J. Appl. Prob. **17**, 1048–1061 (1980)
Printed in Israel
0021-9002/80/041048–14 $01.65

STATIONARY STATE PROBABILITIES AT ARRIVAL INSTANTS FOR CLOSED QUEUEING NETWORKS WITH MULTIPLE TYPES OF CUSTOMERS

S. S. LAVENBERG,* *IBM Thomas J. Watson Research Center, Yorktown Heights*
M. REISER,** *IBM Research Laboratory, Zürich*

Abstract

We consider closed networks of interconnected service centers with multiple types of customers and multiple classes, whose stationary state probabilities at arbitrary times have a product form. A customer can change its class but not its type as it traverses the network. We show that the stationary state probabilities at instants at which customers of a particular type arrive at a particular service center and enter a particular class are equal to the stationary state probabilities at arbitrary times for the network with one less customer of that type. Applications of this result are given.

QUEUEING NETWORKS; MULTIPLE TYPES OF CUSTOMERS; STATIONARY STATE PROBABILITIES; STATIONARY WAITING-TIME DISTRIBUTIONS

1. Introduction

Queueing networks whose stationary state probabilities have a product form have received considerable recent attention in the literature, see e.g. Baskett, Chandy, Muntz and Palacios (1975), Kelly (1975), Kelly (1976), Barbour (1976), and see Lemoine (1977) for a survey. These papers deal with stationary state probabilities at arbitrary times. In this paper we consider a family of product-form closed networks with multiple types of customers and multiple classes. A customer can change its class but it cannot change its type as it traverses the network. The family of networks is somewhat more general than the closed networks in Baskett et al. (1975) and is closely related to the closed networks in Kelly (1975). We are concerned with stationary state probabilities at instants at which customers of a particular type arrive at a particular service center and enter a particular class. We show that the stationary state probabilities at these arrival instants (where the state excludes the arriving customer) are equal to the

Received 11 April 1979; revision received 30 October 1979.
* Postal address: IBM Thomas J. Watson Research Center, P. O. Box 218, Yorktown Heights, NY 10598, U.S.A.
** Postal address: IBM Research Laboratory, Zürich, Switzerland.

stationary state probabilities at arbitrary times for the network with one less customer of that type.

This arrival instant result was conjectured to hold, but was not proven to hold, for a less general family of networks by Reiser and Lavenberg (1978). Subsequently, Sevcik and Mitrani (1979) tried to prove results of this kind for very general networks, but their proofs were not mathematically sound. The arrival instants we consider are examples of what Melamed (1978) called traffic epochs. Traffic epochs are defined with respect to a specified set of transitions, called a traffic set, of a regular Markov jump process and are the times of occurrence of transitions in the traffic set. Melamed obtained results relating the stationary state probabilities at traffic epochs to the stationary state probabilities at arbitrary times. He applied these results to queueing networks and showed that for certain traffic sets the stationary state probabilities at traffic epochs are related in a simple way to the stationary state probabilities at arbitrary times. However, he did not consider traffic sets whose traffic epochs are arrival instants at a particular class. Such a traffic set is the union of disjoint traffic sets of the type considered by Melamed. As Melamed showed by example, a simple relation between the stationary state probabilities at traffic epochs and at arbitrary times need not hold for a union of disjoint traffic sets. We establish an important case where a simple relation does hold.

We define the family of networks in Section 2, derive the stationary state probabilities at arbitrary times in Section 3, derive the stationary state probabilities at arrival instants in Section 4 and give applications of the arrival instant result in Section 5. In particular, the result allows stationary waiting-time distributions to be computed for any service center which is a standard multiserver queue having exponential service times.

2. Closed networks having product form

We will consider a family of closed finite-state Markovian queueing networks whose stationary state probabilities have a product form. A network consists of S interconnected service centers. There are R types of customers and a fixed number, $K(r)$, of type r customers, $r = 1, \cdots, R$. These customers circulate indefinitely among the service centers in a manner to be described below.

There is a finite set $C_s(r)$ of classes for type r customers at service center $s, r = 1, \cdots, R$, $s = 1, \cdots, S$. These sets are disjoint and $C_s(r)$ can be empty for some r and s with the restriction that the set of all classes for type r customers, $C(r) = \bigcup_{s=1}^{S} C_s(r)$, and the set of all classes at service center s, $C_s = \bigcup_{r=1}^{R} C_s(r)$, are not empty. When a type r customer enters service center s it enters one of the classes in $C_s(r)$ and remains in that class until it leaves the service center. We make the following assumptions.

Assumption (a). For each type r customer the sequence of classes entered is an irreducible Markov chain with state space $C(r)$ and transition probability matrix $P(r)=(p_{cd}(r): c \in C(r), d \in C(r))$.

Assumption (b). For each type r customer successive service demands in a particular class form a sequence of independent and identically distributed random variables whose common distribution function is a finite mixture of gamma distributions each of which has integral shape parameter. Such random variables can be represented by a finite number of exponential stages.

Assumption (c). The class sequences (there is one for each customer) and the service demand sequences (there is one for each customer and each class that customer can enter) are mutually independent random sequences.

