AD-A090 538 DELAWARE UNIV NEWARK APPLIED MATHEMATICS INST F/G 12/1
A MULTI-SERVER QUEUE WITH THRESHOLDS FOR THE ACCEPTANCE OF CUST--ETC(U)
JUN 80 M F NEUTS, R NADARAJAN AFOSR-77-3236
UNCLASSIFIED TR-59B AFOSR-TR-80-0969 NL

1 of 1
AD
A090538

END
DATE
FILMED
11-80
DTIC

AFOSR-TR- 80-0969

LEVEL II

(3)

AD A090538

DTIC
ELECTE
OCT 17 1980
S
C

APPLIED
MATHEMATICS INSTITUTE

University of Delaware
Newark, Delaware

DDC FILE COPY

Approved for public release;
distribution unlimited.

80 10 9 105

A MULTI-SERVER QUEUE WITH THRESHOLDS FOR THE ACCEPTANCE OF CUSTOMERS INTO SERVICE

by

Marcel F. Neuts
University of Delaware

and

R. Nadarajan
Annamalai University

Department of Mathematical Sciences
University of Delaware
Newark, DE 19711
U.S.A.

Department of Mathematics and Statistics
Annamalai University
Annamalainagar
India 608101

Applied Mathematics Institute
Technical Report No. 59B
June 1980

This research was supported in part by the National Science Foundation under Grant No. ENG-7908351 and the Air Force Office of Scientific Research under Grant No. AFOSR-77-3236.

ABSTRACT

A queue with Poisson arrivals is served by N identical exponential servers. When a server becomes free, he can serve a group of customers of size at most b. He is not allowed to process a group of size less than a, $1 \leq a \leq b$. The actual size of each group served depends on the number of waiting customers.

The queue is studied as a Markov process. In the stable case, the stationary probability vector of this process has a simple, readily computable form. Using this form, the stationary waiting time distribution, which is analytically intractable, may be expressed in an algorithmically useful form.

Several questions, related to the optimal design of such a service system, may be algorithmically investigated. This model serves to illustrate the advantages of the algorithmic approach. The design criteria are not analytically tractable in general.

KEY WORDS

Multi-server queues, group service, optimal design of queues, computational probability.

1. Introduction

The purpose of this paper is to show the advantages of the algorithmic approach in studying the optimal design problems for a simple, but useful queueing model. Customers arrive at a fixed location according to a Poisson process of rate λ. There are N servers, who process customers in groups. When a free server is at the fixed location and finds $i \geq a \geq 1$ customers there, he removes min (i,b) customers. The parameter b denotes the maximum allowable size of a group to be served and satisfies $b \geq a$. If there are fewer than a customers present, all free servers remain at the location until a customers are present. One server then leaves with a group of a customers.

Services are thought of as the removal of groups of customers to a different location. The times between the departures and subsequent returns of a server are assumed to have an exponential distribution of parameter μ. All service times are mutually independent of each other and also of the interarrival times of customers.

There are a number of practical situations, which motivated the study of this queueing model. The most familiar is a taxi stand served by a fleet of N taxis. The customers arrive at the stand and are removed in groups of at least a and at most b. The quantity $1/\mu$ is then the mean travel time of a taxi on a single trip.

In a situation of greater interest, the N servers are locomotives which remove wagon loads of ore from a mine head to other locations. The parameter λ is now the rate at which wagon loads of ore (customers) become available at the station. The parameters a and b are respectively the minimum and maximum allowable numbers of wagons per train. One may also view the N servers as medical evacuation helicopters removing casualties from a field hospital to a more remote facility. The significance of each of the five parameters of the model is then again evident.

