ANALYSIS OF AN N/G/1 FINITE QUEUE WITH THE SUPPLEMENTARY VARIABLE METHOD

GYEMIN LEE and JONGWOO JEON

Seoul National University
Department of Statistics
Seoul 151-742, Korea

(Received April, 1997; Revised April, 1998)

In this paper, we suggest a new approach to the analysis of an N/G/1 finite queue with the supplementary variable method. Compared to the conventional approach, our approach yields a simpler formula for the queue length distribution, which in turn gives a more efficient computational algorithm. Also, the new approach enables us to derive the joint density of the queue length and the elapsed service time.

Key words: N-process, Elapsed Service Time, Imbedded Markov Chain, Supplementary Variable Method, Schur-Banachiewicz Formula.

AMS subject classifications: 60K25, 68M20.

1. Introduction

Very complex input flows often occur in integrated service communication systems. As an approximation to such a stream, Neuts [4] introduced the N-process. This N-process is analytically tractable and can appropriately represent the correlation and burstiness of the stream. Many familiar arrival processes are special cases of the N-process.

To investigate the performance of the service facility with finite resources, Blondia [1] considered an N/G/1 finite queue, i.e., a single server queue with $K$ waiting rooms in which customers arrive according to an N-process. For the analysis, he used the imbedded Markov chain technique upon service completion epochs. He also gave a computational algorithm for the queue length distribution of the system by using the Schur-Banachiewicz formula [3] for the inverse of the block matrices.

However, the computational algorithm suggested by Blondia [1] needs a large amount of work. This motivated us to study an N/G/1 finite queue. Our aim is to obtain a more efficient computational algorithm. To this end, we employ the supplementary variable method originated by Cox [2].

This paper is organized as follows. In Section 2, we define N-process as originally supported by KOSEF, 1996.
introduced in Neuts [4]. Section 3 consists of the joint density of the queue length and the elapsed service time and the distribution of the queue length of the N/G/1 finite queue.

2. N-Process

Consider a continuous-time Markov process with \( m \) transient states and a single absorbing state. Then the infinitesimal generator of this Markov chain has the form

\[
Q = \begin{pmatrix}
T & T^0 \\
0 & 0
\end{pmatrix},
\]

where \( T \) is an \( m \times m \) non-singular matrix with \( T_{i,i} < 0, T_{i,j} \geq 0 \) for \( i \neq j \). The vector \( T^0 \) is non-negative and satisfied \( Te + T^0 e = 0 \), with \( e = (1,\ldots,1)^T \). Let \( (\alpha, \alpha_{m+1}) \) be a vector of initial state probabilities of the Markov process. In what follows, we shall assume that \( \alpha_{m+1} = 0 \).

Now, construct a continuous-time process by restarting the above Markov process \( Q \) instantaneously after each absorption through a multinomial trial with probability \( \alpha \) and outcomes \( 1,\ldots,m \). Then this process is also a Markov process with the state space \( \{1,2,\ldots,m\} \) and the infinitesimal generator

\[
Q^* = T + T^0 A^0,
\]

where \( T^0 \) is an \( m \times m \) matrix whose columns are all \( T^0 \) and \( A^0 = diag(\alpha_1,\ldots,\alpha_m) \). A transition from the state \( i \) to the state \( j \) in the Markov process \( Q^* \), which does not involve absorption, will be called an \((i,j)\)-transition, while the others are called \((i,j)\)-renewal transition. Then the N-process is an arrival process defined in the following way [4].

1. During any sojourn in the state \( i \), there are Poisson group arrivals of rate \( \lambda_i \) and group size of densities \( \Phi_i(k), k \geq 0 \). We shall denote \( \Phi_i^*(z) \) the p.g.f. of \( \{\Phi_i(k), k \geq 0\} \), and define \( \Lambda = diag(\lambda_1,\ldots,\lambda_m) \) and \( \Phi(z) = diag(\Phi_1^*(z), \ldots,\Phi_m^*(z)) \).
2. At \((i,j)\)-renewal transitions, there are group arrivals with probability density \( \{\Phi_{i,j}(k), k \geq 0\} \) whose p.g.f. is \( \Psi_{i,j}(z) \). Let us denote the \( m \times m \) matrix \( \Psi_{i,j}(z) \) by \( \Psi(z) \).
3. At \((i,j)\)-transitions \((i \neq j)\), there are group arrivals with probability densities \( \{\Omega_{i,j}(k), k \geq 0\} \), whose p.g.f. is \( \Omega_{i,j}(z) \). For notational convenience, we set \( \Omega_{i,i}(z) \equiv 1 \) for all \( i \) and define \( \Omega_{i,j}(z) \leq i, j \leq m \) by \( \Omega(z) \).