Thus, a customer's current class determines the probability distribution function of the customer's service demand and the probability distribution of the next class it enters. Although the sequence of classes entered by a customer is a Markov chain, the sequence of service centers entered by the customer need not be. With classes the sequence of service centers can consist of repetitions of a deterministic finite-length sequence.

We now describe the operation of a service center. An ordered list is kept of the customers currently in the service center. When n customers are in service center s, the lth customer in the list receives service at rate $\gamma_s(l, n)\phi_s(n) \geqq 0$, i.e., that customer's remaining service demand decreases at the rate of $\gamma_s(l, n)\phi_s(n)$ units per second, where $\phi_s(n)>0$, if $n>0$ *and*

$$\sum_{l=1}^{n} \gamma_s(l, n)=1. \tag{1}$$

$\phi_s(n)$ is the total service rate of service center s when n customers are present and $\gamma_s(l, n)$ is the fraction of that rate received by the lth customer. A customer leaves the service center when its remaining service demand is 0. The list of customers for a service center is updated as follows: when n customers are in the service center and the customer in position l leaves, the customers previously in positions $l+1, \cdots, n$ (if any) enter positions $l, \cdots, n-1$ respectively. When $n-1$ customers are in service center s and a customer arrives, the new customer enters position l with probability $\delta_s(l, n)$, $l=1, \cdots, n$, where

$$\sum_{l=1}^{n} \delta_s(l, n)=1. \tag{2}$$

The customers previously in positions $l, \cdots, n-1$ (if any) enter positions $l+1, \cdots, n$ respectively. This description of a service center is due to Kelly (1975). We make the following final assumption.

Assumption (d). Any service center s is of one of the following three types:
Type 1. All service demands at the service center are exponentially distributed and have the same mean for all classes at the service center. For $n = 1, 2, \cdots$

$$\delta_s(l, n) = \begin{cases} 1 & l = n \\ 0 & l < n. \end{cases} \tag{3}$$

Type 2. For $n = 1, 2, \cdots$ and $l \leqq n$

$$\delta_s(l, n) = \gamma_s(l, n) = 1/n. \tag{4}$$

Type 3. For $n = 1, 2, \cdots$

$$\delta_s(l, n) = \gamma_s(l, n) = \begin{cases} 1 & l = n \\ 0 & l < n. \end{cases} \tag{5}$$

This family of networks is somewhat more general than the family of closed networks considered in Baskett et al. (1975) in that the above three types of service centers include the service centers they considered as special cases. If service center s is type 1 and for $n = 1, 2, \cdots$

$$\gamma_s(l, n) = \begin{cases} 1/\min(m, n) & 1 \leqq l \leqq \min(m, n) \\ 0 & \min(m, n) < l \leqq n \end{cases} \tag{6}$$

where m is an integer then the service center has m servers, the queueing discipline is first-come–first-served and if $n > 0$ the service rate of each busy server is $\phi_s(n)/\min(n, m)$. The queueing discipline at a type 2 service center is called processor sharing, i.e., all customers receive equal shares of the total service rate $\phi_s(n)$. The queueing discipline at a type 3 service center is last-come–first-served preemptive resume. Our use of classes is easily shown to be equivalent to the use of classes by Baskett et al. Classes are a powerful concept. As we will see in Appendix B, non-exponential service demands (as described in Assumption (b)) at type 2 or type 3 service centers can be represented using classes. In Kelly (1975) classes are not used and all service demands are assumed to be exponentially distributed. However, (3) is not required to hold for a type 1 service center and for all other service centers it is necessary only that $\delta_s(l, n) = \gamma_s(l, n)$, i.e., the rightmost equalities in (4) and (5) need not hold.

The following notation will be used:

S: number of service centers
R: number of types of customers
$K(r)$: number of customers of type r
$C_s(r)$: set of classes for type r customers at service center s

$C(r)$: set of classes for type r customers
C_s: set of classes at service center s
$P(r) = (p_{cd}(r): c \in C(r), d \in C(r))$: transition probability matrix for the irreducible Markov chain over $C(r)$
$(e_c(r): c \in C(r))$: stationary probability distribution for the irreducible Markov chain over $C(r)$
$1/\lambda_c$: mean service demand in class c
$\phi_s(n)$: total service rate of service center s when n customers are present
$\gamma_s(l, n)$: fraction of the total service rate of service center s received by the customer in position l when n customers are present
$\delta_s(l, n)$: probability a customer arriving at service center s when $n-1$ customers are present enters position l
$s(c)$: service center corresponding to class c
$r(c)$: customer type corresponding to class c
$n(c)$: number of customers in class c
n_s: number of customers at service center s
$c_s(l)$: class of the customer in the lth position at service center s
$r_s(l)$: type of the customer in the lth position at service center s
$\boldsymbol{n}_s = (n(c): c \in C_s)$: population state of service center s
$\boldsymbol{n} = (\boldsymbol{n}_1, \cdots, \boldsymbol{n}_S)$: population state of the network
$\mathcal{N} = \{\boldsymbol{n} : \Sigma_{c \in C(r)} n(c) = K(r),\ r = 1, \cdots, R\}$: set of population states of the network
$\boldsymbol{c}_s = (c_s(1), \cdots, c_s(n_s))$: detailed state of service center s
$\boldsymbol{c} = (\boldsymbol{c}_1, \cdots, \boldsymbol{c}_S)$: detailed state of the network
$\mathcal{S} = \{\boldsymbol{c} : \Sigma_{c \in C(r)} n(c) = K(r),\ r = 1, \cdots, R\}$: set of detailed states of the network.

