On the finite-source $\tilde{G}/M/r$ queue

J. SZTRIK
Institutum Mathematicum, Universitatis Debreceniensis, H-4010 Debrecen Pf 12, Hungaria

Abstract. The aim of the present paper is to give the main characteristics of the finite-source $\tilde{G}/M/r$ queue in equilibrium. Here unit $i$ stays in the source for a random time having general distribution function $F_i(x)$ with density $f_i(x)$. The service times of all units are assumed to be identically and exponentially distributed random variables with means $1/\mu$. It is shown that the solution to this $\tilde{G}/M/r$ model is similar in most important respects to that for the $M/M/r$ model.

Keywords: Source time, service time, utilization, waiting times, Little's formula

1. Introduction

Units emanate from a finite source of size $n$ and are served by one of $r$ ($r \leq n$) servers at a service facility according to a first-come first-served (FCFS) discipline. If there is no idle server, then a waiting line is formed and the units are delayed. The service times of the units are supposed to be identically and exponentially distributed random variables with means $1/\mu$. After completing service, unit $i$ returns to the source and stays there for a random time having general distribution function $F_i(x)$ with density $f_i(x)$. All random variables are assumed to be independent of each other.

Such a finite-source queueing model is often called the 'machine interference problem.' In recent years this mode has been effectively used, for example, for the mathematical description of multiprogramming computer systems. Since there is a sizable literature on machine interference, we refer only to the most related results. The first methods were due to Benson and Cox [2], who obtained a steady state solution for the $M/M/r$ model. Using similar methods, Benson [1] treated the $E_k/M/1$ model and showed that the stationary result was identical to that for the $M/M/1$ model. Maritas and Xirokostas [13] discussed the $M/E_k/r$ case and gave a numerical solution to it. Bunday and Scraton [3] have recently proved that the probability distribution of the number of machines running in steady state is the same in the $M/M/r$ and $G/M/r$ cases. A fairly extensive set of tables, based on the $M/M/r$ case, has been calculated by Peck and Hazelwood [14].

For the interested reader the following books and papers are recommended: Cox and Smith [5], Jaiswal [9], Kleinrock [12], Takács [19], Gaver [7], Ferdinand [6], Kameda [10], Karmeshu and Jaiswal [11], Schatte [15], Sztrik [16] and Tomkó [17]. This paper deals with a generalization of the $G/M/r$ model and gives the main steady-state characteristics of the $G/M/r$ queue.

2. The mathematical model

Let the random variable $\nu(t)$ denote the number of units staying in the source at time $t$ and $(\alpha_i(t), \ldots, \alpha_{\nu(t)}(t))$ indicate their indices ordered lexicographically. Let us denote by $(\beta_1(t), \ldots, \beta_{n-\nu(t)}(t))$ the

I am very grateful to Dr. J. Tomkó for his helpful discussion. My special thanks are also due to the referees for providing valuable comments and very thorough readings of earlier versions of the paper, which have greatly improved the presentation.

Received October 1983; revised April 1984

North-Holland

indices of the units waiting or served at the service facility in the order of their arrival. Clearly the sets \( \{\alpha(t), \ldots, \alpha(v(t))\} \) and \( \{\beta(t), \ldots, \beta_{n-v(t)}(t)\} \) are disjoint.

Introduce the process
\[
Y(t) = \{v(t); \alpha_1(t), \ldots, \alpha_{v(t)}(t); \beta_1(t), \ldots, \beta_{n-v(t)}(t)\}.
\]
The stochastic process \( Y(t), t \geq 0 \) is not Markovian unless the distribution functions \( F_i(x) \) are exponential, \( i = 1, \ldots, n \).

Let us also introduce the supplementary variables \( \alpha_i(t) \) to denote the random time that unit \( \alpha_i(t) \) has been spending in the source until time \( t \), \( i = 1, \ldots, n \). Define
\[
x(t) = \{v(t); \alpha_1(t), \ldots, \alpha_{v(t)}(t); \xi_{\alpha_i(t)}, \ldots, \xi_{\alpha_{v(t)}(t)}(t); \beta_1(t), \ldots, \beta_{n-v(t)}(t)\}.
\]
Then the process \( X(t), t \geq 0 \) has the Markov property.