Define the conditional probabilities

\[
P_{i,j}(n,t) = Pr(J(t) = j, N(t) = n \mid J(0) = i, N(0) = 0),
\]

where \( N(t) \) and \( J(t) \) denote the number of arrivals during \((0,t]\) and the state of the underlying Markov process \( Q^* \) at time \( t \), respectively. We also define conditional probability matrices \( P(n,t) = (P_{i,j}(n,t))_{1 \leq i,j < m, n \geq 0} \). It was shown in [4] that

\[
\sum_{n=0}^{\infty} z^n P(n,t) = exp(R(z)t), 0 \leq z \leq 1,
\]
with $\mathbf{R}(z) = \sum_{n=0}^{\infty} z^n \mathbf{R}_n$ and

\[
\begin{align*}
\mathbf{R}_0 &= \mathbf{A} \Phi(0) - \mathbf{A} + \mathbf{A} \Phi \circ \mathbf{A} \Phi(0) + \mathbf{A} \Phi(0) \\
\mathbf{R}_n &= \mathbf{A} \Phi(n) + \mathbf{A} \Phi \circ \mathbf{A} \Phi(n) + \mathbf{A} \Phi(n), \quad n \geq 1,
\end{align*}
\]

where $\circ$ denotes the Schur (entrywise) product of two matrices. For the upcoming analysis, we shall assume that the matrix $\mathbf{R}_0^{-1}$ exists.

3. Analysis of an N/G/1 Finite Queue

In this section, we will analyze the N/G/1 finite queue with the supplementary variable method. The queue size is assumed to be $K$. When describing the N-process, we will use the same notations as in Section 2. The successive service times are independent and identically distributed according to $\mathbf{H}(x)$. Also the hazard rate function and the mean of $\mathbf{H}(x)$ are denoted by $r(x)$ and $\mu$ respectively.

3.1 Supplementary Variable Method

Let $X(t)$ denote the number of customers in the system at time $t$. We define the elapsed service time $S(t)$ as follows: If $X(t) > 0$, $S(t)$ denotes the amount of service already received by a customer in service. Otherwise, $S(t)$ denotes the amount of time elapsed after the last service completion. Then, the triplet $(J(t), X(t), S(t))$ is a three-dimensional Markov process with state space $\{1, \ldots, m\} \times \{0, \ldots, K\} \times [0, \infty)$.

Suppose that

\[
\pi(i, n, x)dx = \lim_{t \to \infty} Pr(J(t) = i, X(t) = n, x \leq S(t) < x + dx)
\]

exists for all states and define $\pi(n, x) = (\pi(1, n, x), \ldots, \pi(m, n, x))$. Then the Kolmogorov differential equations of the joint density $\pi(n, x)$ can be written down as follows:

\[
\frac{d}{dx} \pi(0, x) = \pi(0, x) \mathbf{R}_0,
\]

\[
\frac{d}{dx} \pi(n, x) = -\pi(n, x)r(x) + \sum_{k=1}^{n} \pi(k, x) \mathbf{R}_{n-k}, \quad 0 < n < K,
\]

\[
\frac{d}{dx} \pi(K, x) = -\pi(K, x)r(x) + \sum_{k=1}^{K} \sum_{l=K-k}^{\infty} \pi(k, x) \mathbf{R}_{l}.
\]

The joint density $\pi(n, x)$ should satisfy the boundary conditions

\[
\pi(0, 0) = \int_{0}^{\infty} \pi(1, x)r(x)dx,
\]
\[ \pi(n,0) = \int_0^\infty \pi(n+1,x)r(x)dx + \int_0^\infty \pi(0,x)R_n dx, \quad 0 < n < K, \quad (6) \]

\[ \pi(K,0) = \sum_{i=0}^\infty \int_0^\infty \pi(0,x)R_i dx, \quad (7) \]

and the normalization condition

\[ \sum_{n=0}^K \int_0^\infty \pi(n,x)dx = 1, \quad (8) \]

where \( e = (1,\ldots,1)^t \).

Now, we shall find the joint density \( \pi(n,x) \) of the queue length and the elapsed service time. From equation (1), we obtain

\[ \frac{d}{dx}P(n,x) = \sum_{k=0}^n P(k,x)R_{n-k}, \quad n \geq 0. \]

With this and equations (2)-(8), we get

\[ \pi(0,x) = \pi(0,0)P(0,x), \quad (9) \]
\[ \pi(n,x) = \sum_{k=1}^n \pi(k,0)P(n-k,x)(1-H(x)), \quad 0 < n < K, \quad (10) \]
\[ \pi(K,x) = \sum_{k=1}^K \sum_{i=K-k}^\infty \pi(k,0)P(i,x)(1-H(x)). \quad (11) \]

We may also derive the above solutions by conditioning on the state of the system time \( x \) back.