3. Preliminary results

We first assume that all service demands are exponentially distributed. In Appendix B we will remove this restriction by using additional classes to represent the exponential stages of service for a non-exponential service demand. Letting $\boldsymbol{c}(t)$ denote the detailed state of the network at time t, $\mathscr{C} = \{\boldsymbol{c}(t) : t \geqq 0\}$ is a finite-state continuous-time Markov process. We assume that $\mathscr{C}$ is irreducible on $\mathcal{S}$. Then the stationary distribution of $\mathscr{C}$ is the unique solution $\{p(\boldsymbol{c}) : \boldsymbol{c} \in \mathcal{S}\}$ to the equations

$$p(\boldsymbol{c})q(\boldsymbol{c}) = \sum_{\boldsymbol{d} \in \mathcal{S}} p(\boldsymbol{d})q(\boldsymbol{d}, \boldsymbol{c}) \qquad \boldsymbol{c} \in \mathcal{S} \tag{7}$$

$$\sum_{\boldsymbol{c} \in \mathcal{S}} p(\boldsymbol{c}) = 1, \tag{8}$$

where $q(\boldsymbol{c}, \boldsymbol{d})$, $\boldsymbol{d} \neq \boldsymbol{c}$, is the transition rate from $\boldsymbol{c}$ to $\boldsymbol{d}$, $q(\boldsymbol{c}, \boldsymbol{c}) = 0$ *and*

$$q(\boldsymbol{c}) = \sum_{\boldsymbol{d} \in \mathcal{S}} q(\boldsymbol{c}, \boldsymbol{d}). \tag{9}$$

In a similar manner to that used in Kelly ((1975), pp. 544–546) we next write the set of equations in (7) for a network. For any $\boldsymbol{c} \in \mathcal{S}$ let $\boldsymbol{c}(s, l; h, m)$ denote the state obtained from $\boldsymbol{c}$ by deleting the customer in position l in service center s and inserting a class h customer in position m in service center $s(h)$, where $h \in C[r_s(l)]$. Any state which can reach $\boldsymbol{c}$ in a one-step transition of $\mathscr{C}$ is of the form $\boldsymbol{c}(s, l; h, m)$. The set of equations in (7) can be written

$$\begin{aligned} &p(\boldsymbol{c}) \sum_{s=1}^{S} \sum_{l=1}^{n_s} \gamma_s(l, n_s)\phi_s(n_s)\lambda_{c_s(l)} \\ &= \sum_{s=1}^{S} \sum_{l=1}^{n_s} \sum_{h \in C[r_s(l)]} \sum_{m=1}^{n_{s(h)}+1} \{p[\boldsymbol{c}(s, l; h, m)]\gamma_{s(h)}(m, n_{s(h)}+1)\phi_{s(h)}(n_{s(h)}+1) \\ &\qquad \times \lambda_h p_{hc_s(l)}[r_s(l)]\delta_s(l, n_s)\} \qquad \boldsymbol{c} \in \mathcal{S}. \end{aligned} \tag{10}$$

It is straightforward to verify that the unique solution to (8) and (10) is given by

$$p(\boldsymbol{c}) = a(\boldsymbol{c})/b \qquad \boldsymbol{c} \in \mathcal{S} \tag{11}$$

where

$$a(\boldsymbol{c}) = \prod_{s=1}^{S} a_s(\boldsymbol{c}_s) \tag{12}$$

$$a_s(\boldsymbol{c}_s) = \prod_{l=1}^{n_s} e_{c_s(l)}[r_s(l)]/[\lambda_{c_s(l)}\phi_s(l)] \tag{13}$$

and

$$b = \sum_{\boldsymbol{c} \in \mathcal{S}} a(\boldsymbol{c}). \tag{14}$$

For any $\boldsymbol{c} \in \mathcal{S}$ let $\boldsymbol{m}(\boldsymbol{c})$ denote the population state corresponding to detailed state $\boldsymbol{c}$. For any $\boldsymbol{n} \in \mathcal{N}$ let

$$\mathcal{S}(\boldsymbol{n}) = \{\boldsymbol{c} \in \mathcal{S} : \boldsymbol{m}(\boldsymbol{c}) = \boldsymbol{n}\} \tag{15}$$

and

$$\pi(\boldsymbol{n}) = \sum_{\boldsymbol{c} \in \mathcal{S}(\boldsymbol{n})} p(\boldsymbol{c}). \tag{16}$$

$\{\pi(\boldsymbol{n}) : \boldsymbol{n} \in \mathcal{N}\}$ is the unique stationary distribution over the set of population states. Note that $p(\boldsymbol{c})$ has the same values for all $\boldsymbol{c} \in \mathcal{S}(\boldsymbol{n})$ and that

$$|\mathcal{S}(\boldsymbol{n})| = \prod_{s=1}^{S} n_s! \prod_{c \in C_s} 1/n(c)!. \tag{17}$$

