RECURSIVE COMPUTATION OF STEADY-STATE PROBABILITIES IN PRIORITY QUEUES

Israel CIDON
IBM, Thomas J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598, USA

Moshe SIDY
Electrical Engineering Department, Technion – Israel Institute of Technology, Haifa 32000, Israel

Received May 1988
Revised September 1989

Exact recursive formulas are derived for the state probabilities in priority queueing systems (preemptive and non-preemptive). The derivation is based only on the general structure of the generating function involved, and thus is simpler and more general than previous methods. Furthermore, applications of the method to other queueing systems are discussed.

priority queues • recursive computation • steady-state probabilities

1. Introduction

Queueing models are frequently used in the process of evaluating the performance of practical systems. The most common performance measures for the behavior of a system are the moments (especially first and second moments) of the buffers occupancy. However, in many cases the buffer occupancy probabilities themselves are needed. Frequently, the designer is interested in the probability that the queue length will exceed some predetermined thresholds. In such a case there is a need to calculate the buffer occupancy probabilities up to some large number. A typical example is the performance evaluation of a communication store and forward node that has an hierarchy of memories and each of them has different speed and size. Messages are stored in the fastest memory available to the node. In order to compute the probability of arriving packets to be stored at some level of the memory, we need to know the probabilities of exceeding the corresponding thresholds.

In many queueing systems it is relatively simple to derive the generating function of the number of customers in the system. If this function is not too complex, one would be able to derive the first two moments of the number of customers. However, in most cases, obtaining the probability distribution might prove to be a challenging task. An example of such a system is a priority (preemptive or non-preemptive) queueing system (Jaiswal [5]). The goal of this paper is to introduce a simple derivation of recursive formulas for the state probabilities in a large class of queueing systems including priority queueing systems. The details of the recursion are given in Sections 2 and 3 for the preemptive and the non-preemptive single-server systems, respectively. In Section 4 we show that the same technique can be applied to multi-server priority systems, and we also discuss the applications of the method to systems other than priority systems.

White and Christie [12] derived complicated expressions for the probability distribution for the number of customers in a preemptive priority queueing system. Recursive computation of the joint state probabilities for the two priority classes, has been presented by Marks [6] and Miller [7], for the preemptive and the non-preemptive priority (with the same exponential service distribution for both
classes of customers in [6] systems. The method of [6] is by far more complicated and hard to generalize than the method of [7] and the method that we present in this paper. Miller [7] derived his recursive formulas by using Neut's theory of matrix-geometric invariant probability vector. In both [6] and [7], it is not obvious how to obtain marginal probabilities. Note that recursive computation of joint probabilities does not yield marginal probabilities, since (in general) infinite summations should be taken on the joint probability distribution to obtain the marginals, while recursive computations yield only finite number of terms (as large as one wants, but not infinite).

The derivation of the recursive formulas in this paper is completely different than that of [6] and [7]. It is based only on the general structure of the underlying generating function. Consequently, it is simply applied to both the marginal and the joint probabilities. Generally, recursive computations have many advantages over other numerical computation techniques. They are easy to program and their accuracy depends only on the accuracy of computing the basic built in functions (such as additions, multiplications, square roots etc.). For a further discussion on accuracy checks, see [7].

In the following, we first give a very short review of standard (non-recursive) techniques to obtain the corresponding probabilities from a generating function. Then we describe the basic idea of the recursive approach.

Let \( N \) be a random variable that represents the number of customers in a queueing system in steady-state. Let \( g_n = \Pr(N = n), n = 0, 1, 2, \ldots \), and \( G(Z) \) be the generating function of \( N \),

\[
G(Z) = \sum_{n=0}^{\infty} g_n Z^n, \quad |Z| < 1, \tag{1}
\]

and \( G(Z) \) is analytic in the disk \( |Z| < 1 \). For a given \( G(Z) \), one can obtain the occupancy probabilities (at least in principal) by \( g_n = G^{(n)}(0)/n! \) \( (n = 0, 1, \ldots) \) where \( G^{(n)}(0) = d^n G(Z)/dZ^n \big|_{Z=0} \). In most cases, the computation of successive differentiations, would be quite cumbersome (the reader is invited to try this approach for the relatively simple generating function in (10)). To alleviate this difficulty, one can use the Cauchy integral formula