Let \( V^n_k \) and \( C^n_k \) denote the set of all variations and combinations of order \( k \) of the integers 1, 2, \ldots, \( n \) respectively, ordered lexicographically. Then the state space of the process \( x(t) \) consists of the sets
\[
(i_1, \ldots, i_k; x_1, \ldots, x_k; j_1, \ldots, j_{n-k}), (i_1, \ldots, i_k) \in C^n_k, (j_1, \ldots, j_{n-k}) \in V^n_{n-k}, x_i \in \mathbb{R}_+,
\]
i = 1, 2, \ldots, n.

The process is in state \( (i_1, \ldots, i_k; x_1, \ldots, x_k; j_1, \ldots, j_{n-k}) \) if \( k \) units with indices \( (i_1, \ldots, i_k) \) have been staying in the source for times \( (x_1, \ldots, x_k) \) respectively, while the rest need service and their indices in order of arrival are \( (j_1, \ldots, j_{n-k}) \).

To derive the Kolmogorov equations we should consider the transitions that can occur in an arbitrary time interval \( (t, t + h) \). The transition probabilities are then the following:
\[
P\{x(t + h) = (i_1, \ldots, i_k; x_1 + h, \ldots, x_k + h; j_1, \ldots, j_{n-k}) / x(t) = (i_1, \ldots, i_k; x_1, \ldots, x_k; j_1, \ldots, j_{n-k})\} = [1 - (n - k)\mu h] \prod_{i=1}^{k} \frac{1 - F_i(x_i + h)}{1 - F_i(x_i)} + o(h),
\]
\[
P\{x(t + h) = (i_1, \ldots, i_k; x_1 + h, \ldots, x_k + h; j_1, \ldots, j_{n-k}) / x(t) = (i_1, \ldots, i_k, i'; x_1, \ldots, x_k; j_1, \ldots, j_{n-k-1})\} = \frac{f_i(y)h}{1 - F_i(y)} \prod_{i=1}^{k} \frac{1 - F_i(x_i + h)}{1 - F_i(x_i)} + o(h),
\]
where \( (i', \ldots, i'_{n-k}, \ldots, i'_k) \) denotes the lexicographical order of indices \( (i_1, \ldots, i_k, j_{n-k}) \), while \( (x_1', \ldots, y', \ldots, x_k') \) indicates the corresponding times, \( 0 \leq n - k < r \). For \( r \leq n - k \leq n \)
\[
P\{x(t + h) = (i_1, \ldots, i_k; x_1 + h, \ldots, x_k + h; j_1, \ldots, j_{n-k}) / x(t) = (i_1, \ldots, i_k; x_1, \ldots, x_k; j_1, \ldots, j_{n-k})\} = [1 - r\mu h] \prod_{i=1}^{k} \frac{1 - F_i(x_i + h)}{1 - F_i(x_i)} + o(h),
\]
\[
P\{x(t + h) = (i_1, \ldots, i_k; x_1 + h, \ldots, x_k + h; j_1, \ldots, j_{n-k}) / x(t) = (i', \ldots, i'_{n-k}, \ldots, i'_k; x_1, \ldots, y', \ldots, x_k; j_1, \ldots, j_{n-k-1})\} = \frac{f_i(y)h}{1 - F_i(y)} \prod_{i=1}^{k} \frac{1 - F_i(x_i + h)}{1 - F_i(x_i)} + o(h).
\]
For the distribution of $x(t)$ consider the following functions:

$$Q_{0; j_1, \ldots, j_n}(t) = P \{ v(t) = 0; \beta_1(t) = j_1, \ldots, \beta_n(t) = j_n \},$$

$$Q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k; t) =$$

$$P \{ v(t) = k; \alpha_1(t) = i_1, \ldots, \alpha_k(t) = i_k;$$

$$\xi_{i_1} < x_1, \ldots, \xi_{i_k} < x_k; \beta_1(t) = j_1, \ldots, \beta_{n-k}(t) = j_{n-k} \}.$$  (2.1)

