PROBABILISTIC MODELS OF COMPUTER SYSTEMS

E. GELENBE and G. PUJOLLE
IRIA - LABORIA
BP 5.
78150 LE CHESNAY
FRANCE

Abstract

We develop a method based on diffusion approximations in order to compute, under some general conditions, the queue length distribution for a queue in a network. Applications to computer networks and to time-sharing systems are presented.

Introduction

Though considerable progress has been made in obtaining exact solutions for large classes of queueing network models [1, 2] , one particularly simple type of network, an arbitrary network with first-come-first-served (FCFS) service discipline and general distribution function of service time at the servers, has proved to be resilient to all approaches except for approximate solution techniques. Here our attention is limited to this type of queueing network.

Several approximation methods have been suggested for its treatment. On the one hand there are diffusion approximations [3, 4, 5] applicable to two-station networks or to general queueing networks [6, 7], and on the other hand we have iterative techniques [8]. The convergence of the latter to the exact solution is not an established fact and we know that the former tend, in certain simple cases, to the exact solution under heavy traffic assumptions.

Although most of the work published in the literature has concentrated on evaluating the joint probability distribution of queue lengths for all the queues in a network, it is seldom possible to make use of this complete information. In measurements on computer systems it is difficult enough to collect data on the performance of a single resource, and the measurement of joint data for several resources could become very time and space consuming. The same can be said of simulation experiments where the task of estimating confidence intervals for estimated joint statistics becomes impractical. Furthermore when it comes to computing average response times it suffices to know the average response time ancountered in each individual queue. Therefore it would suffice in most cases to be able to compute within satisfactory accuracy the probability distribution for the queue length in each individual queue in the network.

In this note we present an approximation method using a diffusion model to obtain the stationary probability of queue length of any given queue in an open or closed queueing network composed of FCFS service stations, each composed of a single server with general service time distribution. The method is applied to a simple example of interest and the model predictions are compared with simulation results.

The approximation technique we propose is based on assuming that the output from any queue in the network forms a renewal process. First consider the general network of Figure 1 in which :

(i):external arrivals constitute a renewal process of rate λ_o ; the variance of the interarrival time is V_o.

(ii):the transition of customers form one station to another is defined by a first order Markov chain with transition matrix $P=(p_{ij}), p_{ij}, 1 \leq i, j \leq n+1$, being the probability that a customer having terminated its service at station i then enters station j, or leaves the system when $j = n+1$; P is assumed to have a single absorbing state n+1.

(iii): the service times for successive customers at station i are independent and identically distributed with common distribution function $F_i(t)$; service times are also independent from one station to the other.

(iv): customers first entering the network are directed to station i with fixed probability p_{oi}.

Let e_i, $1 \leq i \leq n$, be the solution to the system of equations

$$e_i = p_{oi} + \sum_1^n e_j p_{ji}$$

which is unique under the assumtions we have made. Then e_i is the expected number of visits which a customer of the network will make to station i. The arrival rate of customers to station i is then λe_i at steady-state ; also the steady-state probability u_i that station i contains at least one customer is given by

$$u_i = \frac{\lambda_o e_i}{\mu_i} \text{ if } \lambda e_i < \mu_i \text{ where}$$

$$\mu_i^{-1} = \int_0^\infty t \, dF_i(t)$$

is the average service time for a customer at station i. These facts can be easily established rigorously ; one way is to treat an open network of this kind as a limiting case of a closed network when one station is saturated and to apply the work-rate theorem [9].

Diffusion approximation for a single queue

In this section we briefly review some material on diffusion approximations which will be used in this paper [3, 4, 5].

