \cdot R^m + x^m_1 \cdot e)}{r_Q^f \cdot D_1 \cdot e} \quad (31)$$

Applying similar arguments, we can calculate the stream-dependent call-congestion rates $P_m^{(0)}(j)$ of an arbitrary arriving customer belonging to a specified MAP of type $j \in \{1,\ldots,\chi\}$. For this purpose, we define for each MAP $j \in \{1,\ldots,\chi\}$ individual rate matrices $S_j = I \otimes \ldots \otimes D_1^{(j)} \otimes I \otimes \otimes I \in \mathbb{R}^{s \times s}$ according to the construction of the superimposed MAP with the corresponding rate matrix $D_1^{(j)}$ at position $j$. Moreover, let $\xi_j = r_j^f \cdot D_1^{(j)} \cdot e$ be the fundamental arrival rate of MAP $j$. Here $r_j$ is the steady-state vector associated with its generator matrix $Q^{(j)} = D_0^{(j)} + D_1^{(j)}$, i.e. $r_j^f Q^{(j)} = 0, r_j^f e = 1$. Note that $r_Q = \otimes_{j=1}^{\chi} r_j$ and $\xi_j = r_j^f \cdot D_1^{(j)} \cdot e = r_j^f e \otimes \ldots \otimes r_j^f \cdot D_1^{(j)} \cdot e \otimes \ldots \otimes r_j^f e = (\otimes_{j=1}^{\chi} r_j)^f \cdot (I \otimes \ldots I \otimes D_1^{(j)} \otimes I \otimes \otimes I) \cdot e = r_Q^f \cdot S_j \cdot e$.

Then the individual call-congestion rate $P_m^{(0)}(j)$ experienced by an arbitrary arriving customer of stream $j$ is determined by:

$$
P_m^{(0)}(j) = \frac{\xi_j^{-1} \cdot \pi_m \cdot (e \otimes S_j \cdot e)}{r_j^f \cdot D_1^{(j)} \cdot e} \quad (32)$$

The time congestion $P_m$ and the overall and individual call-congestion rates $P_m^{(0)}$, $P_m^{(0)}(j), j \in \{1,\ldots,\chi\}$, can be computed in a straightforward manner by algorithms implementing the formulas (29), (31), and (32).

Furthermore, it is evident that the superimposed streams of carried as well as lost customers of all types are MAPs and that all streams of carried or lost customers of a specific type are MAPs, too.
4.2 The actual waiting-time distributions

The actual waiting-time distributions of the customers are important performance measures of the multi-class delay-loss system \( \sum_i MAP_i/PH/1/m \). In the sequel, we assume that the system is in steady state and note that the actual waiting times of the customers are identically distributed random variables (cf. [1, Chap. III.10]).

Let \( W(0) \) denote the waiting time observed by an arbitrary arriving customer, \( R(0) \) be the number of customers and \( H(0) \) the phase vector of the service process seen at his arrival instant. \( W(x) \) denotes the conditional probability that an arriving customer has to wait at most \( x \) time units until he is served, provided that he can enter the system, i.e. \( W(x) = \text{Prob}\{W(0) \leq x \mid 0 \leq R(0) < m\} \). Let \( e_h(k) \in \mathbb{R}^k \) denote the \( h \)-th unit vector.

We conclude by convenient arguments (cf. (24, Sec. 3.9, p. 133ff)) that this actual waiting time \( W(0) \) is governed by a phase-type distribution \( ((\alpha_0, \alpha^t), \hat{W}) \) of the form

\[
W(x) = 1 - \alpha^t \cdot \exp(\hat{W} \cdot x) \cdot e
\]

with a regular ML-matrix

\[
\hat{W} = I_{m-1} \otimes T + \sum_{j=2}^{m-1} [e_j(m-1) e_j^{t-1}(m-1)] \otimes [T^\beta] \tag{33}
\]

of order \( \omega = (m - 1)k \) with \( m - 1 \) blocks \( T \in \mathbb{R}^{k \times k} \) along its diagonal and an initial probability vector for \( r \in \{1, \ldots, m-1\} \).