Therefore,

$$\pi(\boldsymbol{n}) = \alpha(\boldsymbol{n})/b \qquad \boldsymbol{n} \in \mathcal{N} \tag{18}$$

where

$$\alpha(\boldsymbol{n}) = \prod_{s=1}^{S} \alpha_s(\boldsymbol{n}_s) \tag{19}$$

$$\alpha_s(\boldsymbol{n}_s) = \left[\prod_{l=1}^{n_s} 1/\phi_s(l)\right] n_s! \prod_{c \in C_s} (1/n(c)!)(e_c[r(c)]/\lambda_c)^{n(c)} \tag{20}$$

and b which was given in (14) simplifies to

$$b = \sum_{\boldsymbol{n} \in \mathcal{N}} \alpha(\boldsymbol{n}). \tag{21}$$

We conclude this section by considering the imbedded Markov chain just after transitions in $\mathscr{C}$, i.e., the jump chain associated with $\mathscr{C}$. We denote this chain by $\mathscr{C}^*$. Then $\mathscr{C}^*$ is irreducible over $\mathscr{S}$ and has transition probabilities

$$q^*(\boldsymbol{c}, \boldsymbol{d}) = q(\boldsymbol{c}, \boldsymbol{d})/q(\boldsymbol{c}) \qquad \boldsymbol{c}, \boldsymbol{d} \in \mathscr{S}. \tag{22}$$

The unique stationary distribution of $\mathscr{C}^*$ is given by

$$p^*(\boldsymbol{c}) = a(\boldsymbol{c})q(\boldsymbol{c})/b^* \qquad \boldsymbol{c} \in \mathscr{S} \tag{23}$$

where

$$b^* = \sum_{\boldsymbol{c} \in \mathscr{S}} a(\boldsymbol{c})q(\boldsymbol{c}). \tag{24}$$

4. Principal result

We wish to obtain the stationary probability distribution of the sequence of population states at successive time instants at which customers enter a particular class, say class g. We temporarily assume that $p_{gg}[r(g)] = 0$. This restriction guarantees that customers can only enter class g at transitions in $\mathscr{C}$. We first consider the sequence of detailed states at transitions of $\mathscr{C}^*$ at which a customer enters class g. We denote this Markov chain by $\mathscr{C}_g$ and let $\mathscr{S}_g \subset \mathscr{S}$ denote the set of states which $\mathscr{C}_g$ can visit, i.e.,

$$\mathscr{S}_g = \begin{cases} \{\boldsymbol{c} \in \mathscr{S} : c_{s(g)}(n_{s(g)}) = g\} & \text{if service center } s(g) \text{ is type 1 or 3} \\ \{\boldsymbol{c} \in \mathscr{S} : n(g) \geqq 1\} & \text{if service center } s(g) \text{ is type 2.} \end{cases} \tag{25}$$

Since $\mathscr{C}^*$ is irreducible on $\mathscr{S}$, $\mathscr{C}_g$ is irreducible on $\mathscr{S}_g$. Thus, $\mathscr{C}_g$ has a unique stationary distribution on $\mathscr{S}_g$ which we denote by $\{p_g(\boldsymbol{c}) : \boldsymbol{c} \in \mathscr{S}_g\}$.

For $\boldsymbol{c} \in \mathscr{S}_g$, let $\mathscr{S}_g(\boldsymbol{c}) = \{\boldsymbol{d} \in \mathscr{S} : \mathscr{C}^*$ can make a one-step transition from $\boldsymbol{d}$ to $\boldsymbol{c}$ at which a customer enters class $g\}$. We next prove the following lemma, which also follows from a more general result (Theorem 3.1) in Melamed (1978).

Lemma 1.

$$p_g(\mathbf{c}) = \sum_{\mathbf{d} \in \mathcal{S}_g(\mathbf{c})} a(\mathbf{d}) q(\mathbf{d}, \mathbf{c}) / b_g \qquad \mathbf{c} \in \mathcal{S}_g \tag{26}$$

where b_g is a normalizing constant chosen so that $\{p_g(\mathbf{c}) : \mathbf{c} \in \mathcal{S}_g\}$ is a probability distribution.

Proof. Choose a fixed state $\mathbf{c}_0 \in \mathcal{S}$ and let A_g = number of transitions of $\mathcal{C}_g$ between successive returns of $\mathcal{C}^*$ to $\mathbf{c}_0$, $A_g(\mathbf{c})$ = number of times $\mathcal{C}_g$ enters $\mathbf{c}$ between successive returns of $\mathcal{C}^*$ to $\mathbf{c}_0$. Then,

$$p_g(\mathbf{c}) = E[A_g(\mathbf{c})]/E[A_g], \tag{27}$$

where

$$\begin{aligned} E[A_g(\mathbf{c})] &= \sum_{\mathbf{d} \in \mathcal{S}_g(\mathbf{c})} p^*(\mathbf{d}) q^*(\mathbf{d}, \mathbf{c}) / p^*(\mathbf{c}_0) \\ &= \sum_{\mathbf{d} \in \mathcal{S}_g(\mathbf{c})} a(\mathbf{d}) q(\mathbf{d}, \mathbf{c}) / b^* p^*(\mathbf{c}_0). \end{aligned} \tag{28}$$

The second equality in (28) follows from (22) and (23). The lemma now follows from (27) and (28) since $E[A_g] b^* p^*(\mathbf{c}_0)$ does not depend on $\mathbf{c}$.