Consider a queue with FCFS service and suppose that at time t = 0 a busy period of the queue begins. Interarrival times are independent and identically distributed, of mean and variance λ^{-1} and V_a, respectively. Service times are independent of interarrival times and of each other, and are identically distributed with mean and variance μ^{-1} and V_s respectively. Let A(t), D(t) be the total number of arrivals and departures up to time t, where t is some instant during the busy period beginning at t = 0. Then for t large enough, A(t), D(t) will be approximately normally distributed with mean λt, μt respectively, and variance $\lambda^3 V_a t$, $\mu^3 V_s t$, respectively. Thus N(t), the number in queue at time t, will be approximately normal with mean $(\lambda-\mu)t$ and variance $(\lambda^3 V_a + \mu^3 V_s)t$, since N(t) = A(t) - D(t). Thus N(t) may be approximated by a continuous random variable X(t) whose probability density function

$$f(x, t)\, dx = \Pr[x \leq X(t) < x + dx]$$

satisfies, for X(t) > 1 the diffusion equation

$$\frac{\partial f}{\partial t} - b\frac{\partial f}{\partial x} + \frac{1}{2}\alpha \frac{\partial^2 f}{\partial x^2} = 0$$

where $b = \lambda - \mu$, $\alpha = \lambda^3 V_a + \mu^3 V_s$. For values of X(t) in [0, 1] we follow the approach in [5]. Let P(t) be the probability that X(t) = 0. Then for X(t) on the non-negative real line we have

$$\left.\begin{aligned} &\frac{\partial f}{\partial t} - b\frac{\partial f}{\partial x} + \frac{1}{2}\alpha \frac{\partial^2 f}{\partial x^2} + \Lambda P(t)\,\delta(x-1) = 0 \\ &\frac{dP(t)}{dt} = - \Lambda P(t) + \lim_{x\to 0} + -bf + \frac{1}{2}\alpha\frac{\partial f}{\partial x} \\ &f(0) = 0 \end{aligned}\right\} \qquad (1)$$

where the term containing the Dirac delta funtion $\delta(x-1)$ in the first equation represents an arrival to the queue after an idle period of average duration Λ^{-1}. These equations are fully interpreted in [5]. Let us mention that they assume that the idle time is exponentially distributed with mean Λ^{-1} ; this is not a restriction for our purposes for the following reason. It is shown in [10] that the steady-state solution to (1) is identical to the steady-state probability distribution obtained for a more elaborate version of (1) where the idle time has the same mean Λ^{-1} but obeys a general (differentiable) distribution function. The equilibrium (steady state) solution to (1) is

$$f(x) = \begin{cases} R(e^{-\gamma}-1)e^{\gamma x}, & x \geq 1 \\ R(1-e^{\gamma x}), & 0 \leq x \leq 1 \end{cases}$$

$$P(0) = 1 - \frac{\Lambda}{\Lambda + \mu - \lambda} \triangleq 1 - R$$

where $\gamma = \frac{2b}{\alpha}$; the solution exists if $\gamma < 0$ as would be expected, since it implies the $\lambda < \mu$. We set $\Lambda = \lambda$ since it is known that the steady-state probability that the queue being modelled is empty is $(1-\lambda\mu^{-1})$. Thus the solution we adopt is :

$$f(x) = \begin{cases} u(e^{-\gamma}-1)e^{\gamma x}, & x \geq 1 \\ u(1-e^{\gamma x}), & 0 \leq x \leq 1 \end{cases} \qquad (2)$$

$$P(0) = 1-u, \quad u = \lambda\mu^{-1}$$

In section 4 we shall choose the parameters α, b, λ, μ, in (2) so as to approximate the behavior of a single queue in an open queueing network.

The approach of Reiser and Kobayashi for queueing networks : In [6] Kobayashi proposed a generalization to an open or closed queueing network of arbitrary topology of the results of Gaver and Shedler [4]. Kobayashi imposes reflecting boundaries to the n-dimensional diffusion process and arrives at an equilibrium joint distribution of queue lengths which is in product form. He introduces from queueing theory the probability of an empty queue for each queue to modify the solution of the diffusion equation so as to obtain a more accurate representation of queue length distribution. Reiser and Kobayashi [7] have presented a simplified diffusion model which necessitates considerably less numerical computation for obtaining the model parameters but which leads to a comparable degree of accuracy. We briefly review here their approach.

