ANALYSIS OF PARETO/M/s/c QUEUES USING UNIFORMIZATION

FÁTIMA FERREIRA\textsuperscript{a} and ANTÓNIO PACHECO\textsuperscript{b}
\textsuperscript{a}Department of Mathematics and CEMAT, UTAD,
Quinta dos Prados, Apartado 1013, 5001-911 Vila Real, Portugal
E-mail: mmferrei@utad.pt
\textsuperscript{b}Department of Mathematics and CEMAT, Instituto Superior Técnico,
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
E-mail: apacheco@math.ist.utl.pt

\section*{Abstract} Uniformization is a powerful numerical technique for the analysis of continuous time Markov chains that possesses a probabilistic interpretation. In this paper, we propose the combination of the embedded Markov chain approach for queues with uniformization to analyze GI/M/s/c queues. The main steps of the procedure are: (1) To compute the mixed-Poisson probabilities associated to the number of arrival epochs of the uniformizing Poisson process in-between two consecutive customer arrivals to the queueing system; and (2) To compute the conditional embedded uniformized transition probabilities of the number of customers in the queueing system in-between two consecutive customer arrivals to the system. To show the power of the approach, we analyze Pareto/M/s/c queues using a stable recursion for the mixed-Poisson probabilities with a computation time that is linear on the number of computed coefficients. As a result, fast numerical computation of loss probabilities in moderate size queues with Pareto renewal arrival process and exponential servers is achievable. The results for Pareto/M/s/c queues are compared with those obtained for queues with other renewal arrival processes, including Gamma/M/s/c, M/M/s/c and D/M/s/c. The results show that much higher loss probabilities and mean waiting times in queue may be obtained for Pareto renewal arrival processes than for the other renewal arrival processes considered, especially for small customer arrival rates.

\section*{Keywords} GI/M/s/c queues, Loss Probability, Markov chains, Mixed-Poisson probabilities, Pareto distribution, Stochastically monotone matrices, Uniformization.

\section{1. INTRODUCTION} We consider general multiserver GI/M/s/c queueing systems. Customers arrive according to a renewal process with interarrival times with general distribution $A$ with mean $\lambda^{-1}$. The customers are served by one of $s$ servers (working in parallel) in a first come, first served discipline and the service times of customers are independent and identically distributed (i.i.d.) with exponential distribution with rate $\mu$. If the queue capacity $c$ is finite, customers that when arriving find $c^* = s + c$ customers in the system are lost, whereas if $c = \infty$ no customers are lost.

For the study of GI/G/1 systems, the use of the embedded Markov chain approach, introduced by Kendall [1], is the classic approach. In this paper we combine this approach with uniformization [2, 3] to study GI/M/s/c queues. Uniformization is a powerful numerical technique for computing the transition probability matrices of continuous time Markov chains (CTMCs) with bounded rates out of states. This technique uniformizes the rates at which transitions out from each state of the CTMC take place by introducing self-transitions.

\textbf{Outline of paper:} In Section 2 we show how to combine the embedded Markov chain approach with uniformization to obtain the transition probability matrix of the discrete time Markov chain (DTMC) resulting from embedding the number of customers in the GI/M/s/c system immediately before arrival epochs and the corresponding stationary distribution. In Section 3 formulas for the steady-state loss probability, the mean waiting time in queue of entering customers and the mean number of customers in the system seen
by arriving costumers are given. In Section 4 a numerical illustration of the behavior of
the previous performance measures is presented for Pareto, gamma, exponential, determin-
istic and uniform inter-arrival distributions. We highlight the differences between
the Pareto/M/s/c system and the other systems. Finally, in Section 5 some conclusions are

drawn.

2. COMBINING EMBEDDED MARKOV CHAINS WITH
UNIFORMIZATION

Let \( X(t) \) denote the number of customers in a GI/M/s/c system (the system) at time
\( t, t \geq 0 \). The process \( X = \{X(t), t \geq 0\} \) is a semi-Markov process, which is Markovian
only when the renewal arrival processes is Poisson, i.e., the interarrival times have expo-
nential distribution. Due to space constraints the study in the paper is limited to the
characterization of the system from the perspective of the arriving customers. Thus, in-
stead of \( X \) we use the associated embedded DTMC \( \hat{X} = \{\hat{X}_k, k \geq 0\} \) immediately before
arrival epochs. The complete study of \( X \), which will be reported elsewhere, is based on
the Markov renewal sequence \( \{(\hat{X}_k, S_k), k \geq 0\} \), where \( S_k \) is the \( k \)th renewal-arrival epoch
and, consequently, \( \hat{X}_k = X(S_k^-) \) is the number of customers in the system immediately
before the \( k \)-th arrival epoch.