\[
G^{(n)}(0) = \frac{n!}{2\pi i} \oint_{\Gamma} \frac{G(t)}{t^{n+1}} \, dt, \quad n = 0, 1, 2, \ldots, \tag{2}
\]

where \( \Gamma = \gamma e^{i\tau}, 0 < \gamma < 1, 0 \leq \tau \leq 2\pi \). In general, the computation above requires numerical integration and therefore the accuracy of this computation depends on the exact integration method and its parameters (mainly the granularity of the discrete division). A better way to approximate \( g_n, n = 0, 1, 2, \ldots, \) to some desired accuracy that does not require numerical integration, has been introduced by Jagerman [3]. Yet, both these methods, as well as methods that are based on inversion techniques for Laplace transforms [4], yield only approximate results since they use approximation sequence or truncation of infinite sums.

**Exact recursive computation**

The basic idea of the recursive approach is the following (see Gross and Harris [2, p.262]). Suppose that \( G(Z) \) is given by the quotient of two functions \( A(Z) \) and \( B(Z) \),

\[
G(Z) = \frac{A(Z)}{B(Z)}, \tag{3}
\]

and assume that the coefficients \( a_i, b_i, i = 1, 2, \ldots, n \), of the Taylor expansion are known for both \( A(Z) \) and \( B(Z) \), where

\[
A(Z) = \sum_{i=0}^{\infty} a_i Z^i, \quad B(Z) = \sum_{i=0}^{\infty} b_i Z^i. \tag{4}
\]

Then if \( b_0 > 0 \) it is easy to see that \( g_n, n = 0, 1, 2, \ldots, \) are computed via the following recursion:

\[
g_0 = a_0/b_0, \quad g_n = \left( a_n - \sum_{i=0}^{n-1} b_{n-i} g_i \right)/b_0, \quad n > 0. \tag{5}
\]

\[250\]
Therefore, in order to compute \( g_n \) for some \( n \), we only need to be able to compute \( a_i, b_i, i = 0, 1, \ldots, n \), namely to determine the first \( n + 1 \) coefficients of the Taylor series of \( A(Z) \) and \( B(Z) \) in (3). In Sections 2 and 3 we show how to compute these coefficients in a recursive manner for priority queueing systems.

Note that in some examples \( A(Z) \) and similarly \( B(Z) \) might be expressed as a product of several terms, namely \( A(Z) = \prod_{k=1}^{K} A_k(Z) \). In such cases it is sufficient to determine the first \( n + 1 \) coefficients, \( (a^0_0, a^1_1, \ldots, a^n_n) \) of the Taylor series of each \( A_k(Z) \). Then the coefficients of \( A(Z) \) are obtained by the convolution

\[
a_i = \sum_{L_i = 1}^{K} \prod_{k=1}^{K} a^k_i, \quad 0 \leq i \leq n \quad \text{with} \quad L_i = \left\{ (i_1, i_2, \ldots, i_K) : i_k \geq 0 \text{ for } k = 1, 2, \ldots, K; \sum_{k=1}^{K} i_k = i \right\}.
\]

(5)

The same derivation applies to multi-dimensional generating functions. We demonstrate this for a two-dimensional generating function. Assume that \( G(X, Y) \) is a two-dimensional generating function which is given by the quotient of two functions \( A(X, Y) \) and \( B(X, Y) \). Thus let

\[
G(X, Y) = A(X, Y)/B(X, Y),
\]

(6)

and assume that the coefficients \( a_{i,j}, b_{i,j} \) \((i = 0, 1, 2, \ldots, n; j = 0, 1, 2, \ldots, m)\) of the two-dimensional Taylor expansion are known for both \( A(X, Y) \) and \( B(X, Y) \):

\[
A(X, Y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{i,j} X^i Y^j, \quad B(X, Y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} b_{i,j} X^i Y^j.
\]

(7)

Then if \( b_{0,0} > 0 \) it is easy to see that \( g_{n,m}, n, m = 0, 1, 2, \ldots, \) are computed via the following recursion:

\[
g_{0,0} = a_{0,0}/b_{0,0}, \quad g_{n,m} = \left( a_{n,m} - \sum_{i=0}^{n} \sum_{j=0}^{m} b_{n-i,m-j} g_{i,j} \right) / b_{0,0} \quad (n + m > 0).
\]

(8)