The squared coefficient of variation of the interarrival time to queue i is taken to be

$$K^{(i)} = (\lambda_i)^{-1} \sum_{j=0}^{n} [(K_j-1)p_{ji} + 1]\lambda_j p_{ji} \qquad (3)$$

where $\lambda_j = \lambda_o e_j$, $1 \leq j \leq n$, and K_j is the squared coefficient of variation of the service time at queue j if $j \neq 0$; $K_o = \lambda_o^2 V_o$. An equilibrium queue size distribution

$$p_i(m_i) = \begin{cases} 1-u_i & \text{if } m_i = 0 \\ u_i(1-\hat{\rho}_i)\hat{\rho}_i^{m_i-1} & \text{if } m_i \geq 1 \end{cases} \qquad (4)$$

is proposed for queue i where m_i is the i-th queue's length, and

$$\hat{\rho}_i = e^{\gamma i}$$

where

$$\gamma i = \frac{2(\lambda_i-\mu_i)}{\lambda_i K^{(i)}+\mu_i K_i} \qquad (5)$$

The approximation proposed for the joint probability distribution is

$$p(m_1, \ldots, m_n) = \prod_{i=1}^{n} p_i(m_i) \qquad (6)$$

for the n queues in an open network. Obviously the result is valid only when $\lambda_i < \mu_i$, $1 \leq i \leq n$.

For a closed network, the following treatment has been suggested in [7]. Suppose $\hat{u}_i$ is the utilization of server i ; the joint probability distribution is taken to be (for M customers in the network):

$$\hat{p}(m_1, \ldots, m_n) = G' \prod_{i=1}^{n} p_i(m_i) \qquad (7)$$

where

$$\hat{p}_i(m_i) = \begin{cases} 1-\hat{u}_i, & m_i = 0 \\ \hat{u}_i(1-\hat{\rho}_i)\hat{\rho}_i^{m_i-1}, & m_i = 1, 2, \ldots, M \end{cases} \qquad (8)$$

and G' is a normalizing constant. Several methods are suggested for choosing u_i. The simplest seems to be to assume that M is sufficiently large so that there exists a "bottle-neck" queue (say k) whose utilization is 1. Then

$$\hat{u}_i = \frac{\Pi_i \mu_k}{\Pi_k \mu_i} \qquad (9)$$

where Π_i, Π_k are the equilibrium probabilities that a customer will be at station i, k, respectively. This is, of course, an application of the work-rate theorem [9].

This approach has some obvious shortcomings in certain cases. Consider for instance the sequence of two FCFS service stations shown on Figure 2. Suppose station 1 has constant service time and that arrivals to the system are Poisson of rate λ. Assume also that we wanted to compute the queue length distribution at station 2 using (4). We would then have from (2)

$$K^{(2)} = 1$$

using $u_1 = \frac{\lambda}{\mu_1}$. This would mean that we would be using a diffusion model for station 2 in which it is assumed that arrivals to station 2 are Poisson, which is only true when $u_1 = 0$. When u_1 is close to 1, which is when the diffusion approximation for station 1 works well, we have nearly constant interarrival times to station 2 ($K^{(2)} \cong 0$) which is very different from what the model predicts.

Diffusion Approximation to the Behavior of a Single Queue in an Open Network

The approach we develop in this section is based on the following assumption which in general is unjustified : the departure process from any station in the open network is a renewal process, i.e. times between successive departures are independent and identically distributed. This assumption is valid in the open network with Poisson arrivals and exponentially distributed service times. It is also valid for the output of station i when $\frac{\lambda_o e_i}{\mu_i} \geq 1$, or when all $u_j \cong 0$. Let C_i, $1 \geq i \geq n$, be the squared coefficient of variation of the interdeparture times at station i, and denote by A_i the interarrival time, by S_i the service time, by Λ_i the idle time, and by τ_i the interdeparture time. We shall define $C_o = K_o$ in order to maintain an uniform presentation.

