RAIRO. RECHERCHE OPÉRATIONNELLE

H. G. PERROS

A two-level open queue network with blocking and feedback

RAIRO. Recherche opérationnelle, tome 15, nº 1 (1981), p. 27-38

<http://www.numdam.org/item?id=RO_1981__15_1_27_0>

R.A.I.R.O. Recherche opérationnelle/Operations Research
(vol. 15, n° 1, février 1981, p. 27 à 38)

A TWO-LEVEL OPEN QUEUE NETWORK WITH BLOCKING AND FEEDBACK (*)

by H. G. PERROS (1)

Abstract. — *The queue network model studied in this paper consists of two symmetrical queues in parallel served by a first level of servers and linked to a second-level server with no intermediate waitingroom. Blocking of the flow of units through a first-- e el server occurs each time the server completes a service. The server remains blocked and it can not serve any other units until the blocking unit completes its service at the second-level server. An approximate expression of the generating function, g (z), of the queue-length distribution is obtained. The queue-length distribution is then derived by inverting g (z). The results obtained compare very well with simulation data. The exact condition for stability of the queue network is also derived.*

Keywords: Open queue network, blocking, exponential, two-level service, feedback.

Résumé. — *Le modèle de réseau de file d'attente étudié dans cet article consiste en deux files d'attente symétriques en parallèle, servies par des serveurs de premier niveau, et couplées avec un serveur de second niveau, sans salle d'attente intermédiaire. Le blocage du des flot clients à travers un serveur du premier niveau a lieu chaque fois que le serveur termine son servive. Le serveur demeure bloqué et ne peut servir aucun autre client jusqu'à ce que le client ait achevé d'être servi par le serveur du second niveau. Une expression approchée de la fonction génératrice g (z) de la distribution de la longueur de la file d'attente est obtenue. La distribution elle-même en est déduite en inversant g (z). Les résultats obtenus se comparent très bien avec les données de simulation. La condition exacte de stabilité du réseau de file d'attente est aussi déduite.*

1. INTRODUCTION

The concept of open queue network with blocking has proved useful in modelling stochastic systems, such as computer systems, telecommunication systems, production systems (*see* Konheim and Reiser [9], Perros [10], Hillier and Boling [6]). Networks with blocking consist of a set of arbitrarily linked service channels some of which are of limited capacity. Blocking occurs when the flow of units through one channel is momentarily stopped owing to a capacity limitation of another channel having been reached.

Queue networks with blocking have proved difficult to treat in general. The simplest configuration of such a network is the one consisting of two servers in series with a finite intermediate waitingroom. The queue in front of the first server is unlimited in length. A unit completing a service at the second server may be fed back to the end of either queues. Blocking of the first

(*) Receveid June 1979.
(1) Quantitative Methods Dept., University of Illinois, Chicago, Ill. 60680, U.S.A.

server occurs when the intermediate waitingroom becomes full. This model has been studied by various authors under different assumptions regarding the service distribution, priority discipline and feedback. A survey of the literature and new results on this model can be found in Asare [1]. A well known generalization of this network with blocking is the one consisting of k ($k > 2$) servers in tandem. The queue before the first server is allowed to be infinite in size and a finite queue is allowed between successive servers. A survey of the relevant literature and new results can be found in Caseau and Pujolle [2]. Foster and Perros [4, 5] consider a general class of open queue networks with blocking consisting of n parallel symmetrical queues served by independent first-level servers. Groups of these servers are linked with a second level of servers and these are linked in groups with a third level of servers and so on up to any number of servers. The queue in front of each first-level server is unrestricted in length, but no queue is allowed in front of any high level server. When a unit completes its service at a server that server immediately becomes blocked. The server remains blocked until the blocking unit departs from the network having received service by all the high level servers to which the server is linked. Approximate and exact expressions for the traffic intensity, τ, of a first-level server were derived for a two-level and a three-level queue network assuming no feedback. For the general n-level queue network, an expression for τ was derived assuming a processor-sharing type of service at all levels of servers with the exception of the first-level servers. Pittel [11] reported on two closed exponential queue networks with blocking. A unit upon completion of its service selects its next queue according to a probability transition matrix. In the first model, if the queue the unit selects to join is full at that instant the unit returns back to the last queue. In the second model, the unit randomly searches for another queue which is not full. The time the unit spends searching for another queue is assumed to be negligible. An asymptotic expression of the blocking probabilities is obtained using non-linear mathematical programming techniques.