Let $\lambda_i$ defined by $1/\lambda_i = \int_0^\infty x \, dF_i(x)$. Then we have

**Theorem 1.** If $1/\lambda_i < \infty$, $i = 1, \ldots, n$, then the process $(x(t), t \geq 0)$ possesses a unique limiting (stationary) ergodic distribution independent of the initial conditions, namely

$$Q_{0; j_1, \ldots, j_n} = \lim_{t \to \infty} Q_{0; j_1, \ldots, j_n}(t),$$

$$Q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k) = \lim_{t \to \infty} Q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k; t).$$  (2.2)

**Proof.** Notice that $x(t)$ belongs to the class of piecewise-linear Markov processes, subject to discontinuous changes treated by Gnedenko-Kovalenko [8] in detail. Our statement follows from a theorem on page 211 of this monograph. Furthermore we have

**Corollary 1.** The process $x(t)$ is bounded in state probabilities.

Since the ergodic distributions (2.2) have densities $q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k)$, if $F_i(x)$ possess densities $f_i(x)$, $i = 1, \ldots, n$ (c.f. Gnedenko-Kovalenko [8, p. 224]), Theorem 1 provides the existence and uniqueness of the following limits:

$$q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k) \, dx_1 \cdots dx_k$$

$$= \lim_{t \to \infty} P \{ v(t) = k; \alpha_1(t) = i_1, \ldots, \alpha_k(t) = i_k; x_l < \xi_{i_l} < x_{l+1} + dx_l, l = 1, \ldots, k;$$

$$\beta_1(t) = j_1, \ldots, \beta_{n-k}(t) = j_{n-k} \}, \text{ for } k = 1, \ldots, n.$$  (2.4)

However, the assumption concerning the densities $f_i(x)$ is only made to simplify the derivation of state equations, as the same result can be obtained by replacement of the densities $q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k)$ by the differentials of the corresponding distributions.

In order to formulate the following theorem introduce a further notation, namely

$$q^*_n (x_1, \ldots, x_k) = \frac{q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k)}{(1 - F_i(x_1)) \cdots (1 - F_i(x_k))}$$  (2.4),

which is the so-called normed density function, $k = 1, \ldots, n$. Then we have

**Theorem 2.** The normed density functions satisfy the following system of integro-differential equations (2.5), (2.7) with boundary conditions (2.6), (2.8):

$$\left[ \frac{\partial}{\partial x_1} + \cdots + \frac{\partial}{\partial x_k} \right]^* q^*_n (x_1, \ldots, x_k)$$

$$= -(n-k) \mu q^*_n (x_1, \ldots, x_k)$$

$$+ \int_0^\infty q^*_n (x_1', \ldots, y', \ldots, x_k') f_{j_{n-k}}(y) \, dy,$$  (2.5)
\begin{equation}
q_{i_1, \ldots, i_k, j_1, \ldots, j_n}(x_1 + h, \ldots, x_k + h) = q_{i_1, \ldots, i_k, j_1, \ldots, j_n}(x_1, \ldots, x_k),
(1 - (n - k) \mu h) \prod_{l=1}^{k} \frac{1 - F_{i_l}(x_l + h)}{1 - F_{i_l}(x_l)} + \prod_{l=1}^{k} \frac{1 - F_{j_l}(x_l + h)}{1 - F_{j_l}(x_l)}
\times \int_{0}^{\infty} q_{i_1, \ldots, i_{l-1}, x_{l+1}, \ldots, i_k, j_1, \ldots, j_{n-l}}(x_1', \ldots, y', \ldots, x_k') \frac{f_{j_n}(y) h}{1 - F_{j_n}(y)} dy + o(h),
\end{equation}

\begin{equation}
q_{i_1, \ldots, i_k, j_1, \ldots, j_n}(x_1 + h, \ldots, x_{l-1} + h, 0, x_{l+1} + h, \ldots, x_k + h) = \mu h \sum_{i_1, \ldots, i_k, j_1, \ldots, j_n} q_{i_1, \ldots, i_{l-1}, 1, i_{l+1}, \ldots, i_k, j_1, \ldots, j_n}(x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_k)
\times \prod_{s=1, s \neq l}^{k} \frac{1 - F_{i_s}(x_s + h)}{1 - F_{i_s}(x_s)} + o(h)
\end{equation}

for \( l = 1, \ldots, k, 0 \leq n - k < r. \)