Then for t large enough, and assuming that the output processes from each individual queue are independent, the total number of arrivals to station i in the interval $[0, t]$ will be normally distributed with mean $\lambda_i t$ and variance

$$\sum_{j=0}^{n} [(C_j-1)p_{ji} + 1] \lambda_j p_{ji} t \tag{10}$$

Here we have used the fact that the sum of independent normal random variables is normal with variance being the sum of individual variances. In the usual diffusion equations for approximating the length of queue i (equations (1)) the following parameters will be chosen :

$$\begin{cases} b_i = \lambda_i - \mu_i, \quad \Lambda_i = \lambda_i \\ \alpha_i = \mu_i K_i + \sum_{j=0}^{n} [(C_j-1)p_{ji} + 1]\lambda_j p_{ji} \end{cases} \tag{11}$$

where the subscript i refers to the parameters of the equation for the i-th queue. The value of α_i used by Reiser and Kobayashi (see (4) and (5)) is

$$\alpha_i' = \mu_i K_i + \sum_{j=0}^{n} [(K_j-1)p_{ji} + 1]\lambda_j p_{ji}$$

In order to complete the development we must obtain C_i, $1 \leq i \leq n$. We shall assume that τ_i is a service time S_i with probability u_i or an interarrival time plus a service time $A_i + S_i$ with probability $(1-u_i)$. We then have

$$E\{\tau_i\} = u_i \mu_i^{-1} + (1-u_i)(\lambda_i^{-1} + \mu_i^{-1}) = \lambda_i^{-1}$$

as would be expected, and

$$E\{\tau_i^2\} = (\lambda_i^{-1})^2(1+C_i) = E\{S_i^2\} + (1-u_i)(E\{A_i^2\} + 2E\{A_i\}E\{S_i\})$$

so that

$$C_i + 1 = u_i^2(K_i+1) + (1-u_i)(\lambda_i^2 E\{A_i^2\} + 2u_i) \tag{12}$$

Finally we use (10) in the following manner ; we assume it is the variance of the number of arrivals in $[0,t]$, for large t, of a renewal process. Therefore

$$\sum_{j=0}^{n} [(C_j-1)p_{ji}+1]\lambda_j p_{ji} = \lambda_i^3(E\{A_i^2\} - (\lambda_i^{-1})^2)$$

and for $1 \leq i \leq n$, taking

$$C_i = u_i^2(K_i+1) + (1-u_i)[2u_i + 1 + \lambda_i^{-1} \sum_{j=0}^{n} [(C_j-1)p_{ji}+1] \lambda_j p_{ji}] - 1 \tag{13}$$

The values of the C_i, $1 \leq i \leq n$, to be used in (11) are the solutions to the system of equations (13).

In the case of a queueing network with exponentially distributed service times and Poisson arrivals, setting $K_i = 1$ we see that (13) has the unique solution $C_i = 1$, $1 \leq i \leq n$. In the case of the example of Figure 2 we would have

$$C_1 = u_1^2 + (1-u_1)[2u_1 + 1 + \lambda^{-1}.\lambda] - 1 = 1 - u_1^2$$

as the squared coefficient of variation of interarrival times to station 2, rather than the value $K^{(2)} = 1$ which would have been obtained by the method of Reiser and Kobayashi.

Validation : Comparison with Theoretical Results

Since there is no exact method of solution for the type of network considered in this paper, most of the material on validation of our approximation will be based on simulation results ; these will be presented in the next section. In this section we shall concentrate on comparing the few available results on departure processes from queues in order to determine the conditions under which our choice for the squared coefficient of variation of interarrival times, given in the previous section, is of sufficient accuracy.

As mentioned above, in the case of Poisson external arrivals and exponential service times at all stations our predictions for the C_i, $1 \leq i \leq n$, from (13) are exact. It has been shown in [11, 12] that for FCFS service interdeparture times are exponentially distributed only if interarrival times and servie times are exponential.