In this paper we examine an exponential open queue network model consisting of two symmetrical queues in parallel served by independent servers (first-level servers) and linked to a server (second-level server) with no intermediate waitingroom. When a unit completes its service at a first-level server that server becomes blocked, i. e. it cannot serve any other units. The server remains blocked until the blocking unit completes its service at the second-level server. Using the terminology introduced in Foster and Perros [5] this model can be described as a *two-level* queue network with blocking and feedback. An approximate expression of the generating function of the equilibrium queue-length distribution is derived. The generating function is then inverted in order to obtain a closed-form solution of the queue-length distribution. This expression is validated using simulation techniques. The exact condition for stability is also derived. We now proceed to examine the queue network in detail.

2. THE TWO-LEVEL QUEUE NETWORK WITH BLOCKING AND FEEDBACK

The queue network studied here is shown in figure 1. Let λ, α, and β be the input rate into each queue, the mean service time at a first-level server, and the mean service time at the second-level server respectively. All inter-arrival times as well as service times are assumed exponentially distributed. The two queues are unlimited in length. No queue is allowed to accumulate in front of the second-level server. Both queues are served on a FIFO basis. A unit upon completion of its service at a first-level server proceeds to receive service at the second-level server. Upon completion of this service the unit returns to its first-level server for further service, or it departs from the queue network with probability θ and $1 - \theta$ respectively. A unit, therefore, may cycle several times between its first-level server and the second-level server before it departs from the network. A first-level server gets *blocked* upon completion of a service. The particular server remains blocked until the unit (call it the blocking unit) completes its service at the second-level server. At that instant, its first-level server becomes unblocked, and if the blocking unit returns to the server for further service, the server becomes busy again. Otherwise, if the blocking unit departs from the network, its first-level server initiates a new service or it becomes idle depending upon whether there is a unit waiting in the queue or not.

During the period of time, therefore, that a particular unit is in service the state of its first-level server cycles between the states "busy serving" and

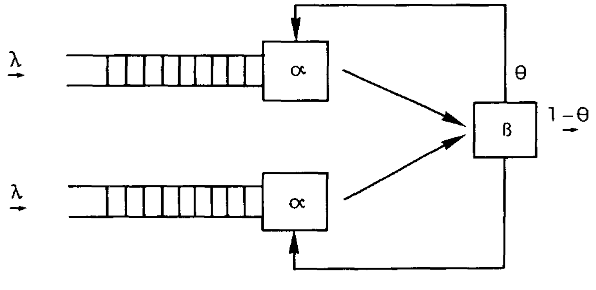

Figure 1. — A two-level queue network with blocking and feedback.

"blocked". Obviously, no more than one unit from each queue may be in service at any time. The maximum number of blocking units, therefore, competing for the second-level server may not exceed 2. The blocking units are served on a "first complete its first-level service first served" basis.

In view of the fact the second-level server is shared by units from both queues, the effective service time of a first-level server is partially dependent upon the activity of the other first-level server. This partial coupling of the two first-level servers makes the queue network difficult to treat. In this paper we analyse this model by studying only one of the two queues. The activity of the other queue is taken into account implicitly. This approach leads to the derivation of an approximate expression of the generating function, $g(z)$, of the queue-length distribution. The closed form solution of the queue-length distribution is then obtained by inverting $g(z)$. The results obtained compare very well with simulation data.

Let π be the conditional probability that a blocking unit upon completion of its service at a first-level server will find the second-level server busy (by a unit of the other queue) and, therefore, it will be forced to wait for its second-level service. Then, the mean blocking time of a first-level server is $\beta(1+\pi)$. The conditional probability π has been derived in Foster and Perros [5] and is given by the expression

$$\pi = (1/2\lambda\beta)\{(1-\theta)-\sqrt{(1-\theta)^2-4\lambda^2\beta^2}\}. \tag{2.1}$$