Therefore, in order to compute \( g_{n,m} \) for some \( n \) and \( m \), we only need to be able to compute \( a_{i,j}, b_{i,j} \) \((i = 0, 1, \ldots, n; j = 0, 1, \ldots, m)\), namely to determine the first \( (n + 1)(m + 1) \) coefficients of the two-dimensional Taylor expansion of \( A(X, Y) \) and \( B(X, Y) \) in (7). Later we will show how to compute these coefficients (again in a recursive manner) for priority queueing systems. Note that the same method can be used to higher dimensions generating functions as well. Again, if \( A(X, Y) = \prod_{k=1}^{K} A_k(X, Y) \), we have that

\[
a_{i,j} = \sum_{L_{i,j}}^{K} \prod_{k=1}^{K} a^k_{i,j} \quad \text{with} \quad L_{i,j} = \left\{ (i_1, j_1, i_2, j_2, \ldots, i_K, j_K) : i_k \geq 0, j_k \geq 0 \right\} \quad \text{for} \quad k = 1, 2, \ldots, K; \sum_{k=1}^{K} i_k = i; \sum_{k=1}^{K} j_k = j.
\]

(9)

2. Preemptive priority queueing system

Consider the two-class preemptive priority queueing system with a single server and two classes of customers, each class having Poisson arrivals and exponentially distributed service times. The high priority class (arrival rate \( \lambda_1 \), service rate \( \mu_1 \)) has preemptive head of the line priority over the low priority customers (arrival rate \( \lambda_2 \), service rate \( \mu_2 \)). We are interested in computing the steady state occupancy probabilities of the high and the low priority customers. Denote \( \rho_i = \lambda_i/\mu_i \) \((i = 1, 2)\), \( \rho_0 = 1 - \rho_1 - \rho_2 \) and \( \lambda = \lambda_1 + \lambda_2 \). The high priority class behaves as an \( M/M/1 \) queueing system and therefore the steady state
probability of having \( n \) high priority customers in the system is \((1 - \rho_1)\rho_1^n\). For the low priority class it is known (Miller [9]) that

\[
G(Z) = p_0/(1 - a(Z) - \rho_2 Z),
\]

where

\[
a(Z) = \left(\mu_1 + \lambda - \lambda_2 Z - \sqrt{(\mu_1 + \lambda - \lambda_2 Z)^2 - 4\lambda_1 \mu_1}\right)/(2\mu_1).
\]

The function \( G(Z) \) is of the form (2) with

\[
A(Z) = p_0, \quad B(Z) = 1 - a(Z) - \rho_2 Z.
\]

In order to use the recursive method of (4) we have to compute \( a_i, b_i \) (\( i = 0, 1, 2, \ldots, n \)). For the \( a_i \)'s we have \( a_0 = p_0, a_i = 0 \) (\( i = 1, 2, \ldots \)). To compute the \( b_i \)'s we first expand \( \beta(Z) = \sqrt{(\mu_1 + \lambda - \lambda_2 Z)^2 - 4\lambda_1 \mu_1} \) as a Taylor series \( \beta(Z) = \sum_{i=0}^{\infty} \beta_i Z^i \) and determine \( \beta_i \) (\( i = 0, 1, 2, \ldots, n \)). We do this again in a recursive manner by using the relation \( (\sum_{i=0}^{\infty} \beta_i Z^i)^2 = (\mu_1 + \lambda - \lambda_2 Z)^2 - 4\lambda_1 \mu_1 \). This gives

\[
\beta_0 = \sqrt{(\mu_1 + \lambda)^2 - 4\lambda_1 \mu_1}, \quad \beta_1 = -\lambda_2 (\mu_1 + \lambda) / \beta_0,
\]

\[
\beta_2 = (\lambda_2^2 - \beta_2^2) / (2\beta_0), \quad \beta_i = -\sum_{j=1}^{i-1} \beta_j \beta_{i-j} / (2\beta_0), \quad i = 3, 4, \ldots.
\]

Therefore, for the expansion of \( a(Z) = \sum_{i=0}^{\infty} a_i Z^i \), we have

\[
a_0 = (\mu_1 + \lambda - \beta_0) / (2\mu_1), \quad a_1 = -\lambda_2 (\mu_1 + \lambda) / (2\mu_1), \quad a_i = -b_i / (2\mu_1), \quad i = 2, 3, \ldots.
\]

From (12) we then have that

\[
b_0 = 1 - a_0, \quad b_1 = -a_1 - \rho_2, \quad b_i = -a_i, \quad i = 2, 3, \ldots,
\]

and now we apply recursion (4) to obtain the steady state probabilities of having \( n \) low priority customers in the system.