- The output of an M/D/1 queue has been examined in considerable detail [13] so that all moments of the interdeparture time distribution are available. We may apply this information to the system of Figure 2 when arrivals to the first queue are Poisson and its service times are constant. For that case equation (13) predicts $C_1 = 1-u_1^2$ for the departures from the first queue ; the value obtained from Pack [13] is exactly the same. Thus squared coefficient of variation or interarrival times to the second queue is $1-u_1^2$. The Reiser-Kobayashi [7] model predicts $K^{(2)} = 1$ for the same quantity.

- The output process of an M/G/1 queue has been studied [14] by means of a Wiener-Hopf factorisation to obtain the Laplace-Stieltjes transform of the interdeparture times distribution. In [15] the variance of interdeparture times has been computed explicitly for the system M/G/1/N, i.e. with finite population N ; the results of interest to us are obtained by setting $N \to \infty$. In this case too we are see that our predictions are exact. The variance of interdeparture times from the first queue of the system in Figure 2 if service time is general and external arrivals are Poisson is computed from [15] as being $[(1-\mu_1^2)/\lambda^2 + K_1/\mu_1^2]$; the value of C_1/λ^2 obtained from equation (13) is exactly this value.

Application to the Analysis of a Packet-Switching Network

In this section we shall apply the new diffusion approximation method introduced in section 4 to the analysis of the packet-switching sub-network CIGALE of the computer network CYCLADES [16].

The CIGALE topology is shown on Figure 3 ; nodes A through G are minicomputers used for packet buffering and switching. The numbers on the arcs connecting the nodes refer to output queues ; thus 3 refers to the queue which contains packets being transfered from node B to node C, while 5 is the queue of packets proceeding in the opposite direction. Traffic moving in the two opposite directions along an arc is non-interfering. Each node receives external traffic at a rate of λ_i $(i = A, B, \dots, G)$ packets per second. We assume that a packet arriving from the outside of the network to any node i has equal probability of having any of the other six nodes as a final destination. A packet arriving at its final destination is instantaneously destroyed there. Packet routing is fixed and summarized in Table 1.

Source	Destination	Via Node
A	F	C
B	F	C
C	G	F
D	G	F
G	C,D	B
F	A,B	C

Table 1

The routes which are not indicated on Table 1 are chosen by shortest line transit time ; this time is 0.4 seconds for a packet on all lines but 1, 2 3, 5 which are faster (0.3 seconds for one packet). This assumes in fact that all packets are of the same length (1000 bits). The assumption concerning fixed packet length is not realistic but was chosen because more realistic information was unavailable ; the same can be said about the assumption concerning the choice of destinations. The fixed routes are fairly realistic and so are the line transit times, and the order of magnitude of packet length. All queues are assumed to be of infinite capacity ; this is not the case in practice.

The network can be represented as the queueing network of Figure 4 which is a special case of the one of Figure 1. Service times at each queue are constant of value indicated on Figure 4. Routing probabilities have been computed from the routing data given in Table 1 and in the text. We have assumed, both in the simulations and the model that external arrivals of packets to each node are Poisson for lack of more realistic information ; the arrival rates (identical to those used in [17]) are chosen from practical considerations. The model has been simulated in order to obtain the average length of each queue at equilibrium ; the simulation results are compared with the diffusion approximation method introduced in Section 4, with the method of Reiser and Kobayashi [7] for the case where external arrivals to each node are Poisson (for want of real data on traffic). The accuracy with which our method predicts average queue lengths is striking. Somewhat less accurate results are obtained by the approach of Reiser and Kobayashi, while the assumption of exponentially distributed transit times and the application of Jackson's [1] formula is the least accurate approach.