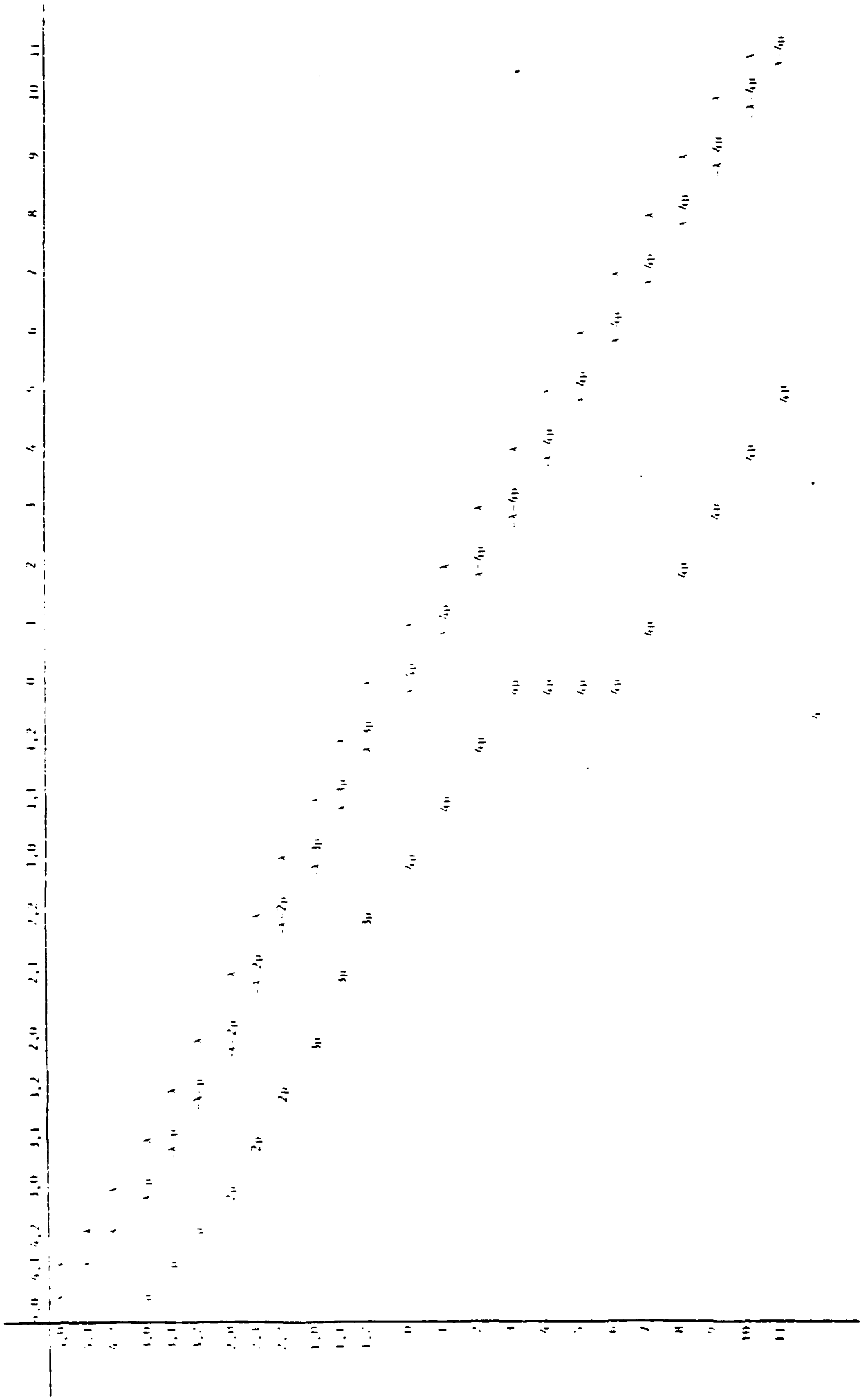
The design problems are clear. It is desirable to choose N sufficiently large that few customers will be waiting for service, yet not so large that servers spend an inordinate amount of time waiting idly for customers. In applications, the parameter b is usually fixed by technological constraints. When b is large, it may be advisable to choose the parameter a considerably smaller than b, in order to expedite service. It is clearly of interest to be able to assess the influence of the choice of a on the behavior of the service system. Our solution procedure is such that it can handle problems with fairly large values of N and b, say on the order of twenty and two hundred respectively.

This queueing model may be denoted by the symbols M/M(a,b)/N. It is not essential that service times be viewed as travel times. There may be N servers who remain fixed, but serve customers according to the rule specified by the threshold values a and b. Our results will then provide the stationary distribution of the number of <u>waiting</u> customers and not of the usual queue length. There are a number of particular cases which were previously treated. For $N = 1$, the model is discussed for general service time distributions, dependent on the batch size, in Neuts [1,2]. A general procedure to analyze the queue GI/PH(a,b)/1 is described in Chapter 4 of [3]. The present model is also an instance of a double-ended queue. It adds to the literature on such queues, but avoids the discouragingly formal treatment, which hitherto has burdened their discussion.

The queue is studied as a Markov process on the state space $E = \{(i,j), 1 \le i \le N, 0 \le j \le a-1\} \cup \{k \ge 0\}$. The process is in the state (i,j), $1 \le i \le N$, $0 \le j \le a-1$, whenever there are i free servers and only j customers. The state $k \ge 0$, corresponds to the case where all servers are occupied and k customers are waiting. It is convenient to order the states (i,j) in lexicographic order with i in descending order. The states $k \ge 0$, are listed thereafter.