To derive the recursive relations for the joint probability distribution let \( g_{n,m} \) be the steady-state probabilities of having \( n \) high priority customers and \( m \) low priority customers in the system. Then, \( G(X, Y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} g_{n,m} X^n Y^m \) is given by (see [9])

\[
G(X, Y) = \frac{p_0}{1 - a(Y)} - \frac{1 - a(Y)}{1 - \rho_2 Y} 1 - Xa(Y),
\]

where \( a(Y) \) is given in (11). Therefore, we have

\[
a_{0,0} = p_0 (1 - a_0), \quad a_{0,j} = -p_0 a_j, \quad (j \geq 1), \quad a_{i,j} = 0, \quad (i \geq 1, j \geq 0),
\]

\[
b_{1,j} = -\sum_{i=0}^{j} b_i a_{j-i}, \quad b_{0,j} = b_j, \quad (j \geq 0), \quad b_{i,j} = b_j, \quad (i \geq 2, j \geq 0),
\]

where \( b_j \) is given in (15), and now we apply recursion (8).

3. Non-preemptive priority queueing system

In this section we consider the system of Section 2 when preemption is not allowed. Note that unlike [6], we do not assume that \( \mu_1 = \mu_2 \).
3.1. High priority customers

From [9] we find that the generating function of the occupancy probabilities of high priority customers is given by

\[ G(Z) = \frac{p_0(\mu_2 + \lambda_1(1-Z)) + \lambda_2}{(1-p_z Z)[\mu_2 + \lambda_1(1-Z)]}. \]  

(19)

So in this case,

\begin{align*}
    a_0 &= p_0(\mu_2 + \lambda_1) + \lambda_2, \quad a_1 = -\lambda_1 p_0, \quad a_i = 0 \quad (i \geq 2), \\
    b_0 &= \mu_2 + \lambda_1, \quad b_1 = -\rho_1(\mu_2 + \lambda_1) - \lambda_1, \quad b_i = \lambda_1 \rho_1, \quad b_i = 0 \quad (i \geq 3),
\end{align*}

(20)

(21)

and now we apply recursion (4) to obtain \( g_n \). Due to the simplicity of this recursion, it follows immediately (see [7]) that

\[ g_n = p_0 \rho_1^n + \lambda_2 \rho_1^n \left[ 1 - \left( \frac{\mu_1}{(\lambda_1 + \mu_2)} \right)^{n+1} \right] / (\lambda_1 + \mu_2 - \mu_1). \]

3.2. Low priority customers

Again from [9] we have that the generating function of the occupancy probabilities of low priority customers is

\[ G(Z) = \frac{p_0[\lambda - \lambda_2 Z + \mu_2 - \mu_1 \alpha(Z)]}{[\lambda_2(1-Z) + \mu_2][1 - \alpha(Z) - \rho_2 Z]}. \]

(22)

In this case,

\begin{align*}
    a_0 &= p_0(\lambda + \mu_2 - \mu_1 \alpha_0), \quad a_1 = -p_0(\lambda_2 + \mu_1 \alpha_1), \quad a_i = -p_0 \mu_1 \alpha_i \quad (i \geq 2), \\
    b_0 &= \lambda_2 + \mu_2, \quad b_1 = -\lambda_2, \quad b_1^i = 0 \quad (i \geq 2), \\
    b_0^i &= 1 - \alpha_0, \quad b_1^i = -\alpha_1 - \rho_2, \quad b_1^2 = -\alpha_1 \quad (i \geq 2).
\end{align*}

(23)

(24)

(25)

Following (7), \( b_i = \sum_{i=0}^k b_i b_{i-1}^j \) \( (i \geq 0) \) and we apply recursion (4) to obtain \( g_n \).