In-between two consecutive customer arrivals the process \( X \) behaves like a pure death
homogeneous Markov process with state space \( S^* = \{0, 1, 2, \ldots, c^*\} \) and death rates \( \mu_i =
\mu \min(s, i), i = 1, 2, \ldots, c^* \). Thus,

\[
\{X(t), S_k \leq t < S_{k+1} | X(S_k) = i\} = \{Y(t), 0 \leq t < S_{k+1} - S_k | Y(0) = i\}
\]

for \( i \in S^* \), where \( Y = \{Y(t), t \geq 0\} \) is a pure death process with state space \( S^* \)
and death rates \( \{\mu_i\} \), independent of the renewal arrival process, and \( = \) denotes equality in
distribution.

Since the interarrival times, \( T_k = S_{k+1} - S_k \), for \( k \geq 0 \) where \( S_0 = 0 \), are i.i.d. random
variables with common distribution function \( A \) and are independent of \( \hat{X}_k \), and in view of
(1), the embedded DTMC \( \hat{X} \) has one-step transition probabilities

\[
\hat{p}_{ij} = P(\hat{X}_{k+1} = j | \hat{X}_k = i) = \int_0^{\infty} P(Y(t) = j | Y(0) = i + [1 - \delta_{i,c^*}]) A(dt)
\]

for \( i, j \in S^* \), where \( \delta \) is the Kronecker delta function, i.e., \( \delta_{ab} \) is one if \( a = b \) and is zero
otherwise. This follows since if a customer arrives to the system finds \( i \) customers
in the system, then he joins the system if \( i < c^* \) and is lost (or blocked) if \( i = c^* \). Then,
deleting \( \delta_{c^*}(i) = i + 1 - \delta_{i,c^*} \) for \( i \in S^* \), we have, for \( i, j \in S^* \),

\[
\hat{p}_{ij} = \int_0^{\infty} P(Y(t) = j | Y(0) = i + [1 - \delta_{i,c^*}]) A(dt) = P(Y(T_1) = j | Y(0) = \delta_{c^*}(i))
\]

Thus, the transition probability matrix \( \hat{P} \) of \( \hat{X} \) is a function of the transition probabili-
ties of the pure death process \( Y \) in a random amount of time with the interarrival time
distribution. As the rates out of states of the CTMC \( Y \) are bounded by the constant \( s \mu \)
the service rate when all \( s \) servers are busy, we may use uniformization to compute the
one-step transition probabilities (3) of \( \hat{X} \).

Let \( Q \) denote the infinitesimal generator matrix of \( Y \), so that \( q_{i,i-1} = -q_{ii} = \mu_i = \mu \min(s, i) \), \( i = 1, 2, \ldots, c^* \), and all other entries of \( Q \) are null. For the computation of \( \hat{P} \) we
use the embedded uniformized DTMC with uniformization rate \( s \mu \) associated to \( Y \), which
has one-step transition probability matrix \( \hat{P} = I + Q/(s \mu) \). Thus, \( \hat{p}_{i,i-1} = \min(s, i)/s =
1 - \hat{p}_{ii}, \ i = 1, 2, \ldots, c^* \), \( \hat{p}_{00} = 1 \) and \( \hat{p}_{ij} = 0 \) otherwise. Hence, \( \hat{P} \) is a stochastically
monotone matrix\(^1\) and, moreover, it is a lower triangular matrix, viz. \(\hat{P}\) is associated to an almost surely nonincreasing DTMC.