We now compute the summation appearing in (26). Let $L(g) = \{l : c_{s(g)}(l) = g\}$. Any state $\mathbf{d} \in \mathcal{S}_g(\mathbf{c})$ is of the form $\mathbf{c}(s(g), l; h, m)$ for some $l \in L(g)$ and $h \in C[r(g)]$. Therefore,

$$\begin{aligned} &\sum_{\mathbf{d} \in \mathcal{S}_g(\mathbf{c})} a(\mathbf{d}) q(\mathbf{d}, \mathbf{c}) \\ &= \sum_{l \in L(g)} \sum_{h \in C[r(g)]} \sum_{m=1}^{n_{s(h)}+1} \{ a[\mathbf{c}(s(g), l; h, m)] \gamma_{s(h)}(m, n_{s(h)} + 1) \\ &\qquad \times \phi_s(n_{s(h)} + 1) \lambda_h p_{hg}[r(g)] \delta_{s(g)}(l, n_{s(g)}) \}. \end{aligned} \tag{29}$$

It follows from (12) and (13) that

$$a[\mathbf{c}(s(g), l; h, m)] = a(\mathbf{c}) (\lambda_g \phi_{s(g)}(n_{s(g)}) / e_g[r(g)]) [e_h[r(g)] / \lambda_h \phi_{s(h)}(n_{s(h)} + 1)]. \tag{30}$$

Substituting (30) into (29), it is easy to show that

$$p_g(\mathbf{c}) = \left[a(\mathbf{c}) \lambda_g \phi_{s(g)}(n_{s(g)}) \sum_{l \in L(g)} \delta_{s(g)}(l, n_{s(g)}) \right] \Big/ b_g \qquad \mathbf{c} \in \mathcal{S}_g. \tag{31}$$

It follows from (3)–(5) that for any $\mathbf{c} \in \mathcal{S}_g$

$$\sum_{l \in L(g)} \delta_{s(g)}(l, n_{s(g)}) = \begin{cases} 1 \text{ if service center } s(g) \text{ is type 1 or 3} \\ n(g)/n_{s(g)} \text{ if service center } s(g) \text{ is type 2.} \end{cases} \tag{32}$$

Let

$$\mathcal{N}_g = \{\boldsymbol{n} \in \mathcal{N} : n(g) \geqq 1\}. \tag{33}$$

For any $\boldsymbol{n} \in \mathcal{N}_g$ let

$$\mathcal{S}(g, \boldsymbol{n}) = \{\boldsymbol{c} \in \mathcal{S}_g : \boldsymbol{m}(\boldsymbol{c}) = \boldsymbol{n}\} \tag{34}$$

($\boldsymbol{m}(\boldsymbol{c})$ is the population state corresponding to detailed state $\boldsymbol{c}$) and

$$\pi_g(\boldsymbol{n}) = \sum_{\boldsymbol{c} \in \mathcal{S}(g,\boldsymbol{n})} p_g(\boldsymbol{c}). \tag{35}$$

$\{\pi_g(\boldsymbol{n}) : \boldsymbol{n} \in \mathcal{N}_g\}$ is the unique stationary distribution over the set of population states at arrival instants to class g. Note that $p_g(\boldsymbol{c})$ has the same value for all $\boldsymbol{c} \in \mathcal{S}(g, \boldsymbol{n})$ and that

$$|\mathcal{S}(g, \boldsymbol{n})| = \left[\prod_{s=1}^{S} n_s! \prod_{c \in C_s} 1/n(c)!\right] \begin{cases} n(g)/n_{s(g)} & \text{if service center } s(g) \text{ is type 1 or 3} \\ 1 & \text{if service center } s(g) \text{ is type 2.} \end{cases} \tag{36}$$

Substituting (12), (13) and (32) into (31), it follows from (35) and (36) that

$$\pi_g(\boldsymbol{n}) = \alpha(\boldsymbol{n}) \lambda_g \phi_{s(g)}(n_{s(g)}) n(g)/(n_{s(g)} b_g) \qquad \boldsymbol{n} \in N_g, \tag{37}$$

where $\alpha(\boldsymbol{n})$ was given in (19) amd (20). Let $\boldsymbol{n}(g)$ denote the population state obtained from $\boldsymbol{n}$ by decreasing $n(g)$ by 1. Since $e_g[r(g)]/b_g$ does not depend on $\boldsymbol{n}$, (37) can be rewritten in the useful form

$$\pi_g(\boldsymbol{n}) = \alpha[\boldsymbol{n}(g)]/\beta_g \qquad \boldsymbol{n} \in \mathcal{N}_g, \tag{38}$$

where

$$\beta_g = \sum_{\boldsymbol{n} \in \mathcal{N}_g} \alpha[\boldsymbol{n}(g)]. \tag{39}$$