3.3. Joint probability distribution

From [9] we have

\[ G(X, Y) = G_1(X, Y) + G_2(X, Y), \]

(26)

where

\[ G_1(X, Y) = \frac{A_1}{B_1}, \quad G_2(X, Y) = \frac{A_2}{B_2}, \]

(27)

with

\begin{align*}
    A_{11} &= \rho_2 p_0 (\lambda + \mu_2 - \lambda_2 Y - \mu_1 \alpha(Y)), \quad A_{12} = \mu_1 Y(X - 1) - \mu_2 Y(X - 1), \\
    B_{11} &= \lambda_1 (1-X) + \lambda_2 (1-Y) + \mu_2, \quad B_{12} = \lambda_1 X(1-X) + \lambda_2 X(1-Y) + \mu_1 (X - 1), \\
    B_{13} &= 1 - \alpha(Y) - \rho_2 Y, \quad B_{12} = \lambda_1 (1-X), \quad B_{2} = B_{12}.
\end{align*}

(28)

(29a)

(29b)

(30)

Using the Taylor expansion of \( \alpha(Y) \) from (14), it is straightforward to compute the Taylor expansions of \( A_{11} \) and \( B_{13} \) up to any arbitrary number. Using (9), one recursively computes the tailor expansion of the
Table 1
Non-preemptive priority

<table>
<thead>
<tr>
<th>n</th>
<th>Probability for n high priority</th>
<th>Probability for n low priority</th>
<th>Probability for n total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.714 E -01</td>
<td>2.487 E -01</td>
<td>1.000 E -01</td>
</tr>
<tr>
<td>1</td>
<td>1.575 E -01</td>
<td>1.140 E -01</td>
<td>8.175 E -02</td>
</tr>
<tr>
<td>2</td>
<td>1.318 E -01</td>
<td>7.643 E -02</td>
<td>6.897 E -02</td>
</tr>
<tr>
<td>3</td>
<td>1.071 E -01</td>
<td>5.929 E -02</td>
<td>5.986 E -02</td>
</tr>
<tr>
<td>4</td>
<td>8.620 E -02</td>
<td>4.884 E -02</td>
<td>5.298 E -02</td>
</tr>
<tr>
<td>5</td>
<td>6.910 E -02</td>
<td>4.151 E -02</td>
<td>4.751 E -02</td>
</tr>
<tr>
<td>6</td>
<td>5.531 E -02</td>
<td>3.598 E -02</td>
<td>4.296 E -02</td>
</tr>
<tr>
<td>7</td>
<td>4.426 E -02</td>
<td>3.160 E -02</td>
<td>3.908 E -02</td>
</tr>
<tr>
<td>8</td>
<td>3.541 E -02</td>
<td>2.804 E -02</td>
<td>3.571 E -02</td>
</tr>
<tr>
<td>9</td>
<td>2.833 E -02</td>
<td>2.507 E -02</td>
<td>3.275 E -02</td>
</tr>
<tr>
<td>10</td>
<td>2.266 E -02</td>
<td>2.255 E -02</td>
<td>3.010 E -02</td>
</tr>
<tr>
<td>11</td>
<td>1.813 E -02</td>
<td>2.038 E -02</td>
<td>2.774 E -02</td>
</tr>
<tr>
<td>12</td>
<td>1.450 E -02</td>
<td>1.849 E -02</td>
<td>2.560 E -02</td>
</tr>
<tr>
<td>13</td>
<td>1.160 E -02</td>
<td>1.683 E -02</td>
<td>2.367 E -02</td>
</tr>
<tr>
<td>14</td>
<td>9.284 E -03</td>
<td>1.537 E -02</td>
<td>2.192 E -02</td>
</tr>
<tr>
<td>15</td>
<td>7.427 E -03</td>
<td>1.407 E -02</td>
<td>2.032 E -02</td>
</tr>
</tbody>
</table>

product terms $A_1$ and $B_1$ of (27). Using (8), it is easy to recursively compute the Taylor series coefficients of the quotients $G_1(X, Y)$ and $G_2(X, Y)$ of (26) and then the coefficients of $G(X, Y)$

3.4. Numerical examples

Numerical examples that are derived via the recursions developed in this section, are given in Tables 1–3. Table 1 corresponds to high load of high priority customers and low load of low priority customers. Table 2 corresponds to high load of low priority customers and low load of high priority customers. Table 3 corresponds to the same load for the two types of customers. The first two columns in each table