Let p_{ij}, $i=0, 1, 2, \ldots$, and $j=0, 1, 2, 3$, be the probability that there are i units in a queue waiting to be served (the one in service is not included), and that the first-level server of this queue is in state j. A first-level server at any instant may be in one of the following states: (*a*) idle ($j=0$); (*b*) busy serving ($j=1$); (*c*) blocked but its own blocking unit is waiting to receive its second-level service ($j=2$), and (*d*) blocked and its own blocking unit is receiving service at the second-level server ($j=3$). The steady-state equations involving p_{ij} are given below. These equations were derived assuming that the process $\{X_t, Y_t\}$, $t \geqq 0$, where $X_t = 0, 1, 2, \ldots$, and $Y_t = 0, 1, 2, 3$, is markovian. This assumption is not true as the memoryless property is not always satisfied:

$$\left\{\begin{array}{ll} \lambda p_{00} = ((1-\theta)/\beta)\, p_{03}, & \\ (\lambda+(1/\alpha))\, p_{01} = \lambda p_{00} + (\theta/\beta)\, p_{03} + ((1-\theta)/\beta)\, p_{13}, & \\ (\lambda+(1/\alpha))\, p_{i1} = \lambda p_{i-1,1} + (\theta/\beta)\, p_{i3} + ((1-\theta)/\beta)\, p_{i+1,3}, & i \geqq 1; \end{array}\right.$$

$$\left\{\begin{array}{ll} (\lambda+(1/\beta))\, p_{02} = (\pi/\alpha)\, p_{01}, & \\ (\lambda+(1/\beta))\, p_{i2} = (\pi/\alpha)\, p_{i1} + \lambda p_{i-1,2}, & i \geqq 1; \end{array}\right.$$

$$\left\{\begin{array}{ll} (\lambda+(1/\beta))\, p_{03} = (1/\beta)\, p_{02} + ((1-\pi)/\alpha)\, p_{01}, & \\ (\lambda+(1/\beta))\, p_{i3} = (1/\beta)\, p_{i2} + ((1-\pi)/\alpha)\, p_{i1} + \lambda p_{i-1,3}, & i \geqq 1. \end{array}\right.$$

The normalizing equation is

$$\sum_{i=0}^{\infty} \sum_{j=0}^{3} p_{ij} = 1.$$

Let p_n, $n = 0, 1, 2, \ldots$, be the steady-state probability that there are n units in a particular queue including the one in service. Let $g(z) = \sum_{n=0}^{\infty} z^n p_n$. An expression for the generating function $g(z)$ is obtained in the following section. This expression is approximate seeing that it is derived assuming that the process $\{X_t, Y_t\}$, $t \geqq 0$, is markovian. The closed-form solution of the probability p_n, $n = 0, 1, 2, \ldots$, is obtained in section 5 by inverting the z-transformation $g(z)$. The inversion of $g(z)$ is exact. However, the resulting solution is in itself an approximation to the true queue-length distribution as it is derived using the approximate expression of $g(z)$. The exact condition for stability of the queue network is derived in section 4. Finally, in section 6 we obtain an approximate expression for the mean queue length. The results obtained in this paper compare very well with simulation data. Some minor deviations were observed for an extreme case in which $\beta \ggg \alpha$.

3. THE DERIVATION OF $g(z)$

Define the following generating functions

$$g_k(z) = \sum_{i=0}^{\infty} z^i p_{ik}, \qquad k = 1, 2, 3.$$

Then, the steady-state equations given in section 2 can be expressed as follows:

$$\begin{cases} \beta z [\lambda(1-z)+(1/\alpha)] g_1(z) - [1-(1-z)\theta] g_3(z) = -\beta(1-z)\lambda p_{00}, \\ \pi g_1(z) - \alpha [\lambda(1-z)+(1/\beta)] g_2(z) = 0, \\ \beta(1-\pi) g_1(z) + \alpha g_2(z) - \alpha [\lambda\beta(1-z)+1] g_3(z) = 0. \end{cases}$$