Starting from (3) and conditioning on the number of renewals until time \(T_1\) of the uniformizing Poisson process with rate \(s\mu\) (independent of \(Y\)), a straightforward computation gives

\[
\hat{P} = \sum_{n=0}^{+\infty} \alpha_A(n, s\mu) \cdot \hat{\delta}_{c^*}(\hat{P}^n) = \hat{\delta}_{c^*} \left( \sum_{n=0}^{+\infty} \alpha_A(n, s\mu) \cdot \hat{P}^n \right)
\]  

(4)

where \([\hat{\delta}_{c^*}(B)]_{ij} = [B]_{\hat{\delta}_{c^*}(i),j}\), for \(B = [b_{ij}]_{i,j \in S^*}\), and

\[
\alpha_A(n, s\mu) = \int_0^{+\infty} e^{-t}(s\mu)^n n! A(dt).
\]  

(5)

The probability \(\alpha_A(n, s\mu)\) is the \(n\)-th mixed-Poisson probability with structural distribution function \(A\) and rate \(s\mu\). Namely, \(\alpha_A(n, s\mu)\) denotes the probability that exactly \(n\) renewals take place in the uniformizing Poisson process with rate \(s\mu\) in-between two consecutive customer arrivals to the queueing system. To ease the notation, we will write from now onwards \(\alpha_n\) instead of \(\alpha_A(n, s\mu)\).

In the examples considered in the paper, we will use the implementations of the mixed-Poisson probabilities in Kwiatkowska, Norman and Pacheco [4]. Aside from the calculation of the mixed Poisson probabilities, the computation of \(\hat{P}\) from (4) involves an infinite number of powers of the matrix \(\hat{P}\). Thus, in general, the sum in (4) has to be truncated and the result leads only to an approximation of \(\hat{P}\). A special case where an exact result is obtained is in the single server case, \(s = 1\), as next stated.

**Theorem 1** If \(s = 1\), then the embedded DTMC \(\hat{X}\) has one-step transition probabilities

\[
\hat{p}_{ij} = 1(\hat{\delta}_{c+1}(i) \geq j > 0) \alpha_{\hat{\delta}_{c+1}(i) - j} + \delta_{j0} \sum_{n=\hat{\delta}_{c+1}(i)}^{\infty} \alpha_n
\]  

(6)

for \(i, j = 0, 1, \ldots, c+1\), where \(1(B)\) is the indicator function of \(B\), i.e., \(1(B)\) is one if \(B\) is true and is zero otherwise. For the \(M/M/1/c\) system, \(\alpha_n = (1 - \beta)^n\) with \(\beta = \mu / (\lambda + \mu)\), and

\[
\hat{p}_{ij} = \begin{cases} 
(1 - \beta)\hat{\delta}_{c+1}(i) - j & \hat{\delta}_{c+1}(i) \geq j > 0 \\
\hat{\delta}_{c+1}(i) & j = 0 \\
0 & \text{otherwise}.
\end{cases}
\]  

(7)

Note that for finite \(c\) the computation of the embedded transition probabilities \(\hat{p}_{ij}\) from (6) requires the computation of mixed Poisson coefficients \(\alpha_n\) only for \(n = 0, 1, \ldots, c\), since \(\sum_{n=k+1}^{\infty} \alpha_n = 1 - \sum_{n=0}^{k} \alpha_n\), for all \(k \in \mathbb{N}_0\). However, for multiple servers \((s > 1)\) the transition probability matrix \(\hat{P}\) is not associated to a deterministic monotone decreasing DTMC as in the case \(s = 1\). Despite that, the transition probability matrix \(\hat{P}\) still possesses good monotonicity properties for multiple servers, which are used to derive Theorem 2 below. For presenting the result, it is useful to let

\[
\hat{R}^{(N)} = \hat{\delta}_{c^*} \left( \sum_{n=0}^{N} \alpha_n \hat{P}^n + \sum_{n=N+1}^{+\infty} \alpha_n \hat{P}^n \right), \quad \tilde{R}^{(N)} = \hat{\delta}_{c^*} \left( \sum_{n=0}^{N} \alpha_n \tilde{P}^n + \sum_{n=N+1}^{+\infty} \alpha_n \tilde{P}^n \right)
\]

for \(N \in \mathbb{N}_0\), where \(Z = (z_{ij}) = (\delta_{j0})\). Moreover we let \(Y^{(N)}\) and \(\tilde{Y}^{(N)}\) denote DTMCs on \(S^*\) with respective transition probability matrices \(\hat{R}^{(N)}\) and \(\tilde{R}^{(N)}\).