Table 2
Non-preemptive priority

<table>
<thead>
<tr>
<th>n</th>
<th>Probability for n high priority</th>
<th>Probability for n low priority</th>
<th>Probability for n total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8.619 E -01</td>
<td>1.057 E -01</td>
<td>1.000 E -01</td>
</tr>
<tr>
<td>1</td>
<td>1.224 E -01</td>
<td>9.051 E -02</td>
<td>8.772 E -02</td>
</tr>
<tr>
<td>2</td>
<td>1.397 E -02</td>
<td>7.860 E -02</td>
<td>7.743 E -02</td>
</tr>
<tr>
<td>3</td>
<td>1.479 E -03</td>
<td>6.910 E -02</td>
<td>6.882 E -02</td>
</tr>
<tr>
<td>4</td>
<td>1.518 E -04</td>
<td>6.134 E -02</td>
<td>6.155 E -02</td>
</tr>
<tr>
<td>5</td>
<td>1.537 E -05</td>
<td>5.486 E -02</td>
<td>5.532 E -02</td>
</tr>
<tr>
<td>6</td>
<td>1.546 E -06</td>
<td>4.932 E -02</td>
<td>4.990 E -02</td>
</tr>
<tr>
<td>7</td>
<td>1.550 E -07</td>
<td>4.451 E -02</td>
<td>4.513 E -02</td>
</tr>
<tr>
<td>8</td>
<td>1.552 E -08</td>
<td>4.028 E -02</td>
<td>4.090 E -02</td>
</tr>
<tr>
<td>9</td>
<td>1.553 E -09</td>
<td>3.652 E -02</td>
<td>3.712 E -02</td>
</tr>
<tr>
<td>10</td>
<td>1.554 E -10</td>
<td>3.315 E -02</td>
<td>3.372 E -02</td>
</tr>
<tr>
<td>11</td>
<td>1.554 E -11</td>
<td>3.012 E -02</td>
<td>3.065 E -02</td>
</tr>
<tr>
<td>12</td>
<td>1.554 E -12</td>
<td>2.738 E -02</td>
<td>2.787 E -02</td>
</tr>
<tr>
<td>13</td>
<td>1.554 E -13</td>
<td>2.490 E -02</td>
<td>2.536 E -02</td>
</tr>
<tr>
<td>14</td>
<td>1.554 E -14</td>
<td>2.266 E -02</td>
<td>2.308 E -02</td>
</tr>
<tr>
<td>15</td>
<td>1.554 E -15</td>
<td>2.062 E -02</td>
<td>2.101 E -02</td>
</tr>
</tbody>
</table>
Table 3
Non-preemptive priority

\[ \lambda_1 = 0.4, \lambda_2 = 0.8, \mu_1 = 1, \mu_2 = 2 \]

<table>
<thead>
<tr>
<th>n</th>
<th>Probability for ( n ) high priority</th>
<th>Probability for ( n ) low priority</th>
<th>Probability for ( n ) total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5.333 E -01</td>
<td>2.678 E -01</td>
<td>2.000 E -01</td>
</tr>
<tr>
<td>1</td>
<td>2.688 E -01</td>
<td>1.608 E -01</td>
<td>1.462 E -01</td>
</tr>
<tr>
<td>2</td>
<td>1.168 E -01</td>
<td>1.082 E -01</td>
<td>1.108 E -01</td>
</tr>
<tr>
<td>3</td>
<td>4.826 E -02</td>
<td>7.949 E -02</td>
<td>8.693 E -02</td>
</tr>
<tr>
<td>4</td>
<td>1.956 E -02</td>
<td>6.176 E -02</td>
<td>7.605 E -02</td>
</tr>
<tr>
<td>5</td>
<td>7.568 E -03</td>
<td>4.966 E -02</td>
<td>5.754 E -02</td>
</tr>
<tr>
<td>6</td>
<td>3.154 E -03</td>
<td>4.078 E -02</td>
<td>4.790 E -02</td>
</tr>
<tr>
<td>7</td>
<td>1.263 E -03</td>
<td>3.393 E -02</td>
<td>4.024 E -02</td>
</tr>
<tr>
<td>8</td>
<td>5.054 E -04</td>
<td>2.848 E -02</td>
<td>3.403 E -02</td>
</tr>
<tr>
<td>9</td>
<td>2.022 E -04</td>
<td>2.405 E -02</td>
<td>2.891 E -02</td>
</tr>
<tr>
<td>10</td>
<td>8.088 E -05</td>
<td>2.041 E -02</td>
<td>2.465 E -02</td>
</tr>
<tr>
<td>11</td>
<td>3.235 E -05</td>
<td>1.738 E -02</td>
<td>2.107 E -02</td>
</tr>
<tr>
<td>12</td>
<td>1.294 E -05</td>
<td>1.484 E -02</td>
<td>1.806 E -02</td>
</tr>
<tr>
<td>13</td>
<td>5.176 E -06</td>
<td>1.270 E -02</td>
<td>1.550 E -02</td>
</tr>
<tr>
<td>14</td>
<td>2.070 E -06</td>
<td>1.089 E -02</td>
<td>1.332 E -02</td>
</tr>
<tr>
<td>15</td>
<td>8.283 E -07</td>
<td>9.358 E -03</td>
<td>1.147 E -02</td>
</tr>
</tbody>
</table>