The normalizing equation becomes

$$g_1(1)+g_2(1)+g_3(1) = 1-p_{00}. \tag{3.1}$$

The unknowns $g_1(z)$, $g_2(z)$, $g_3(z)$, and p_{00} can be determined from the above system of linear equations. After some lengthy calculations we obtain

$$g_1(z) = \lambda\alpha(1+\lambda\beta(1-z))^2 p_{00}/(1-\theta)(1+\lambda\beta(1-\pi)) \times(-P z^3+Q z^2-R z+1), \tag{3.2}$$

$$g_2(z) = \lambda\pi\beta(1+\lambda\beta(1-z)) p_{00}/(1-\theta)(1+\lambda\beta(1-\pi)) \times(-P z^3+Q z^2-R z+1), \tag{3.3}$$

and

$$g_3(z) = \lambda\beta(1+\lambda\beta(1-\pi)(1-z))\,p_{00}/(1-\theta)(1+\lambda\beta(1-\pi)) \times(-Pz^3+Qz^2-Rz+1), \quad (3.4)$$

where

$$P = \lambda^3\alpha\beta^2/(1-\theta)(1+\lambda\beta(1-\pi)), \quad (3.5)$$

$$Q = \lambda\beta(2\lambda\alpha(1+\lambda\beta)+\lambda\beta)/(1-\theta)(1+\lambda\beta(1-\pi)), \quad (3.6)$$

$$R = \{\lambda\beta(1-\theta+\pi\theta)+(1+\lambda\beta)(\lambda\alpha+\lambda\beta+\lambda^2\alpha\beta)\}/(1-\theta)(1+\lambda\beta(1-\pi)). \quad (3.7)$$

[Note that expression (2.1) for π is not used in the above expression as we wish to maintain these expressions in a simple algebraic form.] Now, if we consider the limit of each generating function at $z = 1$ and then take the sum of these three limits we obtain

$$g_1(1)+g_2(1)+g_3(1) = \lambda(\alpha+\beta(1+\pi))\,p_{00}/\{(1-\theta)-\lambda(\alpha+\beta(1+\pi))\}.$$

Using the normalizing equation (3.1) in the above expression we have

$$p_{00} = 1-(\lambda/(1-\theta))(\alpha+\beta(1+\pi)). \quad (3.8)$$

This expression for p_{00} is an exact one. For, let us assume that units which complete their second-level service and are fed back for further service join the end of the queues as in a round-robin model. Obviously, the total input rate into each queue is $\lambda/(1-\theta)$ (*see* Jackson [7]). Now seeing that the effective service time of a first-level server is $\alpha+\beta(1+\pi)$, we have that the utilisation of this server is $(\lambda/(1-\theta))(\alpha+\beta(1+\pi))$. Hence, we obtain expression (3.8) as p_{00} is equal to the compliment of the server's utilization.

We now proceed to derive $g(z)$. We have that $g(z) = \sum_{n=0}^{\infty} z^n p_n$, where p_n is the steady-state probability that there are n units in a queue, including the one in service. We have

$$g(z) = p_{00}+z\{g_1(z)+g_2(z)+g_3(z)\}.$$

After some lengthy calculations we can establish that

$$g(z) = \frac{1}{-Pz^3+Qz^2-Rz+1}\left(1-\frac{\lambda\beta(1-\pi)}{1+\lambda\beta(1-\pi)}z\right)p_{00}, \quad (3.9)$$

where p_{00}, π, P, Q, and R are given by (3.8), (2.1), (3.5), (3.6) and (3.7) respectively.

4. CONDITION FOR STABILITY

We define the condition for stability of the queue network as the condition for stability of a queue in the network, i. e.

$$p_{00} > 0. \tag{4.1}$$

Now, if we use expressions (2.1) and (3.8) in (4.1) we can obtain after some calculations the following condition

$$\lambda < (1-\theta)(\alpha+\beta)/((\alpha+\beta)^2+\beta^2). \tag{4.2}$$

This is an exact condition. For $\theta = 0$ expression (4.2) reduces to the condition for stability of the queue network with no feedback derived by Foster and Perros [5] using a different method.