\(^1\)We recall that given two stochastic matrices with common indices, \(A\) and \(B\), the matrix \(A\) is said to be stochastically smaller than \(B\) in the Kalmykov order sense, \(A \preceq_K B\), if and only if \(\sum_{j \geq k} a_{ij} \leq \sum_{j \geq k} b_{mj}\), for any \(i \leq m\) and any \(k\). Moreover, a stochastic matrix \(A\) is said to be stochastically monotone if \(A \preceq_K A\).
Theorem 2 The matrices $\hat{R}^{(N)}$ and $\tilde{R}^{(N)}$ are stochastic and

$$
\hat{R}^{(N)} \leq K \tilde{P} \leq K \tilde{R}^{(N)}
$$

(8)

for any $N \in \mathbb{N}_0$, and $\hat{R}^{(N)}$ and $\tilde{R}^{(N)}$ converge to $\tilde{P}$ as $N$ tends to infinity. Moreover, $\tilde{X}$, $\tilde{Y}^{(N)}$, $N \geq 1$, and $\hat{Y}^{(N)}$, $N \geq 2$, are irreducible and aperiodic DTMCs.

The DTMCs $\tilde{X}$ and $\tilde{Y}^{(N)}$, $N \in \mathbb{N}$, are ergodic if $c < \infty$ or $c = \infty$ and $\rho = \lambda/(s\mu) < 1$. Conversely, the DTMCs $\tilde{Y}^{(N)}$, $N \geq 2$, are ergodic if $c < \infty$ or for sufficiently large $N$ if $c = \infty$ and $\rho < 1$. Moreover, if the DTMC $\tilde{Y}^{(N)}$ is ergodic and we let $\pi^{(N)} (\tilde{P}, \tilde{Y}^{(N)})$ denote the unique stationary probability row vector associated with $\tilde{Y}^{(N)} (\tilde{P}, \tilde{Y}^{(N)})$, then

$$
\pi^{(N)} \leq_{st} \tilde{P} \leq_{st} \pi^{(N)}.
$$

(9)

Proof. The mixed Poisson probabilities $\{\alpha_n, n \in \mathbb{N}_0\}$ constitute a probability function and are all strictly positive. Thus, $\hat{R}^{(N)}$ and $\tilde{R}^{(N)}$ are stochastic matrices since they are convex linear combinations of the stochastic matrices $\tilde{P}^n, n \in \mathbb{N}_0$, and $Z$. The convergence to $\tilde{P}$ of $\hat{R}^{(N)}$ and $\tilde{R}^{(N)}$, as $N$ tends to infinity, then follows from the fact that $\sum_{n=N+1}^{\infty} \alpha_n$ tends to zero as $N$ tends to infinity.

In addition, (8) follows from the following three properties: (i) The powers of the transition probability matrix $\tilde{P}$ are stochastically monotone nonincreasing in the Kalmykov order sense, i.e., $\tilde{P}^{n_1} \leq K \tilde{P}^{n_2}$ for any $0 \leq n_1 \leq n_2$; (ii) For any stochastic matrix $B$ with indices on $S^*$, $Z \leq K B$; thus, in particular, $Z \leq \tilde{P}^n$ for all $n \in \mathbb{N}_0$; and (iii) For any stochastic matrices $B$ and $C$ with indices on $S^*$, if $B \leq K C$, then $\delta_{\sigma^*}(B) \leq K \delta_{\sigma^*}(C)$.

The fact that $\tilde{X}$, $\tilde{Y}^{(N)}$, $N \geq 1$, and $\hat{Y}^{(N)}$, $N \geq 2$, are irreducible and aperiodic follows directly from the general theory of DTMCs by inspection of the nonzero entries of their transition probability matrices, for which the strict positivity of the mixed Poisson coefficients is used. As a consequence, if in addition $c < \infty$ these DTMCs are ergodic. It is well known that $\tilde{X}$ is also ergodic if $c = \infty$ and $\rho < 1$ [5, 6]. In view of (8) and Theorem 3.12 of Cabral Morais [7], this implies that the DTMCs associated to $\hat{R}^{(N)}$, $N \in \mathbb{N}$, are also ergodic if $c = \infty$ and $\rho < 1$.

Suppose that $c = \infty$ and $\rho < 1$, and fix $\epsilon > 0$ such that $\lambda(1+\epsilon) < s\mu$. As from the properties of the mixed Poisson coefficients we have $\sum_{n=0}^{\infty} \alpha_n = s\mu/\lambda$, there exists $N(\epsilon) \geq 2$ such that $\sum_{n=0}^{N} \alpha_n \geq \sigma n/\lambda(1+\epsilon)$ for any $N \geq N(\epsilon)$. This implies that for $N \geq N(\epsilon)$ and $i \geq s + N(\epsilon)$

$$
E \left[ \tilde{Y}^{(N)}_{k+1} - i | \tilde{Y}^{(N)}_k = i \right] = \sum_{n=0}^{i} (1-n) \alpha_n - N(\epsilon) \sum_{n=i+1}^{\infty} \alpha_n \leq 1 - \sum_{n=0}^{N(\epsilon)} \alpha_n \leq \epsilon^* < 0
$$

where $\epsilon^* = 1 - (s\mu)/\lambda(1+\epsilon)$. Thus, by Pakes’s lemma [5], $\tilde{Y}^{(N)}$ is ergodic for $N \geq N(\epsilon)$.