Queue	Using diffusion expected queue length (G-P method)	Average queue length by simulation
1	0,515	0,530
2	1,040	0,945
3	1,388	1,317
4	0,346	0,340
5	1,495	1,397
6	1,862	1,903
7	0.547	0,560
8	1,155	1,273
9	6,234	6,568
10	1,913	2,008
11	0,491	0,497
12	0,287	0,331
13	0,647	0,658
14	0,312	0,342
15	0,491	0,506
16	1,913	1,923

Figure 1(b)

Application to a Model of an Interactive System

The model shown on Figure 5 represents an interactive computer system examined by Anderson and Sargent [18]. It was used by Reiser and Kobayashi [7] as an example to illustrate there approach ; we shall use their parameters in order to compare our approach to theirs. Jobs arrive to the system according a Poisson process of parameter μ_0. After passing through server 1 they leave the system with probability $1-\theta_1$ or enter the queue of server 2 with probability θ_1. Server 1 and 2 can be used to represent a CPU and an I/O device, respectively.

Assuming $\mu_0 = 0.5$, $\theta_1 = \theta_2 = 0.5$ we obtain using the method introduced in Section 4 the average queue lengths tabulated on Table 2(a-d) which we compare with the simulation and diffusion approximation results in [7]. The results marked R-K correspond to the results of Reiser and Kobayashi, while G-P indicates our approach. On Table 2(a) and (b) we see that the R-K or G-P results can be more accurate, depending on the parameters and the queue considered. The error in the G-P results never exceeds 8 °/o, however, while that in the R-K model predictions can be as high as 15 °/o. The G-P approach is applied to the network in which all servers are exponential on Table 2(c) and (d) ; the error here never exceeds 5 °/o.

Finally on Table 3 we apply our approach to the case where $\theta_2 = 0$ for constant service time at Server 2 and hyper-exponential service time at Server 1 ; μ_0, μ_1 are varied while μ_2 is kept fixed at $\mu_2 = 1$. The squared coefficient of variation of service time at server 1 is kept at $K_1 = 2$, and our results are compared with the simulations carried out at IRIA by M. Badel. The accuracy of our results is such that the error in average queue length never exceeds 8 °/o, and is mostly in the range of 2 to 3 °/o.

We give a comparison with the interactive computer system that may be modeled by the queueing system of Anderson and Sargent [18] represented on Figure 5.

Let be the input exponential, with rate $\mu_0 = 0.5$ and $\theta_1 = 0.5$ and $\theta_2 = 0.5$. So

$$\lambda_1 = \mu_0 e_1 = 1/(1-\theta_1) = 1$$

$$\lambda_2 = \mu_0 e_2 = \frac{\theta_1}{(1-\theta_1)(1-\theta_2)} = 1$$

The following set of parameters was chosen :

<u>Case I</u> : Server 1 : 2 stage Erlang $K_1 = 0.5$
Server 2 : Constant Service $K_2 = 0$

<u>Case II</u> : Server 1 : Exponential $K_1 = 1$
Server 2 : Exponential $K_2 = 1$

And for the two cases we take 2 values for service time

A : $\mu_1 = 0.9 \qquad \mu_2 = 0.84$

B : $\mu_1 = 0.95 \qquad \mu_2 = 0.89$

We obtain the following tables of results for the average queue size.

Case I-A : $\mu_1 = 0.9$ $C_1 = 0.5$ $\mu_2 = 0.84$ $K_2 = 0$

	queue 1	queue 2
simulation	6,84	3,22
diffusion Reiser and Kobayashi	6,76	2,70
error of R-K	1,5 %	15 %
diffusion G-P	6,65	2,95
error of G-P	2,5 %	8 %

Table 2(a)

$\mu_1 = 0.95$ $K_1 = 0.5$ $\mu_2 = 0.89$ $K_2 = 0$

	queue 1	queue 2
simulation	13,3	4,5
diffusion R-K	14,3	4,52
error R-K	7,5 %	1 %
diffusion G-P	12,62	4,39
error G-P	5 %	2,5 %

Table 2(b)