The structure of the generator Q of the Markov process is important to the analysis of the model. It is displayed for $N = 4$, $a = 3$, and $b = 6$. The general structure of Q is readily apparent from this particular case.

It is important to note that all columns labeled $k \ge 1$, contain the same elements i.e. $-\lambda - N\mu$ on the diagonal,



λ immediately above and $N\mu$ in the row labeled $k+b$. The upper left hand submatrix of dimensions $(N+1)a \times Na$ has the form

$$\begin{vmatrix} C & E & & & & \\ \mu I & C-\mu I & E & & & \\ & 2\mu I & C-2\mu I & E & & \\ & & \cdots & & & \\ & & & & (N-1)\mu I & C-(N-1)\mu I \\ & & & & 0 & N\mu I \end{vmatrix} ,$$

where the matrices C and E are square matrices of order a , given by

$$C = \begin{vmatrix} -\lambda & \lambda & & & \mathbf{0} \\ & -\lambda & \lambda & & \\ & & \cdots & & \\ & & & -\lambda & \lambda \\ \mathbf{0} & & & & -\lambda \end{vmatrix} , \quad E = \begin{vmatrix} & & \mathbf{0} & \\ & & & \\ \lambda & & & \end{vmatrix} .$$

The column, labeled 0, has the elements $-\lambda-N\mu$ on the diagonal, λ immediately above and the entry $N\mu$ in the rows labeled a, $a+1$, ..., b .

2. The Stationary Probability Vector

The condition for stability of the queue and the particular form of the stationary probability vector $\underline{x}$ of the generator Q follow immediately from the structure of the matrix Q.

Theorem 1.

The Markov process with generator Q is positive recurrent if and only if the equation

$$(1) \qquad \lambda - (\lambda + N\mu)\xi + N\mu\, \xi^{b+1} = 0,$$

has a root in the interval (0,1). This is the case if and only if

$$(2) \qquad \lambda < Nb\mu .$$

The root ξ is then unique. It is equal to the expected number of visits to the state $k+1$ between successive visits to the state k, for any of the states $k \geq 0$.

The vector $\underline{x}$, which is the unique solution to the system $\underline{x}Q = \underline{0}$, $\underline{x}\,\underline{e} = 1$, is written in the partitioned form $[\underline{x}_N, \ldots, \underline{x}_1, x_0, x_1, \ldots]$, where $\underline{x}_j$, $1 \leq j \leq N$, are a-vectors. We have that

$$(3) \qquad x_i = x_0\, \xi^i , \qquad \text{for } i \geq 0 .$$

The vectors $\underline{x}_j$, $1 \leq j \leq N$, and the quantity x_0 are obtained by solving the system of linear equations

$$\underline{x}_N C + \mu \underline{x}_{N-1} = \underline{0} ,$$

$$\underline{x}_{j+1} E + \underline{x}_j [C-(N-j)\mu I] + (N-j+1)\mu \underline{x}_{j-1} = \underline{0} ,$$

$$\text{for} \quad 2 \leq j < N .$$

(4) $$\underline{x}_2 E + \underline{x}_1 [C-(N-1)\mu I] + x_0 N\mu(1, \xi, \xi^2, \ldots, \xi^{a-1}) = \underline{0} ,$$

$$x_{1,a-1}\lambda - [\lambda + N\mu - N\mu\xi^a(1-\xi)^{-1}(1-\xi^{b-a+1})] x_0 = 0 ,$$

$$\sum_{j=1}^{N} \underline{x}_j \underline{e} + x_0(1-\xi)^{-1} = 1 .$$

This system has a (unique) positive solution.

Proof

The modified geometric form of the vector $\underline{x}$ and the equilibrium condition (2) follow from results in C. B. Winsten [4], also discussed in greater generality in M. F. Neuts [3]. By substitution of (3) into the first $Na+1$ steady-state equations and into the normalizing equation, we obtain the equations (4). It is readily verified that the coefficient matrix of the first $Na+1$ linear equations in (4) is an irreducible generator. The solution vector is therefore determined up to a multiplicative constant and all its components have the same sign. That vector is now uniquely determined by the normalizing equation.

Corollary 1.