If the DTMC $\tilde{Y}^{(N)}$ is ergodic, then $\tilde{X}$ and $\tilde{Y}^{(N)}$ are also ergodic and, moreover, the unique stationary probability row vectors $\pi^{(N)}$, $\tilde{\pi}^{(N)}$ of $\tilde{Y}^{(N)}$, $\tilde{X}$ and $\tilde{Y}^{(N)}$, respectively, are also limit probability row vectors of the corresponding DTMCs. Let $\alpha$ denote an arbitrary probability vector on $S^*$, then, in view of (8),

$$
\alpha(\hat{R}^{(N)})^n \leq_{st} \alpha \tilde{P}^n \leq_{st} \alpha(\tilde{R}^{(N)})^n
$$

for any $n, N \in \mathbb{N}$. Thus, (9) follows by taking the limit as $n$ tends to infinity in the previous expression, since the usual stochastic order is closed for limits [8].

---

2Given two probability row vectors $p$ and $q$, we say that $p$ is stochastically smaller than $q$ (in the usual sense), and write $p \leq_{st} q$, if and only if $\sum_{j \geq k} p_j \leq \sum_{j \geq k} q_j$ for all $k$. Moreover, if $X$ and $Y$ are discrete random variables with common support and respective probability row vectors $p$ and $q$, we write $X \leq_{st} Y$ if and only if $p \leq_{st} q$. 


We note that there are probabilistic interpretations for the DTMCs $\tilde{Y}^{(N)}$ and $\hat{Y}^{(N)}$. Namely, $\tilde{Y}^{(N)}$ corresponds to the number of customers seen by arriving customers to a GI/M/s/c system with a uniformizing Poisson clearing process with rate $s\mu$, whose $(N+1)$th renewal epoch in-between two consecutive customer arrivals would lead to the queueing system being emptied instantaneously. For the one server system, the clearing time would correspond to the $(N+1)$th departure time of a customer in-between two consecutive customer arrivals. Similarly, $\hat{Y}^{(N)}$ corresponds to the number of customers seen by arriving customers to a GI/M/s/c system with a uniformizing Poisson vacation process with rate $s\pi$, whose $(N+1)$th renewal epoch in-between two consecutive customer arrivals would lead to all servers going instantaneously on vacation until the next customer arrival time. For the one server system, the vacation of the server would be started at the $(N+1)$th departure time of a customer in-between two consecutive customer arrivals.

We note that we can bound the difference between the approximations $\hat{R}^{(N)}$ and $\tilde{R}^{(N)}$ of $\hat{P}$ of the transition probability matrix $\hat{P}$ of the DTMC $X$ by

$$
\| \hat{R}^{(N)} - \tilde{R}^{(N)} \|_{\infty} \leq (1 - [\hat{P}^N]_{c^*, 0}) \sum_{n=N+1}^{+\infty} \alpha_n. \tag{10}
$$

## 2.1 SOME IMPLEMENTATION DETAILS

In the implementation of the computation of the powers of the transition probability matrix $\hat{P}$ the special block-structure of the matrix has been explored, leading to strong reduction in computation time from the general procedure to compute the powers of a matrix. Namely, if $\hat{P}$ is partitioned in the following block-form

$$
\hat{P} = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}, \quad \hat{P}^n = \begin{bmatrix} B^n & 0 \\ C^n & D^n \end{bmatrix}, \quad n \geq 1 \tag{11}
$$

where $B$ ($D$) is a square matrix with of order $s$ ($c+1$) indices $0, 1, \ldots, s-1, s, s+1, \ldots, c^*$). It thus follows that the powers of $\hat{P}$ have the form given in (11), where:

- $[D^n]_{ij} = \delta_{i-j,n}$, so that in particular $D^n = 0$ for $n > c$;
- $[C^n]_{ij} = 0$ if $n < i - j$ and $[C^n]_{ij} = [B^{n-i+s-1}]_{s-1,j}$ otherwise.