Let $\mathcal{N}'$ denote the set of population states and let $\{\boldsymbol{\pi}'(\boldsymbol{n}') : \boldsymbol{n}' \in \mathcal{N}'\}$ denote the stationary distribution over $\mathcal{N}'$ for the network with one less customer of type $\boldsymbol{r}$. Then $\{\boldsymbol{n}(g) : \boldsymbol{n} \in \mathcal{N}_g\} = \mathcal{N}^{r(g)}$ and it follows from (18), (21), (38) and (39) that

$$\pi_g(\boldsymbol{n}) = \pi^{r(g)}[\boldsymbol{n}(g)] \qquad \boldsymbol{n} \in \mathcal{N}_g. \tag{40}$$

This is the desired result, relating the stationary population state probabilities at arrival instants to the stationary population state probabilities at arbitrary times. (Note from (31) and (32) that a similar result does not hold for the stationary detailed state probabilities if service center $s(g)$ is type 2.)

In Appendix A we remove the restriction $p_{gg}[r(g)] = 0$ and show that $\pi_g(\boldsymbol{n})$ is still given by (40). At the beginning of Section 3 we restricted service demands at type 2 and type 3 service centers to be exponentially distributed. (Service demands at type 1 service centers are exponentially distributed by definition.) In Appendix B we show that $\pi_g(\boldsymbol{n})$ is still given by (40) and that $\pi(\boldsymbol{n})$ is still given by (18)–(21) if this restriction is removed.

5. Applications

Consider service center s and customer type r, where $C_s(r)$ is not empty. For $g \in C_s(r)$ and $k \geqq 1$ let $\pi_g(k)$ denote the stationary probability that k customers are in service center s at arrival instants to class g, i.e.,

$$\pi_g(k) = \sum_{\boldsymbol{n} \in \mathcal{N}_g : n_s = k} \pi_g(\boldsymbol{n}) \qquad k \geqq 1. \tag{41}$$

It follows from (40) and (41) that

$$\pi_g(k) = \pi'(s, k-1) \qquad k \geqq 1 \tag{42}$$

where $\pi'(s, k-1)$ is the stationary probability that $k-1$ customers are in service center s for the network with one less customer of type r. Note that this probability is the same for all $g \in C_s(r)$. Thus, if we let $\pi_{r,s}(k)$ denote the stationary probability that k customers are in service center s at arrival instants to the set of classes $C_s(r)$, i.e., at arrival instants of type r customers at service center s, then

$$\pi_{r,s}(k) = \pi'(s, k-1) \qquad k \geqq 1, \tag{43}$$

i.e., the marginal queue-size probabilities at these arrival instants are obtained directly from the marginal queue-size probabilities (at arbitrary times) for the network with one less customer of type r. Reiser (1977) presents algorithms to compute such marginal queue-size probabilities.

Suppose service center s is type 1, (6) holds and $\phi_s(n) = \min(m, n)$, i.e., there are m servers, the queueing discipline is first-come–first-served and each server has unit service rate. Since the service center is type 1 all service demands have the same exponential distribution independent of the class and customer type. Thus, the service center is a standard multiserver queue with exponential service times. The stationary waiting time (where waiting time includes time in service) distribution $F_{r,s}(t)$, $t \geqq 0$, for type r customers at service center s can therefore be expressed as follows:

$$\begin{aligned} F_{r,s}(t) &= \sum_{k=1}^{m} \pi_{r,s}(k) \Pr\{T_0(\lambda) \leqq t\} \\ &\quad + \sum_{k>m} \pi_{r,s}(k) \Pr\{T_0(\lambda) + T_1(m\lambda) + \cdots + T_{k-m}(m\lambda) \leqq t\} \qquad t \geqq 0, \end{aligned} \tag{44}$$

where λ is the rate parameter of the exponential service times, $T_i(z)$ is an exponential random variable with rate parameter z and $T_0(\lambda)$, $T_1(m\lambda), \cdots$, are statistically independent. It follows from (43) and (44) that the mean stationary waiting time, denoted $w_{r,s}$, is given by

$$w_{r,s} = (1/\lambda) + (1/m\lambda) \sum_{k \geqq m} \pi^r(s, k)(k - m + 1). \tag{45}$$

This expression was derived in Reiser and Lavenberg (1978) by other means, and was used along with suitable applications of Little's formula to obtain a new algorithm for the recursive computation of mean waiting times and mean queue sizes.