The stationary marginal density $\{p_j\}$ of the number of waiting customers is given by

$$p_j = \sum_{\nu=1}^{N} x_{\nu j} + x_j , \qquad \text{for } 0 \le j \le a-1 ,$$

$$= x_j , \qquad \text{for } j \ge a .$$

The stationary probability r_j that j servers are free is given by

$$r_j = x_0 (1-\xi)^{-1}, \qquad \text{for } j = 0$$

$$= \underline{x}_j \underline{e} , \qquad \text{for } 1 \le j \le N.$$

The case $N=1$ may be solved explicitly in terms of ξ . Elementary manipulations lead to

$$x_{1j} = \frac{\mu}{\lambda} x_0 \frac{1-\xi^{j+1}}{1 - \xi} , \qquad \text{for } 0 \le j \le a-1 ,$$

where

$$x_0 = (1-\xi) \left[1 + \frac{\mu a}{\lambda} - \frac{\mu\xi}{\lambda} \frac{1-\xi^a}{1-\xi}\right]^{-1} .$$

For $a=1$, we obtain $\xi = \frac{\lambda}{\mu}$, $x_{10} = 1-\xi$, $x = \xi(1-\xi)$, which corresponds to the familiar M/M/1 queue. All comments on computational aspects in the sequel will be reserved for the case $N \geq 2$. The case $N=1$ is trivial.

3. The Waiting Time Distributions

In this section, we consider the probability $W_j(x)$, $a \leq j \leq b$, $x \geq 0$, that a customer arriving to the stationary queue at time $t=0$, enters service before or at time x and is served as part of a group of size j. Service is according to the first-come, first-served discipline. The analytic expressions for the mass-functions $W_j(\cdot)$ are exceedingly complicated. We shall show that, in contrast, their algorithmic characterization is straightforward.

To this end, we first consider a simple absorption time problem in a Markov process with $2b$ states. The states of this process are denoted by $1,2,\ldots,b$ and $1^*, 2^*, \ldots, b^*$. The states $a^*, \ldots, b^*$ are absorbing. The rows of the generator, which correspond to the labels $a^*, \ldots, b^*$, are therefore identically zero. The rows with labels $1, 2, \ldots, b$, $1^*, \ldots, (a-1)^*$ are displayed below, again for the representative case $a=3$, $b=6$. We shall consider the probabilities $\phi_{ij^*}(x)$ that absorption occurs in the state $j^* \in \{a^*, \ldots, b^*\}$ before time x, given that the process is started in the state $i \in \{1, \ldots, b, 1^*, \ldots, (a-1)^*\}$. The relation of these probabilities to the mass-functions $W_j(\cdot)$ will soon be clear.

	1	2	3	4	5	6	1*	2*	3*	4*	5*	6*
1	$-\lambda-N\mu$	λ					$N\mu$					
2		$-\lambda-N\mu$	λ					$N\mu$				
3			$-\lambda-N\mu$	λ					$N\mu$			
4				$-\lambda-N\mu$	λ					$N\mu$		
5					$-\lambda-N\mu$	λ					$N\mu$	
6						$-N\mu$						$N\mu$
1*							$-\lambda$	λ				
2*								$-\lambda$	λ			

D $\tilde{D}$

If we partition the $(a+b-1) \times (2b)$ array into a square matrix D of order $(a+b-1)$ and a matrix $\tilde{D}$ of dimensions $(a+b-1) \times (b-a+1)$, then it is readily seen that the matrix $\Phi(x) = \{\Phi_{ij^*}(x)\}$ is given by

$$(5) \quad \Phi(x) = \int_0^x \exp(Du)\tilde{D}\,du = [\exp(Dx)-I]D^{-1}\tilde{D}, \quad \text{for } x \geq 0.$$

Suppose now that the arriving customer finds $b\nu+r$, $\nu \geq 1$, $0 \leq r \leq b-1$, customers ahead of him in the queue. The first $b\nu$ customers will be removed in ν batches of size b. The remaining r customers and the arriving customer will be served together in a batch of size k with $\max(a,r+1) \leq k \leq b$. The size of that batch will depend on the number of customers who arrive during the service of the ν batches and <u>of those arriving during the wait for a free server</u>. The following lemma is useful in the sequel.