Thus, recursive computation of $\hat{P}^n$ resumes in practice to the recursive computation of $B^n$, which reduces to

$$
[B^{n+1}]_{ij} = \begin{cases} 
[B^n]_{i,j} \cdot b_{ij} + [B^n]_{i,j+1} \cdot b_{j+1,j} & \text{if } i > j \land n + 1 \geq i - j \\
[B^n]_{i,i} \cdot b_{ii} & \text{if } i = j \\
0 & \text{if } i < j \lor n + 1 < i - j 
\end{cases}
$$

for $n \in \mathbb{N}$. As a result, the algorithm presented in Figure 1 may be used to compute the matrix powers $B^n$, $n \in \mathbb{N}$.

The presented procedure has a special interest to analyze GI/M/s/c queues that do not possess explicit expressions for the transitions probabilities of the embedded DTMCs, as occurs for Pareto renewal arrival processes. To compute the stationary distribution of the number of customers in system seen by arriving customers, the following auxiliary result is useful.

**Lemma 3** Let $W = \{W_k, k \geq 0\}$ be a DTMC on $S_W = \{0, 1, \ldots, m\}$ with transition probability matrix $P$ such that, for $i, j \in S_W$, $p_{ij} = 0$ if $j > i + 1$ and $p_{ij} > 0$ if $|i - j| \leq 1$. Then $W$ is ergodic and has invariant measure $\{\eta_j, j \in S_W\}$, where $\eta_m = 1$ and

$$
\eta_n = \frac{1}{p_{n,n+1}} \sum_{i=n+1}^{m} \eta_i \sum_{j=0}^{n} p_{ij}, \quad \text{for } n = m-1, m-2, \ldots, 0. \tag{12}
$$
Proof. For a measure \( \alpha \) on \( S_W \), let \( \alpha(A, B) = \sum_{j \in A} \sum_{k \in B} \alpha_j p_{jk} \) for subsets \( A \) and \( B \) of \( S_W \). Since \( p_{j} = 0 \) if \( j > i + 1 \) and \( p_{ij} > 0 \) if \( |i - j| \leq 1 \), the DTMC \( W \) satisfies \( \eta(A, \bar{A}) = \eta(\bar{A}, A) \) for any nonempty subset \( A \) of \( S_W \) (see [9]), where \( \bar{A} \) is the complement of \( A \) on \( S_W \). If we let \( A_n = \{1, 2, \ldots, n\} \), \( n = m - 1, m - 2, \ldots, 0 \), we get that \( \eta(A_n, \bar{A}_n) = \eta(\bar{A}_n, A_n) \), i.e. \( \eta_n p_{n,n+1} = \sum_{i=1}^{m} \eta_i \sum_{j=0}^{n} p_{ij} \). Thus, (12) holds. \( \square \)

Remark 4 The transition probability matrices \( \bar{P} \) and \( \bar{R}^{(N)} \) and \( \bar{\bar{R}}^{(N)} \), \( N \geq 2 \), have the form of the transition probability matrix of Lemma 3. Thus, the recursion (12) may be used to compute the stationary distributions of \( X \), \( \bar{Y}^{(N)} \) and \( \bar{\bar{Y}}^{(N)} \), with \( N \geq 2 \), when \( c \) is finite. After using the recursion (12) to compute invariant measures of \( \bar{R}^{(N)} \) and \( \bar{\bar{R}}^{(N)} \), normalization leads to the unique stationary distributions of \( \bar{Y}^{(N)} \) and \( \bar{\bar{Y}}^{(N)} \). \( \square \)

3. STEADY-STATE PERFORMANCE MEASURES

In this section we assume that the GI/M/s/c system has finite capacity, i.e., \( c < \infty \). For convenience, we let in this section \( P \), \( \bar{R} \) and \( \bar{\bar{R}} \) (\( \bar{\pi}, \bar{\bar{\pi}} \) denote the transition probability matrices (stationary probability vectors) of \( X \), \( \bar{Y}^{(N)} \) and \( \bar{\bar{Y}}^{(N)} \), where \( N \) is fixed.