Appendix A

For ease of notation we drop the customer type designation $r(g)$ from all probabilities. Suppose $p_{gg} > 0$. We replace class g by two classes g_1 and g_2 and let $\lambda_{g_1} = \lambda_{g_2} = \lambda_g$. Instead of entering class g, customers now enter either class g_1 or g_2 as determined by the transition probabilities given below. We let C' denote the expanded set of classes for customer type $r(g)$ and define a transition probability matrix $P' = (p_{cd} : c \in C',\ d \in C')$ as follows:

$$p'_{g_1 d} = \begin{cases} 0 & d = g_1 \\ p_{gg} & d = g_2 \\ p_{gd} & d \neq g_1, g_2 \end{cases} \tag{46}$$

$$p'_{g_2 d} = \begin{cases} p_{gg} & d = g_1 \\ 0 & d = g_2 \\ p_{gd} & d \neq g_1, g_2 \end{cases} \tag{47}$$

and for $c \neq g_1, g_2$

$$p'_{cd} = \begin{cases} p_{cg}/2 & d = g_1 \text{ or } g_2 \\ \\ p_{cd} & d \neq g_1, g_2. \end{cases} \tag{48}$$

In what follows primed quantities refer to the network with class g replaced by classes g_1 and g_2 and unprimed quantities refer to the original network.

For any $\boldsymbol{n} \in \mathcal{N}_g$, let $\pi'_g(\boldsymbol{n})$ denote the stationary probability of entering state $\boldsymbol{n}$ at arrival instants at the set of classes $\{g_1, g_2\}$, where component $n(g)$ of $\boldsymbol{n}$ is the number of customers in this set of classes. Due to the way P' was defined and since $\lambda_{g_1} = \lambda_{g_2} = \lambda_g$, it is clear that $\pi_g(\boldsymbol{n}) = \pi'_g(\boldsymbol{n})$. We will show that $\pi'_g(\boldsymbol{n})$ is given by (40).

Since $p'_{g_1g_1} = p'_{g_2g_2} = 0$, it follows directly from the results in Section 4 that for $i = 1$ or 2

$$\pi'_{g_i}(\boldsymbol{n}') = \alpha[\boldsymbol{n}'(g_i)]/\beta'_{g_i} \qquad \boldsymbol{n} \in \mathcal{N}'_{g_i}. \tag{49}$$

(Note that $\boldsymbol{n}'$ has components $n'(g_1)$ and $n'(g_2)$ corresponding to classes g_1 and g_2.) For $\boldsymbol{n} \in \mathcal{N}_g$, let $\pi'_{g_i}(\boldsymbol{n})$ denote the sum of $\pi'_{g_i}(\boldsymbol{n}')$ over all $\boldsymbol{n}' \in \mathcal{N}'_{g_i}$ such that $n'(g_1) + n'(g_2) = n(g)$ and $n'(c) = n(c)$, $c \neq g_1, g_2$. From the definition of P' it follows that $e'_{g_1} = e'_{g_2} = e_g/2$ and $e'_c = e_c$, $c \neq g_1, g_2$. Using this result and $\lambda_{g_1} = \lambda_{g_2} = \lambda_g$, it is straightforward to compute the sum and show that for $i = 1$ or 2,

$$\pi'_{g_i}(\boldsymbol{n}) = \alpha[\boldsymbol{n}(g)]/\beta_g \qquad \boldsymbol{n} \in \mathcal{N}_g, \tag{50}$$

i.e., $\pi'_{g_i}(\boldsymbol{n})$ is given by (38) and, hence, (40). The above is a stationary probability at arrival instants at the single class g_i. Recall that $\pi'_g(\boldsymbol{n})$ is a stationary probability at arrival instants at the set of classes $\{g_1, g_2\}$. However, since $\pi'_{g_1}(\boldsymbol{n}) = \pi'_{g_2}(\boldsymbol{n})$, $\pi'_g(\boldsymbol{n})$ is also given by (40).

Appendix B

Suppose for some class, say class 1, at a type 2 or type 3 service center, the service demands are not exponentially distributed, but rather their distribution is a mixture of two gamma distributions each having shape parameter equal to 2. Thus, a service demand in class 1 has four exponential stages of service. A customer which enters class 1 immediately enters stage 1 with probability w and when stage 1 is completed immediately enters stage 2, or the customer immediately enters stage 3 with probability $1 - w$ and when stage 3 is completed immediately enters stage 4. We replace class 1 by four classes, denoted $(1,1)$, $(1,2)$, $(1,3)$, $(1,4)$, one for each exponential stage of service. The service demands for the expanded classes are exponentially distributed with rate parameters equal to those of the corresponding exponential stages of service and denoted $\lambda_{(1,i)}$, $i = 1, 2, 3, 4$. For ease of notation we drop the customer type designation $r(1)$ from all probabilities. We let C' denote the expanded set of classes for customer type $r(1)$ and define a transition probability matrix $P' = (p'_{cd} : c \in C',\ d \in C')$ as follows: for $c \neq (1,1), (1,2), (1,3), (1,4)$

$$p'_{cd} = \begin{cases} wp_{c1} & d = (1,1) \\ (1-w)p_{c1} & d = (1,3) \\ 0 & d = (1,2) \text{ or } (1,4) \\ p_{cd} & \text{otherwise,} \end{cases} \tag{51}$$

$$p'_{(1,1)(1,2)} = p'_{(1,3)(1,4)} = 1, \tag{52}$$

$$p'_{(1,2)d} = p'_{(1,4)d} = \begin{cases} 0 & d = (1,1), (1,2), (1,3) \text{ or } (1,4) \\ p_{1d} & \text{otherwise.} \end{cases} \tag{53}$$