Similarly
\begin{equation}
q_{i_1, \ldots, i_k, j_1, \ldots, j_n}(x_1 + h, \ldots, x_k + h)
= q_{i_1, \ldots, i_k, j_1, \ldots, j_n}(x_1, \ldots, x_k)(1 - r \mu h) \prod_{l=1}^{k} \frac{1 - F_{i_l}(x_l + h)}{1 - F_{i_l}(x_l)} + \prod_{l=1}^{k} \frac{1 - F_{j_l}(x_l + h)}{1 - F_{j_l}(x_l)}.
\end{equation}
\[
\times \int_0^\infty q_{i_1, \ldots, i_{n-k-1}, i_{n-k} : j_1, \ldots, j_{n-k-1}}(x_1', \ldots, y', \ldots, x_k') f_{i_{n-k}}(y) \frac{h \, dy}{1 - F_{i_{n-k}}(y)} + o(h), \tag{2.10}
\]
\[
q_{i_1, \ldots, i_{k}, j_1, \ldots, j_{n-k}}(x_1 + h, \ldots, x_{l-1} + h, 0, x_{l+1} + h, \ldots, x_k + h) h
\]
\[
= \mu h \sum_{y_{i_{n-k+1}, \ldots, i_{n}}} q_{i_1, \ldots, i_{n-k+1}, i_{n-k+1}, j_1, \ldots, j_{n-k}}(x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_k)
\times \prod_{s=1, s \neq l}^k \frac{1 - F_j(x_s + h)}{1 - F_j(x_s)} + o(h)
\]
for \( l = 1, \ldots, k, r \leq n - k < n. \)

Finally
\[
Q_{0; i_1, \ldots, i_n} = Q_{0; i_1, \ldots, i_n}(1 - r \mu h) + \int_0^\infty q_{i_1, \ldots, i_n}(y) \frac{f_\mu(y) h}{1 - F_\mu(y)} \, dy + o(h). \tag{2.11}
\]

Hence the derivation of equations (2.5), (2.7) and boundary conditions (2.6), (2.8) is quite simple. Dividing the left-hand side of equations (2.9) (2.11) by factor \( \prod_{s=1, s \neq l}^k [1 - F_j(x_s + h)] \), taking the limits as \( h \to 0 \) and taking into account the definition of the normed densities (2.4) we get the desired result.

In the left-hand side of (2.5), (2.7), used for the notation of the limit in the right-hand side, the usual notation for partial differential quotients has been applied. Strictly considering this is not allowed, since the existence of the individual partial differential quotient is not assured. This is why the operator is notated by \( \left[ \right]^* \). Actually this is a \( (1, 1, \ldots, 1) \in \mathbb{R}^k \) directional derivative. (See Cohen [4, pp. 252].)

In the following we solve equations (2.5), (2.7) subject to boundary conditions (2.6), (2.8) to determine the ergodic probabilities

\[
(Q_{0; i_1, \ldots, i_n}, Q_{1, \ldots, j_1, \ldots, j_{n-k}}), (i_1, \ldots, i_k) \in C_k^n, (j_1, \ldots, j_{n-k}) \in \nu_n^{n-k}, k = 1, \ldots, n.
\]

If we set
\[
Q_{0; i_1, \ldots, i_n} = C_0, \quad q^*_{i_1, \ldots, i_k, j_1, \ldots, j_{n-k}}(x_1, \ldots, x_k) = C_k,
\]
then they satisfy the equations (2.5), (2.7) and boundary conditions (2.6), (2.8); moreover, it is easy to see that

\[
C_k = (r! r^{n-r-k} \mu^{n-k})^{-1} C_n \quad \text{for} \ 0 \leq k \leq n - r,
\]
and, similarly, that

\[
C_k = ((n - k)! \mu^{n-k})^{-1} C_n \quad \text{for} \ n - k \leq k \leq n.
\]

Since these equations completely describe the system, this is the required solution.