<u>Lemma 1</u>

The probability that the arriving customer finds at least b customers in the queue, that all groups of size b ahead of him are served no later than time x and that at the beginning of the service of the last such group, there are j, $1 \leq j \leq b$, customers in the queue who will be served in the same group as the arriving customer, is given by

$$G_j(x) = \sum_{r=0}^{j-1} x_0 \xi^{b+r} \int_0^x \exp[-N\mu(1-\xi^b)u] e^{-\lambda u} \frac{(\lambda u)^{j-r-1}}{(j-r-1)!} N\mu du ,$$

(6) for $1 \leq j \leq b-1$.

$$G_b(x) = \sum_{r=0}^{b-1} x_0 \xi^{b+r} \int_0^x \exp[-N\mu(1-\xi^b)u] \sum_{k=b-r}^{\infty} e^{-\lambda u} \frac{(\lambda u)^{k-1}}{(k-1)!} N\mu du .$$

Proof

By the law of total probability, we see that for $1 \leq j \leq b-1$, the probability, described in the statement of the lemma, is given by

$$\sum_{\nu=1}^{\infty} \sum_{r=0}^{j-1} x_0 \xi^{b\nu+r} \int_0^x e^{-N\mu u} \frac{(N\mu u)^{\nu-1}}{(\nu-1)!} N\mu \cdot e^{-\lambda u} \frac{(\lambda u)^{j-r-1}}{(j-r-1)!} du .$$

This readily reduces to the stated expression. The proof for $j=b$, is similar. We emphasize that the arriving customer is now included in the count described by the index j .

We can now describe the mass-functions $W_j(\cdot)$ for $a \leq j \leq b$. The function $W_a(\cdot)$ contains an extra term, given by

$$V_a(x) = \sum_{i=1}^{N} x_{i,a-1} + \sum_{i=1}^{N} \sum_{j=0}^{a-2} x_{i,j} \int_0^x e^{-\lambda u} \frac{(\lambda u)^{a-j-2}}{(a-j-2)!} \lambda du ,$$

for $x \geq 0$. This term corresponds to the case where the arriving customer finds one or more free servers.

He may then have to wait until a sufficient number of additional customers arrive to initiate service of a group of size a .

The states $\{1, \ldots, b\}$ of the absorbing Markov process, which was defined above, correspond to the situation where all complete groups of size b , who are ahead of the arriving customer, have been served; all additional arrivals to be included in the next group have been counted into the index j , <u>but there is no free server yet</u>. The corresponding state j^* has the same significance, except that there is also a free server present. It is now clear that absorption into one of the states $\{a^*, \ldots, b^*\}$ signals the start of the service of a group of the corresponding size. This group includes the arriving customer.

By $\underline{e}^o$, we denote the row vector $(1, 0, \ldots, 0)$ of dimension $b-a+1$. The vector $\underline{\theta}$ of dimension $a+b-1$ is defined by $x_0(1, \xi, \ldots, \xi^{b-1}, 0, \ldots, 0)$ and $\underline{G}(x) = [G_1(x), \ldots, G_b(x), 0, \ldots, 0]$. The row vector $\underline{W}(x)$ of dimension $b-a+1$ has the components $W_a(x), \ldots, W_b(x)$.

<u>Theorem 2</u>

The vector $\underline{W}(x)$ is given by

$$(7) \quad \underline{W}(x) = V_a(x)\, \underline{e}^o + \underline{\theta} \int_0^x \exp(Du)\tilde{D}\, du + \int_0^x \underline{G}(u)\, \exp[D(x-u)]\tilde{D}\, du ,$$

for $x \geq 0$.

Proof

Formula (7) is immediate from the consideration of several absorption times. The first term corresponds to the case where the arriving customer finds one or more free servers. It adds only to $W_a(x)$.

The second term corresponds to the case where the arriving customer finds j, $0 \le j < b$, customers and no free server in the system.

In the last term, the arriving customer finds $j \ge b$ customers ahead of him. The waiting time in that case is analyzed as a first passage time in the preceding dicussion.

Each of the terms in Formula (7) may be computed by the numerical solution of fairly simple systems of differential equations. In order to evaluate $V_a(x)$, we rewrite the defining expression as

$$V_a(x) = \sum_{i=0}^{N} \sum_{j=0}^{a-1} x_{ij} - \sum_{j=0}^{a-2} x_j \int_x^{\infty} e^{-\lambda u} \frac{(\lambda u)^{a-2-j}}{(a-2-j)!} \lambda \, du$$

$$= 1 - x_0(1-\xi)^{-1} - y \exp(\Lambda x)\, e \,, \quad \text{for } x \ge 0 .$$

The vector $\underline{\alpha}$ has the components $\alpha_j = \sum_{i=1}^{N} x_{ij}$, for $0 \leq j \leq a-2$. The matrix Λ of order $a-1$ is the same as the matrix C , defined in Section 1. Clearly $V_a(x) = x_0(1-\rho)^{-1} - \underline{z}_0(x)\,\underline{e}$, where $\underline{z}_0'(x) = \underline{z}_0(x)\Lambda$, for $x \geq 0$, with $\underline{z}_0(0) = \underline{\alpha}$.