Note that at a type 2 or type 3 service center a customer which completes service in class $(1, 1)$ (corrresponding to stage 1 of service) immediately begins to receive service at the same rate in class $(1, 2)$ (corresponding to stage 2) so that service is received exactly as it would be without the expanded classes. In what follows primed quantities refer to the network with expanded classes and unprimed quantities refer to the original network. For any $\boldsymbol{n} \in \mathcal{N}$, let $\pi'(\boldsymbol{n})$ denote the stationary probability (at arbitrary times) of being in state $\boldsymbol{n}$, where component $\boldsymbol{n}(1)$ of $\boldsymbol{n}$ is the number of customers in the four expanded classes. Also, let $\pi_1'(\boldsymbol{n})$ denote the stationary probability of entering state $\boldsymbol{n}$ at arrival instants at the set of classes $\{(1, 1), (1, 3)\}$. Due to the way the expanded classes were defined, it is clear that $\pi(\boldsymbol{n}) = \pi'(\boldsymbol{n})$ and $\pi_1(\boldsymbol{n}) = \pi_1'(\boldsymbol{n})$. We now show that $\pi'(\boldsymbol{n})$ is given by (18)–(21) and $\pi_1'(\boldsymbol{n})$ is given by (40).

Since the network with expanded classes has only exponential service demands, if follows directly from the results in Section 3 that

$$\pi'(\boldsymbol{n}') = \alpha(\boldsymbol{n}')/b' \qquad \boldsymbol{n}' \in \mathcal{N}'. \tag{54}$$

(Note that $\boldsymbol{n}'$ has a component $n'(1, i)$ for class $(1, i)$, $i = 1, 2, 3, 4$.) Then for $\boldsymbol{n} \in \mathcal{N}$, $\pi'(\boldsymbol{n})$ is the sum of $\pi'(\boldsymbol{n}')$ over all $\boldsymbol{n} \in \mathcal{N}'$ such that $\Sigma_{i=1}^4 n'(1, i) = n(1)$ and $n'(c) = n(c)$, $c \neq (1, i)$, $i = 1, 2, 3, 4$. It follows from the definition of P' that $e'_{(1,1)} = e'_{(1,2)} = we_1/(1 + e_1)$, $e'_{(1,3)} = e'_{(1,4)} = (1 - w)e_1/(1 + e_1)$ and $e'_c = e_c/(1 + e_1)$, $c \neq (1, i)$, $i = 1, 2, 3, 4$. The mean service demand, $1/\lambda_1$, in class 1 is equal to $w[(1/\lambda_{(1,1)}) + (1/\lambda_{(1,2)})] + (1 - w)[(1/\lambda_{(1,3)}) + (1/\lambda_{(1,4)})]$. Using these results it is straightforward to compute $\pi'(\boldsymbol{n})$ and show that for $\boldsymbol{n} \in \mathcal{N}$, $\pi'(\boldsymbol{n}) = \alpha(\boldsymbol{n})/b$, the desired result. The result for $\pi_1'(\boldsymbol{n})$ is obtained in a similar manner except that it is necessary to compute the probabilities $\pi'_{(1,1)}(\boldsymbol{n})$ and $\pi'_{(1,3)}(\boldsymbol{n})$ at arrival instants at classes $(1, 1)$ and $(1, 3)$ respectively. These turn out to equal $\alpha[\boldsymbol{n}(1)]/\beta_1$, and hence, are given by (40). Thus, $\pi_1'(\boldsymbol{n})$ is also given by (40).

The method of proof we have used, i.e., appropriately expanding the set of classes, can be applied when the distributions of service demands for one or more classes are any finite mixture of gamma distributions, to establish that $\pi(\boldsymbol{n})$ is given by (18)–(21) and $\pi_1(\boldsymbol{n})$ is given by (40).

References

BARBOUR, A. D. (1976) Networks of queues and the method of stages, *Adv. Appl. Prob.* **8**, 584–591.

BASKETT, F., CHANDY, K. M., MUNTZ, R. R. AND PALACIOS, F. G. (1975) Open, closed and mixed networks of queues with different classes of customers. *J. Assoc. Comput. Mach.* **22**, 248–260.

KELLY, F. P. (1975) Networks of queues with customers of different types. *J. Appl. Prob.* **12**, 542–554.

KELLY, F. P. (1976) Networks of queues. *Adv. Appl. Prob.* **8**, 416–432.

LEMOINE, A. J. (1977) Networks of queues — A survey of equilibrium analysis. *Management Sci.* **24**, 464–481.

MELAMED, B. (1978) On Markov jump processes imbedded at traffic epochs and their queueing-theoretic interpretation. Department of Industrial Engineering and Management Sciences, Northwestern University.

REISER, M. (1977) Numerical methods in separable queueing networks. *Studies in the Management Sciences* 7, 113–142.

REISER, M. AND LAVENBERG, S. S. (1978) Mean value analysis of closed multichain queueing networks. IBM Research Report RC 7023, Yorktown Heights, New York.

SEVCIK, K. C. AND MITRANI, I. (1979) The distribution of queuing network states at input and output instants. *Proc. 4th Internat. Symp. Modelling and Performance Evaluation of Computer Systems.* North-Holland, Amsterdam.