Before finding the coefficients \( \pi(n,0) \), we consider the embedded Markov chain \( \{J(\tau_n),X(\tau_n)\} \), where \( \{\tau_n \geq 0\} \) is the \( n^{th} \) epoch of service or idle completion. Then the transition probability matrix of \( \{J(\tau_n),X(\tau_n)\} \) is

\[
Q_E = \begin{pmatrix}
0 & U_1 & \cdots & U_{K-2} & U_{K-1} & \sum_{n=K}^\infty U_n \\
A_0 & A_1 & \cdots & A_{K-2} & \sum_{n=K-1}^\infty A_n & 0 \\
0 & A_0 & \cdots & A_{K-3} & \sum_{n=K-2}^\infty A_n & 0 \\
& \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & 0 & \sum_{n=0}^\infty A_n & 0
\end{pmatrix}
\]

where

\[ U_n = \int_0^\infty P(0,x)R_n dx = [-R_0^{-1}]R_n, \quad n \geq 1. \]
The matrix $U_n$ (or $A_n$) is the probability that $n$ customers arrive during an idle time (or a service time).

**Theorem 1:** The coefficients $\pi(n,0)$ of the joint density $\pi(n,x)$ are given by

\[
(\pi(0,0),\ldots,\pi(K,0)) = \frac{1}{\mu - x_0[\mu I + R_0^{-1}(0)]e}(x_0,\ldots,x_K),
\]

where $(x_0,\ldots,x_K)$ is the stationary vector of the transition probability matrix $Q_E$. $I$ is an identity matrix of size $m$.

**Proof:** By inserting (9)-(11) into the boundary conditions, we show that $(\pi(0,0),\ldots,\pi(K,0))$ is a positive invariant vector of $Q_E$, that is,

\[
(\pi(0,0),\ldots,\pi(K,0)) = c(x_0, x_K)
\]

for some constant $c > 0$.

Applying (9)-(11) to the normalization condition, we have

\[
\pi(0,0)[-R_0^{-1}]e + \mu \sum_{n=1}^{\infty} \pi(n,0)e = 1.
\]

Therefore, we have

\[
c = \frac{1}{\mu - x_0[\mu I + R_0^{-1}]e}.
\]

So the proof is complete.

The matrices $P(n,x)$ and $A_n$ can be efficiently evaluated by means of an iterative procedure in [3]. Therefore, we can compute $\pi(n,x)$ by deriving the stationary vector of the transition probability matrix $Q_E$. As Blondia [1] did, we can also reduce the complexity of the computation for the stationary vector with the Schur-Banachiewicz formula for the inverse of block matrices.

### 3.2 Queue Length Distribution

In this subsection, we shall consider two computational algorithms to obtain the queue length distribution $\pi(n)$ using the coefficients $\pi(n,0)$ derived in the previous subsection.

Let us define

\[
M_n = \int_0^{\infty} P(n,x)(1 - H(x))dx, \ n \geq 0
\]

and let $M(z)$ be the generating function of $\{M_n, n \geq 0\}$. Then equations (9)-(10) yield

\[
\pi(0) = \pi(0,0)[-R_0^{-1}],
\]

(12)

\[
\pi(n) = \sum_{k=1}^{n} \pi(k,0)M_{n-k}, \ 1 \leq n \leq K-1.
\]
Since $\sum_{n=0}^{K} \pi(n)$ is the stationary vector $\theta$ of the underlying Markov process $Q^*$, we get

$$\pi(K) = \theta - \sum_{n=0}^{K-1} \pi(n). \quad (13)$$

Using the fact that $(\pi(0,0), \ldots, \pi(K,0))$ is an invariant vector of the transition probability matrix $Q_E$ and $A(z) = M(z)R(z) + I$, we have

$$\pi(n) = \left( \sum_{k=1}^{n-1} \pi(k)R_{n-k} = \pi(0)R_{n-1} \right) \left[ -R_0^{-1} \right]$$

$$- (1_{n \geq 2})[\pi(n-1,0) - \pi(n,0)] \left[ -R_0^{-1} \right], \quad 1 \leq n \leq K - 1, \quad (14)$$

where $1_{(\cdot)}$ is the indicator function.

The above equation (14) for the queue length distribution has a simpler form than on derived by Blondia [1], since Blondia’s formulas require additional computation of matrices $\{R_n(s), n \geq 0\}$ satisfying $[R(z) + sI]^{-1} = \sum_{n=0}^{\infty} R_n(s)z^n$. Consequently, we can obtain a more efficient computational algorithm for the queue length distribution.

References