5. DERIVATION OF THE PROBABILITIES $\{p_n\}$, $n = 0, 1, 2, \ldots$

The closed-form solution of the probabilities $\{p_n\}$, $n = 0, 1, 2, \ldots,$ is derived by inverting the z-transformation $g(z)$ given by (3.9). Let us first consider the series

$$\sum_{n=0}^{\infty} (P z^3 - Q z^2 + R z)^n. \tag{5.1}$$

The term $|P z^3 - Q z^2 + R z|$ or $|z|.|P z^2 - Q z + R|$ is obviously less than one for all $|z| < \delta$, where $\delta < 1$. Therefore, the above series converges to

$$1/\{1-(P z^3 - Q z^2 + R z)\} \tag{5.2}$$

for all $|z| < \delta$. We observe that the above limit is identical to the first term of $g(z)$ in (3.9). Now, we have that the generating function $g(z)$ is analytic within the unit circle, and seeing that the numerator of $g(z)$ has one real root greater than one, the denominator may not have any roots within the unit circle. It is clear, therefore, that expression (5.2), which is the limit of the series given by (5.1), exists at all points within the unit circle and the series converges for all $|z| < 1$.

Now, let us consider one term only of the series (5.1). We have

$$\left.\begin{aligned} (P z^3 - Q z^2 + R z)^n &= \Sigma \frac{n!}{k!\,l!\,m!} (P z^3)^k (-Q z^2)^l (R z)^m, \\ n &= 0, 1, 2, \cdots, \end{aligned}\right\} \tag{5.3}$$

where the summation extends over all non-negative integers k, l, m such that $k+l+m=n$ (*see* Feller [3], p. 66). Expression (5.3) can be written as follows

$$\left.\begin{aligned}(P z^3 - Q z^2 + R z)^n = \Sigma(-1)^l \frac{n!}{k!\,l!\,m!} P^k Q^l R^m z^{3k+2l+m},\\ n = 0, 1, 2, \ldots,\end{aligned}\right\} \quad (5.4)$$

for all k, l, m as in (5.3). Now, if we substitute (5.4) in (5.1) and also use expression (5.2) we obtain

$$1/\{1-(P z^3 - Q z^2 + R z)\} = \sum_{n=0}^{\infty} \Sigma(-1)^l \frac{n!}{k!\,l!\,m!} P^k Q^l R^m z^{3k+2l+m},$$

where the second summation extends over all k, l, m as in (5.3). Now, let us consider the sum of all terms in the right-hand side of the above summation which correspond to the same power of z, say z^n. This is given by the following expression

$$\Sigma(-1)^l \frac{(k+l+m)!}{k!\,l!\,m!} P^k Q^l R^m z^{3k+2l+m},$$

where the summation extends over all non-negative integers k, l, m such that $3k+2l+m=n$. Let

$$A_n = \Sigma(-1)^l \frac{(k+l+m)!}{k!\,l!\,m!} P^k Q^l R^m, \qquad n = 0, 1, 2, \ldots, \qquad (5.5)$$

where the summation extends over all non-negative integers k, l, m such that $3k+2l+m=n$. Then, we have

$$1/\{1-(P z^3 - Q z^2 + R z)\} = \sum_{n=0}^{\infty} A_n z^n, \qquad (5.6)$$

where A_n is given by (5.5). The generating function $g(z)$ can now be expanded into the following series

$$g(z) = p_0 (1 - B z) \sum_{n=0}^{\infty} A_n z^n, \qquad 5.7)$$

where $p_0 = p_{00}$ and $B = \lambda\beta(1-\pi)/(1+\lambda\beta(1-\pi))$. Now, since $A_0 = 1$ expression (5.7) can be written as follows

$$g(z) = p_0 + \sum_{n=1}^{\infty} p_0 (A_n - A_{n-1} B) z^n. \qquad (5.8)$$

Therefore,

$$p_n = p_0(A_n - B A_{n-1}), \qquad n \geqq 1, \tag{5.9}$$

where $B = \lambda\beta(1 - \pi)/(1 + \lambda\beta(1 - \pi))$, A_n, $n = 0, 1, 2, \ldots$, is given by (5.5) and the quantities p_0 and π are given by (3.8) and (2.1) respectively.