Case II-A : $\mu_1 = 0.9$ $K_1 = 1$ $\mu_2 = 0.84$ $k_2 = 1$

	queue 1	queue 2
using Jackson's results	9	5,25
diffusion G-P	9,46	5,3
error G-P	5 %	1 %

Table 2(c)

Case II-B : $\mu_1 = 0.95$ $K = 1$ $\mu_2 = 0.89$ $K_2 = 1$

	queue 1	queue 2
Results Jackson	19	9
diffusion G-P	19,53	8,65
error G-P	2,5 %	3,5 %

Table 2(d)

We see that the error with the new method is always under 8 °/o.

We compare now always with the same system but with $\theta_2 = 0$. We take an hyperexponential service for server 1 and constant service for server 2. The simulations were done at IRIA by M. Badel we obtain the following results for the mean number of customers :

μ_0	μ_1	simulation E_1	simulation E_2	diffusion E_1	diffusion E_2
0.25	0.8	1.40	0.32	1.41	0.30
0.25	0.75	1.33	0.32	1.35	0.30
0.25	0.5	1.05	0.31	1.09	0.30
0.25	0.25	0.88	0.31	0.90	0.30
0.20	0.8	1.21	0.24	1.25	0.23
0.20	0.75	1.18	0.24	1.20	0.23
0.20	0.5	1.00	0.24	1.01	0.23
0.20	0.25	0.89	0.24	0.90	0.23
0.15	0.8	1.10	0.18	1.12	0.17
0.15	0.75	1.09	0.18	1.10	0.17
0.15	0.5	0.98	0.17	0.94	0.17
0.15	0.25	0.91	0.17	0.84	0.17

Table 3

References

1. J.R. Jackson, "Jobshop-Like Queueing Systems", Management Science 10, p. 131-142, (1963).

2. F. Baskett, K.M. Chandy, R.R. Muntz, F.G. F.G. Palacios, "Open, Closed and Mixed Networks of Queues with Different Classes of Customers", Journal ACM 22, p. 248-260, 2(1975).

3. G.F. Newell, Applications of Queueing Theory, Chapman and Hall, Ltd. London, 1971, Chapter 6.

4. D.P. Gaver and G.S. Shedler, "Multiprogramming System Performance via Diffusion Approximations", IBM Research Report RJ-938, 1971, Yorktown Heights, N.Y.

5. E. Gelenbe, "On Approximate Computer System Models", Journal ACM 22, p. 261-263, 2(1975).

6. H. Kobayashi, "Applications of the Diffusion Approximation to Queueing Networks : Part I and II", IBM Research Reports RC 3943, 1972, and RC 4054, 1972, Yorktown Heights, N.Y.

7. M. Reiser and H. Kobayashi, "Accuracy of the Diffusion Approximation for Some Queueing Systems", IBM J. Res. Devel. 18, p. 110-124, 2(1974).

8. M. Chandy, U. Herzog, L. Woo, "Parametric Analisis of Queueing Networks", IBM J. Res. Devel. 19, p. 36-42, 1(1975).

9. A. Chang, S.S. Lavenberg, "Work-rates in Closed Queueing Networks with General Independent Servers", Opns. Res. 22, p. 838-847, 4(1974).

10. E. Gelenbe and R.R. Muntz, "Exact and Approximate Models for computer System Behavior, Part I and II", to appear in Acta Informatica.

11. P.J. Burke, "The Output of a Queueing System", Opns. Res., 4, 699-704 (1956).

12. E. Reich, "Waiting Times when Queues are in Tandem", Annals Math. Stat., 28, 768-773, (1957).

13. C.D. Pack, "The Output of an M/D/1 Queue", Opns. Res., 23, 750-760, (1975).

14. K.T. Marshall, "Some Relationships between the Distributions of Waiting Time, Idle Time, and Inter-output Time in GI/G/1 Queue", SIAM J. Appl. Math., 16, 324-327, (1968).