Let \( Q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}} \) denote the steady state probability that units with indices \( (i_1, \ldots, i_k) \) are in the source and the order of arrival of the rest to the service facility is \( (j_1, \ldots, j_{n-k}) \). Furthermore, denote by \( Q_{i_1, \ldots, i_k} \) the steady state probability that units with indices \( (i_1, \ldots, i_k) \) are staying in the source. It can be readily verified that

\[
Q_{i_1, \ldots, i_k; j_1, \ldots, j_{n-k}} = (\lambda_{i_1} \ldots \lambda_{i_k})^{-1} C_k \quad \text{for} \ k = 1, \ldots, n. \tag{2.12}
\]

Using the relation we got for \( C_k \) we obtain

\[
Q_{i_1, \ldots, i_k} = (n - k)! [r! r^{n-r-k} \mu^{n-k} \lambda_{i_1} \ldots \lambda_{i_k}]^{-1} C_n. \tag{2.13}
\]

\[(i_1, \ldots, i_k) \in C_k^n, k = 0, 1, \ldots, n - r.\]
Similarly

\[ Q_{i_1, \ldots, i_k} = \left[ \mu^{n-k} \cdot \lambda_{i_1} \cdots \lambda_{i_k} \right]^{-1} C_n, \quad (i_1, \ldots, i_k) \in C_n, \quad k = n-r, \ldots, n. \]

Let \( \hat{Q}_k \) and \( \hat{P}_k \) denote the steady state probabilities that \( k \) units are staying in the source and \( l \) units are at the service facility, respectively. Clearly

\[ Q_{i_1, \ldots, i_n} = Q_1, \ldots, n = \hat{Q}_n = \hat{P}_0, \quad \hat{Q}_k = \hat{P}_{n-k} \quad \text{for} \quad k = 0, 1, \ldots, n. \]

It is easy to see that

\[ C_n = \hat{Q}_n(\lambda_1 \cdots \lambda_n) \quad \text{and} \quad \hat{Q}_k = \sum_{(i_1, \ldots, i_k) \in C^n_k} Q_{i_1, \ldots, i_k}, \]

where \( \hat{Q}_n \) can be obtained with the aid of the norming condition

\[ \sum_{k=0}^{n} \hat{Q}_k = 1. \]

In the homogeneous case, relations (2.13), (2.14) yield

\[ \hat{Q}_k = \frac{n!}{k!r!(n-r-k)!} \left( \frac{\lambda}{\mu} \right)^{n-k} \hat{Q}_n \quad \text{for} \quad 0 \leq k \leq n-r, \]
\[ \hat{Q}_k = \binom{n}{k} \left( \frac{\lambda}{\mu} \right)^{n-k} \hat{Q}_n \quad \text{for} \quad n-k \leq k \leq n. \]

Thus, the probability that \( k \) units are not in the source is

\[ \hat{P}_k = \binom{n}{k} \left( \frac{\lambda}{\mu} \right)^{k} \hat{P}_0 \quad \text{for} \quad 0 \leq k \leq r, \]
\[ \hat{P}_k = \frac{n!}{(n-k)!r!(k-r)!} \left( \frac{\lambda}{\mu} \right)^{k} \hat{P}_0 \quad \text{for} \quad r \leq k \leq n. \]

This is exactly the same result as that obtained by Bunday and Scraton [3]. The equivalence of the finite-source \( E_k/M/1 \) to the \( M/M/1 \) and in addition to that of the \( G/M/r \) to the \( M/M/r \) model as noted by Benson, and Bunday and Scraton, respectively, are just special cases of the more general result obtained here.