corresponds to marginal probabilities of high and low priority customers. The third column in the tables is obtained from the joint probability distribution by computing \( \sum_{m=0}^{n} s_{m,n-m} \). Notice that the recursions developed in this paper for the joint probability distributions are similar to those in [6] and hence the computational effort is similar. Miller [6] also discusses ways of monitoring the accuracy of the numerical results. In our examples we did not face any numerical instabilities.

4. Other systems

The method described in this paper is easily generalized to handle multi-server priority systems (Gail et al. [1], Miller [8]). The basic difference is that in a multi-server system, the probability distribution is expressed in terms of a vector of generating functions of the form (see [1]):

\[
G(Z) = B^{-1}(Z)A(Z), \quad (31)
\]

where

\[
G(Z) = (G^1(Z), G^2(Z), \ldots , G^K(Z))^T, \quad A(Z) = (A^1(Z), A^2(Z), \ldots , A^K(Z))^T,
\]

\[
B(Z) = \begin{bmatrix}
B^{11}(Z) & B^{12}(Z) & \ldots & B^{1K}(Z) \\
B^{21}(Z) & B^{22}(Z) & \ldots & B^{2K}(Z) \\
\vdots & \vdots & \ddots & \vdots \\
B^{K1}(Z) & B^{K2}(Z) & \ldots & B^{KK}(Z)
\end{bmatrix}.
\]

We are interested in computing the coefficients \( g^i_k \) (0 \( \leq \) \( i \) \( \leq \) \( n \)) of \( G^i(Z) \) for 1 \( \leq \) \( k \) \( \leq \) \( K \). The expansion of each element in the vector \( A(Z) \) (namely, the coefficients \( a^i_m \) for 1 \( \leq \) \( m \) \( \leq \) \( K \), 0 \( \leq \) \( i \) \( \leq \) \( n \)) and each element in the matrix \( B(Z) \) (namely, the coefficients \( b^{i,m} \) for 1 \( \leq \) \( k \) \( \leq \) \( K \), 1 \( \leq \) \( m \) \( \leq \) \( K \), 0 \( \leq \) \( i \) \( \leq \) \( n \)) is done in the same way that we expanded \( A(Z) \) and \( B(Z) \) in (12) or in (22)–(25). Once these coefficients are known, the coefficients \( g^i_k \) (1 \( \leq \) \( k \) \( \leq \) \( K \); 0 \( \leq \) \( i \) \( \leq \) \( n \)) are computed recursively, assuming that the matrix \( b_0 \) defined below is invertible:

\[
g_0 = b_0^{-1}a_0, \quad g_i = b_0^{-1} \left[ a_i - \sum_{l=0}^{i-1} b_{i-l} g_l \right] \quad (i \geq 1), \quad (32)
\]
where
\[ g_i = \left( g_{i1}, g_{i2}, \ldots, g_{ik} \right)^T, \quad a_i = \left( a_{i1}, a_{i2}, \ldots, a_{ik} \right)^T, \]
\[ b_i = \begin{pmatrix} b_{i1}^{11} & b_{i1}^{12} & \cdots & b_{i1}^{1k} \\ b_{i2}^{21} & b_{i2}^{22} & \cdots & b_{i2}^{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{ik}^{k1} & b_{ik}^{k2} & \cdots & b_{ik}^{kk} \end{pmatrix}. \]

We also remark that our method is not limited to priority queueing systems. We have applied this method in other systems such as the discrete-time tandem in Morrison [10], the two interfering queues in Sidi and Segall [11] and M/G/1 systems in which the Laplace transform of the service time is some rational function.

Acknowledgement

The authors thank Dr. Inder Gopal for helpful discussions which lead to the formulation of the problem.

References