In order to evaluate the second term, we write

$$\underline{z}_1(x) = \underline{\theta} \int_0^x \exp(Du)\, du = \underline{\theta}[\exp(Dx) - I]\, D^{-1} .$$

It is now clear that $\underline{z}_1(x)$ satisfies the system of differential equations

$$\underline{z}_1'(x) = \underline{z}_1(x) D + \underline{\theta} \qquad \text{for } x \geq 0 ,$$

with $\underline{z}_1(0) = \underline{0}$. The second term in Formula (7) is then given by $\underline{z}_1(x)\, D$.

Next we introduce the vector-valued function

$$\underline{z}_2(x) = \int_0^x \underline{G}(u) \exp[D(x-u)]\, du , \qquad \text{for } x \geq 0 .$$

By multiplying both sides by $\exp(-Dx)$ and differentiating, we see that the vector $\underline{z}_2(x)$ satisfies the system of differential equations

$$\underline{z}_2'(x) = \underline{z}_2(x) D + \underline{G}(x) , \qquad \text{for } x \geq 0 ,$$

with $\underline{z}_2(0) = \underline{0}$. The components of the vector $\underline{G}(x)$ are evaluated by numerical integration in the expressions (6).

Corollary 2

The vector $\underline{W}(\infty)$ is given by

$$(8) \qquad \underline{W}(\infty) = [1 - x_0(1-\xi)^{-1}]\, \underline{e}^0 - \underline{\xi}\, D^{-1}\tilde{D} - \underline{G}(\infty) D^{-1}\tilde{D} ,$$

where the components of the vector $\underline{G}(\infty)$ are explicitly given by

$$(9) \qquad \begin{aligned} G_j(\infty) &= j \frac{N\mu}{\lambda} x_0\, \xi^{b+j}, \qquad , \qquad \text{for } 1 \leq j \leq b-1 , \\ G_b(\infty) &= b \frac{N\mu}{\lambda} x_0\, \xi^{2b} (1-\xi)^{-1} . \end{aligned}$$

Proof

By letting x tend to infinity in Formula (7). The expressions for the components of $\underline{G}(\infty)$ follow from (6) by routine integrations and summations in which we recall that $\lambda[\lambda + N\mu(1 - \xi^b)]^{-1} = \xi$, by virtue of (1).

Remark

By noting that $D^{-1}\tilde{D}\underline{e} = \underline{e}$, $\underline{G}(\infty)\underline{e} = x_0\, \xi^b (1 - \xi)^{-1}$, $\underline{\xi}\, \underline{e} = (1 - \xi^b)(1 - \xi)^{-1}$, we readily verify that $\underline{W}(\infty)\underline{e} = 1$.

Corollary 1 gives us the probability density of the size of the group in which a customer arriving to the stationary queue will be served. It is given by the components of the vector $\underline{W}(\infty)$.

The means of the mass-functions $W_j(\cdot)$, $a \le j \le b$, may be computed from Formula (7). Denoting Laplace-Stieltjes transforms by the addition of asterisks, (7) readily yields

$$\underline{W}^*(s) = V_a^*(s)\underline{e}^0 + \underline{\xi}(sI-D)^{-1}\tilde{D} + \underline{G}^*(s)(sI-D)^{-1}\tilde{D} ,$$

so that

$$\underline{W}^{*'}(s) = V_a^{*'}(s)\underline{e}^0 - \underline{\xi}(sI-D)^{-2}\tilde{D} - \underline{G}^*(s)(sI-D)^{-2}\tilde{D} + \underline{G}^{*'}(s)(sI-D)^{-1}\tilde{D}.$$

It follows that

$$(10)\qquad \hat{\underline{W}} = -\underline{W}^{*'}(0) = -V_a^{*'}(0)\underline{e}^0 + [\underline{\xi}+\underline{G}(\infty)]D^{-2}\tilde{D} + \underline{G}^{*'}(0)D^{-1}\tilde{D} .$$

Clearly

$$(11)\qquad -V_a^{*'}(0) = \frac{1}{\lambda}\sum_{i=1}^{N}\sum_{j=0}^{a-2}(a-j-1)x_{ij} ,$$

so that it suffices to evaluate $\underline{G}^{*'}(0)$.