A very simple recursive formula for deriving the quantity A_n, $n = 0, 1, 2, \ldots$, can be obtained by deriving the exact closed-form solution of the probabilities $\{p_{i1}\}$, $\{p_{i2}\}$, and $\{p_{i3}\}$, $i = 0, 1, 2, \ldots$ This can be achieved by following similar arguments as above. In particular, using (5.6) in (3.2) we can expand $g_1(z)$ into a series from which $\{p_{i1}\}$, $i = 0, 1, \ldots$, can be readily obtained. We have

$$g_1(z) = \{\lambda\alpha(1+\lambda\beta)^2/(1-\theta)(1+\lambda\beta(1-\pi))\}\left(1-\frac{\lambda\beta}{1+\lambda\beta}z\right)^2 \sum_{n=0}^{\infty} A_n z^n,$$

which can be expressed as follows

$$g_1(z) = K_1\left\{1+(R-2C)z+\sum_{n=2}^{\infty}(A_n-2CA_{n-1}+C^2A_{n-2})z^n\right\}p_0, \tag{5.10}$$

where $K_1 = \lambda\alpha(1+\lambda\beta)^2/(1-\theta)(1+\lambda\beta(1-\pi))$ and $C = \lambda\beta/(1+\lambda\beta)$. Similarly, the generating functions $g_2(z)$ and $g_3(z)$ can be expanded into the following series

$$g_2(z) = K_2\left\{1+\sum_{n=1}^{\infty}(A_n-CA_{n-1})z^n\right\}p_0, \tag{5.11}$$

and

$$g_3(z) = K_3\left\{1+\sum_{n=1}^{\infty}(A_n-BA_{n-1})z^n\right\}p_0, \tag{5.12}$$

where $K_2 = \lambda\pi\beta(1+\lambda\beta)/(1-\theta)(1+\lambda\beta(1-\pi))$ and $K_3 = \lambda\beta/(1-\theta)$. The quantities A_n, $n = 0, 1, 2, \ldots$, B, R, p_0 and π are defined as above. Now, from expressions (5.10), (5.11) and (5.12) we have

$$\left.\begin{aligned} p_{01} &= K_1 p_0, \\ p_{11} &= K_1(R-2C)p_0, \\ p_{i1} &= K_1(A_i-2CA_{i-1}+C^2A_{i-2})p_0, \qquad i \geqq 2, \end{aligned}\right\} \tag{5.13}$$

$$\left.\begin{aligned} p_{02} &= K_2 p_0, \\ p_{i2} &= K_2(A_i-CA_{i-1})p_0, \qquad i \geqq 1, \end{aligned}\right\} \tag{5.14}$$

and

$$\left.\begin{aligned} p_{03} &= K_3 p_0, \\ p_{i3} &= K_3(A_i-BA_{i-1})p_0, \qquad i \geqq 1. \end{aligned}\right\} \tag{5.15}$$

The quantity A_n, $n = 0, 1, 2, \ldots$, can now be obtained by equating (5.9) with (5.13), (5.14), and (5.15). We have that

$$p_n = p_{n-1,1} + p_{n-1,2} + p_{n-1,3}, \qquad n \geqq 1. \tag{5.16}$$

Also, we can easily establish that $B + (K_1 + K_2 + K_3) = R$, $2K_1C + K_2C - K_3B = Q$ and $K_1C^2 = P$. Using these expressions as well as (5.9), (5.13), (5.14) and (5.15) in (5.16) we obtain

$$A_n = R\,A_{n-1} - Q\,A_{n-2} + P\,A_{n-3}, \qquad n \geqq 1. \tag{5.17}$$

(It is assumed that for n negative $A_n = 0$.) For $n = 0$ we have $A_0 = 1$. The quantities R, Q, and P are given by (3.7), (3.6), and (3.5) respectively.

For the limiting case of $\alpha \to 0$, $\beta \to 0$ and $\theta \to 1$ while $a/(1 - \theta)$ and $\beta/(1 - \theta)$ remain finite we obtain a processor-sharing type of model. In this case, a unit in service cycles infinitely quickly between its first-level server and the second-level server receiving an infinitesimal amount of service at each server (*see* Kleinrock [8]). Let $g^*(z)$ be the limit of $g(z)$ as given by (3.9). We observe that

$$g^*(z) = p_0/\{1 - (\rho_1 + \rho_2(1+\pi))\,z\}, \tag{5.18}$$

where p_0 and π are given by (3.8) and (2.1) respectively and $\rho_1 = \lambda\alpha/(1 - \theta)$, $\rho_2 = \lambda\beta/(1 - \theta)$. Let p_n^* be the limit of p_n, $n = 0, 1, 2, \ldots$ Then p_n^* can be readily obtained from (5.18). We have

$$p_n^* = (\rho_1 + \rho_2(1+\pi))^n\, p_0, \qquad n \geqq 1.$$

We observe that for this limiting case the approximate expression for $g(z)$ yields a queue length distribution that coincides with that of an $M/M/1$ queue.