For \( 1 \leq r \leq m - 1 \) these mixing probabilities

\[
\alpha_{(r,h)} = \mathbb{P}\{R(0) = r, H(0) = h \mid 0 \leq R(0) < m\} = \frac{P_{(r,h)}(0)}{1 - P_m^{(0)}}
\]

for \( r \in \{1, \ldots, m-1\} \).
are determined by the arrival-stationary probabilities $P \{ R(0) = r, H(0) = h \} = P_{(r,h)}^{(0)}$ that an arbitrary tagged customer finds $R(0) = r \in \{1, \ldots, m\}$ customers in the system and the service process in state $H(0) = h \in \{1, \ldots, k\}$ at his arrival instant provided that he can enter the system, i.e. $R(0) < m$. With the probability $a_0 = P \{ R(0) = 0 \mid 0 \leq R(0) < m \} = P_{0}^{(0)}/[1 - P_{m}^{(0)}]$ the server is idle at the arrival of a tagged customer, hence the actual waiting time is zero. This event is described by a start in the absorbing state zero of the PH-distribution.

$P_{Q}^{(0)} = 1 - P_{0}^{(0)} - P_{m}^{(0)} = \xi^{-1} \sum_{r=1}^{m-1} \pi_{r}^{t} \cdot (e \otimes D_{1} e)$ is the queueing probability, i.e. the probability that an arbitrary arriving customer can enter the system and has to wait after his arrival.

If the service times are exponentially distributed, the actual waiting-time distribution is a mixture of Erlang distributions with 1 to $m-1$ phases (cf. [19]).

Regarding the actual waiting-time distribution $W_{j}$ experienced by an arbitrary arriving customer of a specified MAP of type $j \in \{1, \ldots, \chi\}$, similar arguments can be applied. $W_{j}$ is again a PH-distribution with representation $([\alpha(j)]_{0}, [\alpha(j)]^{t}, \tilde{W})$ and

$$[\alpha(j)]_{0} = \frac{P_{0}^{(0)}(j)}{1 - P_{m}^{(0)}(j)}$$

$$[\alpha(j)]_{r}^{t} = ([\alpha(j)]_{(r,1)}, \ldots, [\alpha(j)]_{(r,k)}) = \frac{\pi_{r}^{t} \cdot [I_{k} \otimes (S_{j} e)]}{\xi_{j} \cdot [1 - P_{m}^{(0)}(j)]}$$

for $r \in \{1, \ldots, m-1\}$. The mixing probability

$$[\alpha(j)]_{(r,i)} = \frac{\pi_{r}^{t} \cdot (e_{i}(k) \otimes I_{s}) \cdot S_{j} e}{\xi_{j} \cdot (1 - P_{m}^{(0)}(j))} = \frac{\pi_{r}^{t} \cdot [e_{i}(k) \otimes (S_{j} e)]}{\xi_{j} - \pi_{m}^{t} \cdot (e \otimes S_{j} e)}$$

is the conditional probability that an arriving customer of type $j$ finds $r$ customers in the system and the service process in state $i$ provided that he can enter the system. The queueing probability of type $j$ is given by:

$$P_{Q}^{(0)}(j) = 1 - P_{0}^{(0)}(j) - P_{m}^{(0)}(j) = \xi_{j}^{-1} \sum_{r=1}^{m-1} \pi_{r}^{t} \cdot (e \otimes S_{j} e)$$

The actual waiting-time distributions $W$ and $W_{j}$, $j \in \{1, \ldots, \chi\}$, can be computed by any algorithm that is suitable to evaluate PH-distributions. It has to implement the calculation of the matrix $\hat{W}$ according to (33) and the mixing probability vectors $\alpha$ and $\alpha(j)$ according to (34), (35) and (36), (37). Their $k$-th moments $M^{(k)}$ and $M_{j}^{(k)}$ can be calculated by

$$M^{(k)} = (-1)^{k} k! \alpha^{t} \hat{W}^{-k} e \quad k \geq 1$$

$$M_{j}^{(k)} = (-1)^{k} k! [\alpha(j)]^{t} \hat{W}^{-k} e \quad k \geq 1$$
in a straightforward manner, too (cf. [24, p. 461]. Regarding the first moments, we can exploit the representation (33) and obtain, after some algebraic manipulation, the following simplified formulas:

\[
M^{(1)} = \sum_{r=1}^{m-1} \alpha_r \cdot (-T^{-1}) \cdot e + \left( \sum_{i=2}^{m-1} (i-1) \alpha_i \cdot e \right) \cdot \frac{1}{\mu}
\]