From (6), we obtain routinely that for $1 \le j \le b-1$,

$$(12)\qquad G_j^*(s) = \frac{N\mu}{\lambda}\sum_{r=0}^{j-1} x_0\xi^{b+r}\left[\frac{\lambda}{s+\lambda+N\mu(1-\xi^b)}\right]^{j-r}$$

$$= \frac{N\mu}{\lambda}x_0\xi^{b+j}\sum_{r=0}^{j-1}\left[\frac{\lambda+N\mu(1-\xi^b)}{s+\lambda+N\mu(1-\xi^b)}\right]^{j-r}$$

$$= \frac{N\mu}{\lambda}x_0\xi^{b+j}\sum_{r=1}^{j}\left[\frac{\lambda\xi^{-1}}{s+\lambda\xi^{-1}}\right]^{r} .$$

A similar calculation yields

$$(13)\quad G_b^*(s) = \frac{N\mu}{\lambda} x_0 \xi^{2b} \sum_{r=1}^{b} \left(\frac{\lambda\xi^{-1}}{s+\lambda\xi^{-1}}\right)^r \left[1 + \frac{\xi}{1-\xi} \frac{(1-\xi)\lambda\xi^{-1}}{s+(1-\xi)\lambda\xi^{-1}}\right].$$

It now readily follows that

$$-G_j^{*'}(0) = \frac{N\mu}{\lambda^2} x_0 \xi^{b+j+1} \frac{j(j+1)}{2}, \qquad \text{for } 1 \le j \le b-1,$$

(14)

$$-G_b^{*'}(0) = \frac{N\mu}{\lambda^2} x_0 \xi^{2b+1} (1-\xi)^{-2} b \left[\xi + (1-\xi)\frac{b+1}{2}\right].$$

The conditional mean waiting time, given that the arriving customer is served in a group of size j is given by $\hat{W}_j W_j^{-1}$, for $a \le j \le b$. The unconditional mean waiting time is given by

$$(15)\qquad \hat{W}\mathbf{e} = -V_a^{*'}(0) - \left[\hat{\mathbf{e}} + \mathbf{G}(\infty)\right] D^{-1}\mathbf{e} - \mathbf{G}^{*'}(0)\mathbf{e}.$$

<u>Remark</u>

From the formulas (12) and (13), we obtain the following alternate expressions for the $G_j(x)$.

$$G_j(x) = \frac{N\mu}{\lambda} x_0 \xi^{b+j} \sum_{r=1}^{j} E_r(\lambda\xi^{-1}, x), \qquad \text{for } 1 \le j \le b-1,$$

(16)

$$G_b(x) = \frac{N\mu}{\lambda} x_0 \xi^{2b} \sum_{r=1}^{b} E_r(\lambda\xi^{-1}, \cdot) * \left[U(x) + \frac{\xi}{1-\xi} E_1(\lambda\xi^{-1}(1-\xi), x)\right].$$

These expressions are better suited for numerical computation than those given in Formula (6).

From Formula (6), we obtain by a routine summation that

$$(17)\qquad \sum_{j=1}^{b} G_j(x) = N\mu\, x_0 \xi^b (1-\xi^b)(1-\xi)^{-1} \int_0^x \exp[-N\mu(1-\xi^b)u]\, du ,$$

so that

$$(18)\qquad -G^{*'}(0)\underline{e} = \frac{1}{N\mu}\, x_0 \xi^b (1-\xi^b)^{-1}(1-\xi)^{-1} .$$

If, on the other hand, we add the expressions in (14), we obtain after some calculations that

$$(19)\qquad -G^{*'}(0)\underline{e} = \frac{N\mu}{\lambda^2}\, x_0 \xi^{b+2}(1-\xi^b)(1-\xi)^{-3} .$$

The expressions in (18) and (19) are equal, since by virtue of Equation (1), we have

$$\frac{N\mu}{\lambda} \cdot (1-\xi^b)\xi(1-\xi)^{-1} = 1 .$$

This provides us with an accuracy check on the lengthy calculations involved in finding the expressions for the mean waiting time.

4. Algorithmic Aspects

The equilibrium condition (2) readily defines the smallest value of Nb for which the queue is stable. Equation (1) is rewritten as

$$(20)\qquad \xi = 1 - p + p\xi^{b+1} ,$$

where $p = N\mu(\lambda+N\mu)^{-1}$, and is solved e.g. by Newton's method.