6. THE MEAN NUMBER OF UNITS IN A QUEUE

An approximate expression for the mean number of units, N, in a queue, including the one in service, can be obtained by employing the quantity $g'(1)$. After some tedious calculations we can establish that

$$N = \{(\lambda\alpha + \lambda\beta(1+\pi)) - \lambda\beta(\lambda\alpha + (\lambda\alpha + \lambda\beta\pi)\,\pi)\}/\{(1-\theta) - (\lambda\alpha + \lambda\beta(1+\pi))\}, \tag{6.1}$$

where π is given by (2.1).

7. DISCUSSION

Queue networks with blocking have proved useful in modelling stochastic systems. However, the blocking process associated with these networks seriously complicates the mathematical analysis of these models. As a result, queue networks with blocking have proved difficult to treat in general. The simplest configuration of such a network is the one consisting of two servers in tandem with an intermediate finite waitingroom. This model has been studied by several authors. However, it is interesting to note that the exact closed-form solution of the joint queue-length distribution of this model assuming exponentially distributed inter-arrival and service times has not as yet been reported in the literature.

In this paper a two-level queue network model with blocking is studied. An approximate expression of the generating function $g(z)$ of the equilibrium queue-length distribution p_n, $n = 0, 1, 2, \ldots$, is obtained by analysing only one of the two queues. The activity of the other queue is taken into account implicitly. The closed-form solution of p_n, $n = 0, 1, 2, \ldots$, is then obtained by inverting $g(z)$. The inversion of $g(z)$ is exact. However, the resulting solution is in itself an approximation to the true queue-length distribution as it is derived using the approximate expression of $g(z)$. The quantity p_n, $n = 0, 1, 2, \ldots$, given by (5.9) as well as the mean number of units in a queue given by (6.1) were calculated for various values of λ, α, β and θ. The results obtained were then compared against simulation data. The simulation model was written in GPSS. No significant deviations between the numerical results and the simulation data were observed with the exception of the extreme case in which $\beta \rangle\rangle\rangle \alpha$. In this case some minor deviations were observed.

REFERENCES

1. B. K. Asare, *Queue Networks with Blocking*, Ph. D. Thesis, Trinity College, Dublin, 1978.
2. P. Caseau and G. Pujolle, *Throughput Capacity of a Sequence of Queues with Blocking Due to Finite Waiting Room*, I.E.E.E., Software Engineering Trans., 1979, pp. 631-642.
3. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, Wiley, New York, 1968.
4. F. G. Foster and H. G. Perros, *Hierarchical Queue Networks with Partially Shared Servicing*, J. of Opl. Res. Soc., Vol. 30, 1979, pp. 157-166.
5. F. G. Foster and H. G. Perros, *On the Blocking Process in Queue Networks*, to appear in the European J. of Opl. Res.

6. F. S. HILLIER and R. W. BOLING, *Finite Queues in Series with Exponential or Erlang Service Times-A Numerical Approach*, Opns. Res., Vol. 15, 1967, pp. 283-303.
7. J. R. JACKSON, *Networks of Waiting Lines*, Opns. Res., Vol. 5, 1957, pp. 518-521.
8. L. KLEINROCK, *Queueing Systems*, Vol. 2, 1975, Wiley, New York.
9. A. G. KONHEIM and M. REISER, *A Queueing Model with Finite Waiting Room and Blocking*, J. A.C.M., Vol. 23, 1976, pp. 328-344.
10. H. G. PERROS, *Modelling a Disc I/O System: A Queue Network Approach*, European J. of Opl. Res., Vol. 1, 1977, p. 139-145.
11. B. PITTEL, *Closed Exponential Networks of Queues with Blocking. The Jackson-Type Stationary Distribution and its Asymptotic analysis*, I.B.M. Watson Research Center, report RC 6174 (# 26548), 1976.