\[
= \frac{1}{1 - P_m^{(0)}} \cdot \left( \frac{1}{\xi} \cdot \left[ \sum_{r=1}^{m-1} \pi_r \cdot (-T^{-1}) \otimes (D_r e) \right] + \frac{1}{\mu} \cdot \left[ \sum_{i=2}^{m-1} (i-1) P_i^{(0)} \right] \right)
\]

\[
M^{(1)}_j = \sum_{r=1}^{m-1} \alpha_r^{(j)} \cdot (-T^{-1}) \cdot e + \left( \sum_{i=2}^{m-1} (i-1) \alpha_i^{(j)} \cdot e \right) \cdot \frac{1}{\mu}
\]

\[
= \frac{1}{1 - P_m^{(0)}(j)} \cdot \left( \frac{1}{\xi_j} \cdot \left[ \sum_{r=1}^{m-1} \pi_r \cdot (-T^{-1}) \otimes (S_r e) \right] + \frac{1}{\mu} \cdot \left[ \sum_{i=2}^{m-1} (i-1) P_i^{(0)}(j) \right] \right)
\]

Here the first moment of the service time \(1/\mu\) in (9) and \(\alpha_i^{(j)}\) in (34) as well as \(\alpha_i^{(j)}\) in (36) are used.

4.3 The actual sojourn-time distributions

The actual sojourn-time distributions are important performance measures of the \(\Sigma^\infty_i MAP_i/PH/1/m\) model. We suppose that the system is in steady state and realize that the actual sojourn times of the customers are identically distributed (cf. [1, Chap. III.10], [31, 4.1.3 Ex., p. 173f]). Let \(D^{(0)}\) and \(W^{(0)}\) denote the corresponding random variables of the generic sojourn and waiting times observed by an arriving customer, \(D_j^{(0)}\) and \(W_j^{(0)}\) those experienced by an arriving customer of type \(j \in \{1, \ldots, \chi\}\) and \(B\) be the random variable of his generic service time. \(R^{(0)}\) and \(R_j^{(0)}\) denote the numbers of customers seen in the system at the arrival instant of an arbitrary customer and a customer of type \(j\), respectively. In either case, the sojourn times \(D^{(0)} \triangleq W^{(0)} + B\) and \(D_j^{(0)} \triangleq W_j^{(0)} + B\) are determined by the sum of the waiting and service times provided that the tagged customer can enter the system, i.e. \(R^{(0)} < m\) and \(R_j^{(0)} < m\), respectively. Since these times are independent PH-distributed random variables, \(D^{(0)}\) and \(D_j^{(0)}\) are governed by PH distributions again (cf. [24]). The corresponding distribution functions \(D(x) = \mathbb{P}\{D^{(0)} \leq x \mid R^{(0)} < m\}, D_j(x) = \mathbb{P}\{D_j^{(0)} \leq x \mid R_j^{(0)} < m\}\), \(x \geq 0, j \in \{1, \ldots, \chi\}\), have the representations \((\gamma, L_m)\) and \((\gamma_j, L_j)\) with the probability vectors \(\gamma^t = ([\alpha_0]_0 \beta^t, [\alpha_r]_r \beta^t)\) and the regular ML-matrix

\[
L_m = \begin{pmatrix}
T & 0 & \ldots & \ldots & 0 \\
T^0 \beta & 0 & \ldots & \ldots & \ldots \\
0 & \ldots & \ldots & \ldots & \ldots \\
\vdots & \ddots & \ldots & \ldots & \ldots \\
0 & \ldots & \ldots & T & 0 \\
0 & \ldots & 0 & T^0 \beta & T
\end{pmatrix} = \begin{pmatrix}
T & 0 \\
-W \beta^t & \bar{W}
\end{pmatrix}
\]
of order \( m k \). Then the mean sojourn times are determined by:

\[
V^{(1)} = \mathbb{E}(W^{(0)}) + \frac{1}{\mu} = \sum_{i=1}^{m-1} \alpha_i \cdot (-T^{-1}) \cdot e + \left( \alpha_0 + \sum_{i=1}^{m-1} \alpha_i \cdot e \right) \cdot \frac{1}{\mu}
\]

\[
= \frac{1}{1 - P_m^{(0)}} \cdot \left( \sum_{i=1}^{m-1} \pi_i \cdot (-T^{-1}) \cdot e \right) + \frac{1}{\mu} \left[ P_0^{(0)} + \sum_{i=1}^{m-1} i \cdot P_i^{(0)} \right]
\]