Defining the matrices $M(j)$, $1 \le j \le N$, by

$$(21)\qquad M(j) = [(N-j)\mu I - C]^{-1} ,$$

the first N equations in (4) may be rewritten as

$$\underline{x}_N = \mu \underline{x}_{N-1} M(N) ,$$

$$(22) \quad \underline{x}_j = [(N-j+1)\mu \underline{x}_{j-1} + \underline{x}_{j+1} E] M(j), \qquad \text{for } 2 \le j \le N ,$$

$$\underline{x}_1 = [x_0 N (1, \xi, \ldots, \xi^{a-1}) + \underline{x}_2 E] M(1) .$$

We performed Gauss-Seidel iteration in the equations (22), computed the next value of x_0 from the penultimate equation in (4) and renormalized each solution vector by the last equation in (4). Any non-zero, nonnegative initial vector $[\underline{x}_N, \ldots, \underline{x}_1, x_0]$ may be chosen to start the iterative solution.

The special form of the matrices C and E may be used to economize on storage and computational effort. Only the first component of each vector $\underline{x}_j E$ differs from zero and it requires only one multiplication. The matrices $(N-j)\mu I-C$, $1 \le j \le N$, are of the form

$$\begin{vmatrix} \delta & -\lambda & & & \\ & \delta & -\lambda & & \\ & & \cdots & & \\ & & & \delta & -\lambda \\ & & & & \delta \end{vmatrix} ,$$

where $\delta \geq \lambda \geq 0$. The matrices $M(j)$ are therefore upper triangular and of the form

$$(23) \qquad \begin{vmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{a-1} \\ & \gamma_0 & \gamma_1 & \cdots & \gamma_{a-2} \\ & & \cdots & & \\ & & & \gamma_0 & \gamma_1 \\ & & & & \gamma_0 \end{vmatrix} ,$$

where $\gamma_\nu = \delta^{-1}(\lambda/\delta)^\nu$, for $0 \leq \nu \leq a-1$. It is hence not necessary to store the N matrices $M(j)$. The computation of the stationary probability vector of the matrix Q is both stable and very fast.

The vector $\underline{W}(\infty)$ is easily computed from the formulas (8) and (9). In view of the special form of the matrix D , the vector $-[\underline{e} + \underline{G}(\infty)] D^{-1}$ is obtained with very little effort.

As was discussed in Section 3, the vector $\underline{W}(x)$ of mass-functions is computed by the numerical integration of simple systems of linear differential equations with constant coefficients. This may be accomplished by any one of a number of classical methods. The special structure of the coefficient matrices is exploited to reduce computation time. The approach to the easily computed vector $\underline{W}(\infty)$ provides us with an accuracy check on the numerical integration procedure. For examples in which the parameter a is large or in which the queue is nearly critical, the evaluation of $\underline{W}(x)$, $x \geq 0$, requires fairly substantial computation times. In all other cases, the approach to $\underline{W}(\infty)$ is very rapid.

5. Acknowledgements

Research on this problem was initiated during a visit of the first author to Annamalai University during December 1979. The authors express their profound appreciation to Professor G. Sankaranarayanan for his great hospitality and enthusiastic encouragement.

REFERENCES

[1] Neuts, M. F.
A General Class of Bulk Queues with Poisson Input
Ann. Math. Stat., 38, 759-770, 1967.

[2] Neuts, M. F.
Queues Solvable without Rouché's Theorem
Opns. Res., 27, 767-781, 1979.

[3] Neuts, M. F.
Matrix-geometric Solutions in Stochastic Models-
An Algorithmic Approach
The Johns Hopkins University Press, Baltimore MD -
To appear.

[4] Winsten, C. B.
Geometric Distributions in the Theory of Queues
J. Roy. Stat. Soc., XXI, 1-35, 1959.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 80-0969	2. GOVT ACCESSION NO. AD-A090 530	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A MULTI-SERVER QUEUE WITH THRESHOLDS FOR THE ACCEPTANCE OF CUSTOMERS INTO SERVICE		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Marcel Neuts R. Nadarajan		8. CONTRACT OR GRANT NUMBER(s) AFOSR-77-3236
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Delaware Applied Mathematics Institute Newark, DE 19711		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE June 1980 13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multi-server queues, group service, optimal design of queues, computational probability

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A queue with Poisson arrivals is served by N identical exponential servers. When a server becomes free, he can serve a group of customers of size at most b. He is not allowed to process a group of size less than a, 1 a b. The actual size of each group served depends on the number of waiting customers. The queue is studied as a Markov process. In the stable case, the stationary probability vector of this process has a simple, readily computable form. Using this form, the stationary waiting time distribution, which is analytically intractable, may be expressed in an algorithmically useful form. (OVER)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

20. Abstract cont.

Several questions, related to the optimal design of such a service system, may be algorithmically investigated. This model serves to illustrate the advantages of the algorithmic approach. The design criteria are not analytically tractable in general.

FILM