We will consider the following steady-state performance measures (we will drop the adjective steady-state) of the state of the system immediately before arrival epochs: the loss probability, the loss-delay probability, the mean number of customers in the system (\( \mathbb{E}[L] \)), and the mean waiting time in queue of entering customers (\( \mathbb{E}[W_q] \)). We will also consider the distribution function of \( L \), the number of customers in the system at arrival epochs, and \( W_q \), the waiting time in queue of an entering customer. In what follows we will provide, whenever feasible, lower and upper bounds for these performance measures, since in general only approximate values of these measures are computed. The derivation of the formulas used follows that of M/M/s/c systems (see, e.g., Gross and Harris [6]).

The loss probability (long-run proportion of customers that are lost) is \( \pi_{c^*} \leq \bar{\pi}_{c^*} \leq \bar{\bar{\pi}}_{c^*} \). The loss-delay probability (the long-run fraction of customers that are lost or delayed) is \( \sum_{n=0}^{c^*} \pi_n \) and \( \sum_{n=0}^{c^*} \bar{\pi}_n \leq \sum_{n=0}^{c^*} \bar{\bar{\pi}}_n \leq \sum_{n=0}^{c^*} \pi_n \). The number of customers in the system seen by arriving customers, \( L \), has distribution \( \pi \) and we have

\[
\sum_{k=0}^{i} \pi_k \geq P(L \leq i) = \sum_{k=0}^{i} \pi_k \geq \sum_{k=0}^{i} \bar{\pi}_k, \quad i = 0, 1, \ldots, c^* - 1. \tag{13}
\]
As a consequence the mean number of customers in the system seen by arriving customers, $E[L]$, satisfies

$$\sum_{k=0}^{c^*} k\pi_k \leq E[L] = \sum_{k=0}^{c^*} k\pi_k \leq \sum_{k=0}^{c^*} k\pi_k. \quad (14)$$

The waiting time in queue of an entering customer, $W_q$, has survival function

$$P(W_q > t) = \sum_{j=0}^{c-1} e^{-\mu t} (s\mu)^j \frac{\sum_{k=s+j}^{c-1} \pi_k}{1 - \pi_{c^*}}, \quad t \geq 0 \quad (15)$$

and the mean waiting time experienced by entering customers is

$$E[W_q] = \sum_{k=s}^{c^*-1} (k - s + 1)\pi_k \frac{1}{s\mu(1 - \pi_{c^*})}. \quad (16)$$

4. NUMERICAL ILLUSTRATION

To show the power of the proposed approach we analyze Pareto/M/s/c systems and compare their results with those of GI/M/s/c systems with renewal arrival processes with uniform, exponential, deterministic and Erlang (with 3 and 10 phases) having a common finite mean interarrival time. We have computed the steady-state loss probability, the mean number of customers in the system seen by arriving customers and the mean waiting time in queue of entering customers. We also evaluate the loss probability for the $n$-th next arriving customer given that the present customer has been blocked.

We consider GI/M/s/c queues with service rate $\mu = s^{-1}$ and five different interarrival time distributions with mean $1/\lambda$ aside from the Pareto distribution (P) on $[0, 2/\lambda]$, Exponential (M) with rate $\lambda$, Deterministic (D), the constant $1/\lambda$, and Erlang (or gamma) with 3 and 10 phases, $G(3, 3\lambda)$ and $G(10, 10\lambda)$. The results have been computed with MATLAB algorithms and all the steady-state distributions have been obtained with an accuracy of $\varepsilon = 10^{-10}$.

The mixed-Poisson probabilities for the considered interarrival time distributions have been computed using the recursions proposed in [4]. For the Pareto($\beta, \kappa$) distribution

$$\alpha_n = \int_{\kappa}^{+\infty} e^{-s\mu t} (s\mu)^n \frac{\beta \kappa^{\beta}}{\Gamma(\beta + 1)} dt$$

Figure 2: Number of iterations to achieve an accuracy of $10^{-10}$ for the steady-state distribution of the state of the system as seen by arriving customers.
with $\kappa > 0$ and $\beta > 1$ (whose mean is $\beta \kappa / (\beta - 1)$), the mixed-Poisson probabilities are $\alpha_0 = \beta(s \mu \kappa)^0 \Gamma(-\beta, s \mu \kappa)$ and

$$\alpha_{n+1} = \frac{1}{n+1} \left((n-\beta) a_n + \beta e^{-s \mu \kappa} (s \mu \kappa)^n / n!\right), \quad n \geq 0$$

where $\Gamma(a, b) = \int_b^\infty e^{-y} y^{a-1} dy$ is the incomplete gamma function. This recursion is stable, easy to implement, and its computation time is linear on the number of computed coefficients.