15. R.L Disney and W.P. Cherry, "Some Topics in Queueing Network Theory", In Mathematical Methods in Queueing Theory A.B. Clarke (ed.), Springer-Verlag, Berlin, (1974).

16. L. Pouzin, "CIGALE, the Packet-Switching Machine of the Cyclades Computer Network", Proc. IFIP Congress 74, Stockholm, North-Holland, 155-159 (Aug. 1974).

17. M. Irland, "Simulation of Cigale, report on assumptions and results", CCNG Report E-32, University of Waterloo, (January 1975).

18. H.A. Anderson and R. Sargent; "The statistical Evaluation of the performance of an Experimental APL/360 system". Statistical Computer Performance Evaluation, W. Freiberger (ed), p. 73-98, Academic press, 1972.

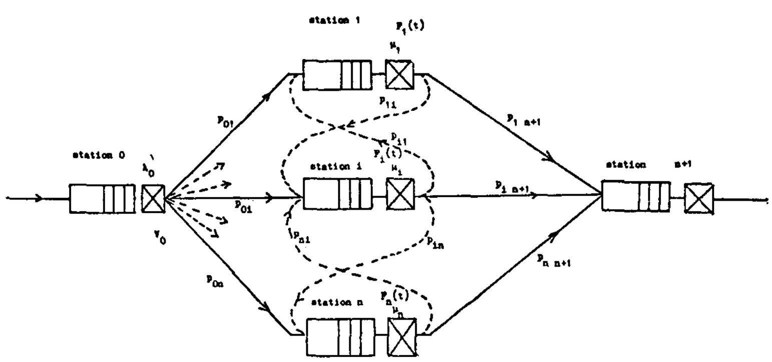

Figure 1

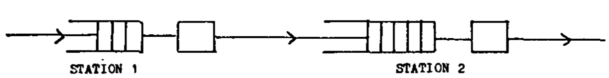

Figure 2

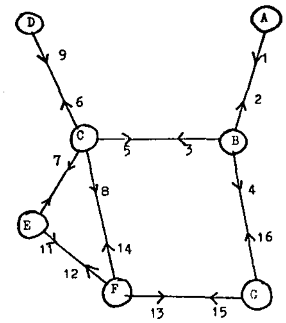

Figure 3

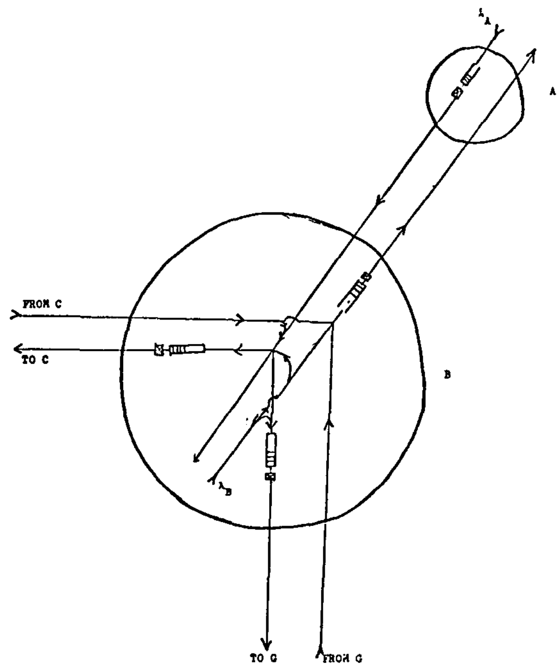

Figure 4

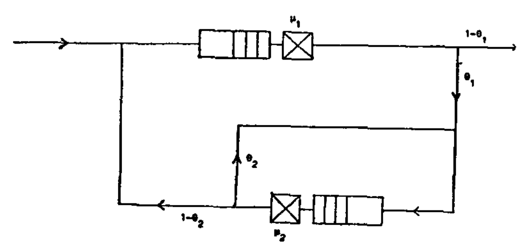

Figure 5