\[V_j^{(1)} = \mathbb{E}(D_j^{(0)}) = \mathbb{E}(W_j^{(0)}) + \frac{1}{\mu} = \sum_{i=1}^{m-1} \alpha_{i(j)} \cdot (-T^{-1}) \cdot e + \left( [\alpha_{i(j)}] + \sum_{i=1}^{m-1} \alpha_{i(j)} \cdot e \right) \cdot \frac{1}{\mu}
\]

\[
= \frac{1}{1 - P_m^{(0)}(j)} \cdot \left( \sum_{i=1}^{m-1} \pi_i \cdot (-T^{-1}) \cdot e \right) + \frac{1}{\mu} \left[ P_0^{(0)}(j) + \sum_{i=1}^{m-1} i \cdot P_i^{(0)}(j) \right]
\]

Here \( \alpha_i \) in (34) as well as \( [\alpha_{i(j)}] \) in (36), \( \tilde{W} \) in (33) and the first moment of the service time \( 1/\mu \) in (9) are used again.

### 4.4 The mean numbers of customers in the system and the queue

Further relevant performance characteristics are the mean numbers of customers in the system, \( L^{(1)} = \mathbb{E}(R) \), and in the queue, \( Q^{(1)} = \mathbb{E}(\max(0, R - 1)) \). The corresponding quantities of customers of type \( j \in \{1, \ldots, \chi\} \) are denoted by \( L_j^{(1)} \) and \( Q_j^{(1)} \), respectively. Let \( \varrho = \mathbb{E}(\min(1, R)) = (1 - \mathbb{P}\{R^{(0)} \leq m\}) \) \( \xi/\mu \) denote the utilization of the server and \( \varrho_j = \mathbb{E}(\min(1, R_j)) \) the corresponding term for type \( j \). Then

\[
\varrho = \sum_{j=1}^{\chi} \varrho_j = \left( 1 - P_m^{(0)} \right) \frac{\xi}{\mu}, \quad \varrho_j = \left[ 1 - P_m^{(0)}(j) \right] \frac{\xi_j}{\mu}
\]

and the well-known relation

\[
P_m^{(0)} = \sum_{j=1}^{\chi} \frac{\xi_j}{\xi} P_m^{(0)}(j)
\]

follow by the rate conservation law. By Little's Law we conclude

\[
L^{(1)} = Q^{(1)} + \rho = \sum_{r=0}^{m} r \cdot \mathbb{P}\{R = r\} = \sum_{r=1}^{m} r \cdot \pi_r \cdot e = \left( 1 - P_m^{(0)} \right) \cdot \xi \cdot V^{(1)}
\]
\[ L_j^{(1)} = Q_j^{(1)} + \rho_j = [1 - P_m^{(0)}(j)] \cdot \xi_j \cdot V_j^{(1)} \]

\[ Q^{(1)} = \sum_{r=1}^{m} (r - 1) \cdot \Pr \{ R = r \} = \sum_{r=2}^{m} (r - 1) \pi_r' \cdot e = (1 - P_m^{(0)}) \cdot \xi \cdot M^{(1)} \]

\[ Q_j^{(1)} = \left[ 1 - P_m^{(0)}(j) \right] \cdot \xi_j \cdot M_j^{(1)} \]

(cf. [31, p. 100]). These relations can be used to evaluate all first-order steady-state performance characteristics of the model. Quantities of higher order can be derived from the moment formulas for PH-distributions (cf. [24, p. 46]).

5 Conclusion

The multi-class single-server delay-loss system $\sum_i^3 \text{MAP}_i/\text{PH}/1/m$ with a general Markovian arrival process resulting from a superposition of independent MAPs and phase-type distributed service times is a versatile generic model for a finite FIFO buffer in a packet-switched network or a flexible manufacturing system. In our contribution we have presented a novel computational method to analyze its performance characteristics. Considering the steady-state probability vector of the number of customers in the system, a new representation in terms of a linear combination of two matrix-geometric terms is stated as the main result. It is derived from Naoumov’s more general results [21], [22] obtained for finite CTMCs with QBD-structured generator matrices. The relevant performance characteristics of the system are expressed in terms of this steady-state vector. The results can be applied to study the stream-dependent performance characteristics of a logical buffer in a switching module of a packet switch or to analyze measurement data arising from a manufacturing device.

References