We note that the partial sums of mixed-Poisson probabilities with Pareto structural distribution converge slowly to 1. Thus, in general the approximation of steady-state performance measures (within a given accuracy level) for GI/M/s/c queues requires more iterations for Pareto renewal processes than for the other renewal processes considered, as shown in Figure 2. Moreover, the number of iterations required to achieve a fixed accuracy decreases with the arrival rate and increases with the number of servers and the system capacity. As illustrated in Figure 2 for the Pareto arrival process, the increase in the number of iterations as the number of servers increases is stronger than the increase due to increased capacity. Roughly, one additional mixed-Poisson coefficient is needed for each unitary increase in the queue capacity.

As expected, and as can be seen in Figure 3, the results show that for all renewal arrival processes considered, the loss probability increases with the arrival rate and decreases with the queue capacity (for fixed number of servers) and the number of servers (for fixed queue capacity). We note that for not too large arrival rates the loss probabilities are strongly dependent on the interarrival time distribution. As the Pareto distributions is among those considered the one that has heavier left-tail, the Pareto/M/s/c system tends to be more
Figure 5: Loss probability for the $n$-th next arriving customer given that the last customer has been lost.

Figure 6: Mean number of customers seen in the system by arriving customers and mean waiting time in queue of entering customers versus arrival rate.

congested that the other systems considered for a common traffic intensity. Moreover, from the point of view of arriving customers, Pareto/M/s/c systems tend to be congested even for small arrival rates. As a result, queues with Pareto renewal arrival process exhibit higher loss probabilities than the other considered queues, with the difference becoming more significant for small arrival rates. Moreover, as highlighted in Figure 4, increases in the queue capacity lead to much smaller reductions in the loss probability when the renewal arrival process is Pareto than for the other considered cases.

Figure 5 shows that high loss probabilities may persist for a rather long sequence of arriving customers after the loss of a customer takes place for Pareto renewal arrival process. This is consistent with the view that strong dependence in customer losses may be observed for Pareto/M/s/c systems and that the service may deteriorate for long periods of time. For the other considered renewal arrival processes the loss probability for the $n$-th next arriving customer given that the last customer has been blocked decreases to its limit value much faster. This decrease is faster for smaller arrival rates. Results not given show that the loss-delay probability decreases with the number of servers and increases with the arrival rate and with the system capacity.

Figure 6 displays the mean number of customers seen in the system by arriving customers and mean waiting time in queue of entering customers as a function of the customer arrival rate. For large queue capacity the graphs obtained with exponential, gamma, uniform and deterministic distributions are similar but quite different of those obtained for Pareto arrivals (except for fairly large arrival rates). For low arrival rates ($\lambda < 1$), the mean number of customers seen in the system by arriving customers and mean waiting time in queue of entering customers are much larger for Pareto/M/s/c systems than for
the other considered systems.

We note that the results for the M/M/s/c systems are the ones that approach the most the results for Pareto/M/s/c systems. However, the results tend to be fairly distinct, specially for small customer arrival rates. Also, the results for the renewal arrival processes with deterministic and Erlang with 10 phases tend to be fairly close to each other.

5. CONCLUSIONS

In this paper we have shown that the classical embedded Markov chain approach may be combined with uniformization to analyze GI/M/s/c queues. The technique derived provides a recursive way to approximate, with any desired degree of accuracy, the one step transition probabilities of the number of customers in the system as seen by arriving customers and its associated steady-state distribution. The derived procedure has special interest to analyze GI/M/s/c queues that do not possess explicit expressions for the transitions probabilities of the embedded DTMCs, as occurs for Pareto, Gamma and Uniform renewal arrival processes, and provides an alternative to the use of simulation. MATLAB algorithms have been implemented to obtain transient and steady-state performance measures related to the state of GI/M/s/c systems as seen by arriving customers. These measures were compared for systems with different inter-arrival distributions. In the numerical results special attention has been given to Pareto/M/s/c systems. The customers experience much higher loss probabilities and mean waiting time in queue in these systems than in similar systems with exponential, uniform, gamma or deterministic interarrival times. This phenomenon is especially evident for small customer arrival rates.

ACKNOWLEDGEMENTS

This research was supported in part by Fundação para a Ciência e a Tecnologia, grant SFRH/BSAB/251/01 and the projects POSI/34826/CPS/2000 and POSI/42069/CPS/2001.

REFERENCES