Before determining the main characteristics of the system we need one more theorem. In order to formulate this theorem we introduce some further notation. Let \( Q^{(i)}(P^{(i)}) \) denote the steady state probability that unit \( i \) is in the source (at the service facility) for \( i = 1, \ldots, n \). It is clear that the process \( (Y(t), t \geq 0) \) is semi-Markovian with state space

\[ \bigcup_{(i_1, \ldots, i_k) \in C^n_k; (j_1, \ldots, j_{n-k}) \in P^n_{n-k}} \{(i_1, \ldots, i_k; j_1, \ldots, j_{n-k})\}. \]

Let \( H_i \) be the event that unit \( i \) is in the source and \( Z_{H_i}(t) \) its characteristic function; i.e.

\[ Z_{H_i}(t) = \begin{cases} 1 & \text{if} \quad Y(t) \in H_i, \\ 0 & \text{otherwise}. \end{cases} \]

Then we have
Theorem 4.

\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T Z_{H_i}(t) \, dt = \frac{1/\lambda_i}{1/\lambda_i + W_i + 1/\mu} = Q^{(i)} = 1 - p^{(i)},
\]

where \( W_i \) denotes the mean waiting time of unit \( i \).

**Proof.** The statement is a special case of a theorem concerning the expected sojourn time for semi-Markov processes, see Tomkó [18, pp. 297].

Sometimes we need the long-run fraction of time the unit \( i \) spends in the source. This happens, e.g., in the 'machine interference model'. In that case for the utilization of machine \( i \) we have

\[
U_i = Q^{(i)} = \sum_{k=1}^{n} \sum_{i \in (i_1, \ldots, i_k) \in C^{(i)}} Q_{i_1, \ldots, i_k}.
\]

3. The main performance measures

(i) **Utilizations**

Utilizations can now be considered for individual servers or for the system as a whole. The process \((x(t), t \geq 0)\) is assumed to be in equilibrium. Considering the system as a whole, it will be empty only when there are no units at the service facility and will be busy at other times. As usual, using renewal-theoretic arguments for the system utilization, we have

\[
U = 1 - \hat{Q}_n \quad \text{and} \quad \hat{Q}_n = \frac{M_{n^*}}{M_{n^*} + M\delta},
\]

where \( n^* = \min(n_1, \ldots, n_n) \), random variable \( \eta_i \) denotes the source-time of the unit \( i \), \( i = 1, \ldots, n \), and \( M\delta \) denotes the average busy period of the system.

Thus the expected length of the busy period is given by

\[
M\delta = M_{n^*} \frac{1 - \hat{Q}_\eta}{\hat{Q}_\eta}.
\]

In particular, if \( F_i(x) = 1 - \exp(-\lambda_i x), i = 1, \ldots, n \), we get

\[
M\delta = \frac{1 - \hat{Q}_\eta}{\hat{Q}_\eta} \frac{1}{\Sigma \lambda_i}.
\]

It is also easy to see that for server utilization the following relation holds:

\[
U_i = \frac{1}{r} \left( \sum_{k=1}^{r} k \hat{P}_k + r \sum_{k=r+1}^{n} \hat{P}_k \right) = \frac{\bar{r}}{r},
\]

where \( \bar{r} \) denotes the mean number of busy servers.

(ii) **Mean waiting times**

By the virtue of Theorem 4 we obtain \( q^{(i)} = (1 + \lambda_i W_i + \lambda_i / \mu)^{-1} \). Consequently, the average waiting time of unit \( i \) is

\[
W_i = (1 - Q^{(i)})(\lambda_i Q^{(i)})^{-1} - 1/\mu.
\]

It follows that the mean sojourn time of unit \( i \), that is, the waiting and service time together, can be
obtained by
\[ T_i = W_i + 1/\mu = (1 - Q^{(i)})(\lambda_i Q^{(i)})^{-1} \quad \text{for } i = 1, \ldots, n. \]

Since \( \sum_{i=1}^{n} (1 - Q^{(i)}) = \bar{n} \), where \( \bar{n} \) denotes the mean number of units staying at the service facility we have, by reordering and adding (3.1)
\[ \sum_{i=1}^{n} \lambda_i T_i Q^{(i)} = \bar{n}. \]

This is Little's formula for the finite-source \( \bar{G}/M/r \) queue. In particular, if \( F_i(x) = F(x), i = 1, \ldots, n \), (3.2) can be written as \( \lambda \bar{Q} T = \bar{n} \), where \( \bar{Q} \) denotes the expected number of units staying in the source.

References