Revue française d'automatique, informatique, recherche opérationnelle. Recherche opérationnelle

R. G. Rani

On the waiting time distribution of a cascade queueing process

Revue française d'automatique, informatique, recherche opérationnelle. Recherche opérationnelle, tome 8, nº V1 (1974), p. 57-63

<http://www.numdam.org/item?id=RO_1974__8_1_57_0>

R.A.I.R.O.
(8e année, V-1, 1974, p. 57-63)

ON THE WAITING TIME DISTRIBUTION OF A CASCADE QUEUEING PROCESS

par R. G. RANI (1)

Summary. — *The problem of obtaining analytic solution of a queueing model made up of a number of queueing systems in series is extremely difficult. Sufficiently good approximations can be obtained by simulation.*

The concept of isomorphism (cf. Ghosal [7]) consists essentially in finding a queueing system which has some points of similarity with a given system. Such an idea may be used to study a complicated queueing system.

The concept of isomorphism is used to obtain the distribution function of the waiting time at the second queue of a queueing system of type $E_2/G_1/1 \rightarrow ./G_2/1$.

1. INTRODUCTION

One of the most general type of queuing phenomenon observed is queues in network. A simple type of network of queues is two or more queues in series. A customer enters the first queue and goes to the second queue after completing his service at the first, and subsequently goes to the third, and so on. A problem of interest is : given the inter arrival time distribution at the first queue and the service time distributions at the first, second,... etc. queues, how to obtain the waiting time distribution at any queue in the network. This problem has been attempted by a numbers of authors, Jackson [1], Reich [2], Ghosal [3], Saaty [4]; but the solution was obtained in relatively simple cases only, viz. when the system is of the type $M/M/1 \rightarrow ./M/1 \rightarrow /M/1 \ldots$ or is of the type $M/M/1 \rightarrow ./G/1$ or $M/E_2/1 \rightarrow ./M/1$.

Simulation methods help to a great extent when the system studied is more complicated than the ones mentioned above. An excellent simulation study has been done by Nelson [5] and Tocher [6]. The analytical methods though exact render the solution very complicated owing to the nature of the output distribution from the queue. In what follows, the concepts of isomorphism (see

(1) University of Paris VI.

Ghosal [7] in queues is utilized to obtain some new results for an $E_2/G_1/1 \rightarrow \cdot/G_2/1$ queueing system.

A waiting time queueing model can be represented by w_n, s_n, and t_n, where :

w_n is the waiting time of the n'th customer,

s_n is the service time of the n'th customer,

t_n is the inter arrival time between the $(n - 1)$'th and the n'th customer.

The distribution of w_n depends mainly on the distribution function of $u_n = s_n - t_n$. Hence we can represent the waiting time queueing model as $\xi_n : (w_n, u_n)$.

Considering two queueing systems ξ_1 and ξ_2 represented by,

$$\xi_1 : (w_n^1, u_n^1) \; \xi_2 : (w_n^2, u_n^2)$$

if the limiting distribution function of w_n^i in queue ξ_i is $F^i(y)$, $i = 1, 2$, the two queues ξ_1, ξ_2 are said to be isomorphic with respect to their waiting time if

$$F^{(1)}(y) = F^{(2)}(y)$$

Thus we can have 'p' isomorphic queueing systems, provided the isomorphism is defined with respect to the same property. We can also have isomorphic queues with respect to any other property of a queueing system viz. the queue size distribution.

The concept of isomorphism tends to provide a method to study the network of queues as a whole which arises very often in practice. A problem of interest is, given the relevant distribution functions of a queueing system, to obtain corresponding arrival and service time distribution functions of an isomorphic queue, the isomorphism having been defined with respect to a particular property.

In the following section, a relation between the parameters of two isomorphic queues is obtained. The system of queues studied is described below.

2. THE QUEUEING SYSTEM

Consider a cascade system of the type $E_2/G_1/1 \rightarrow ./G_2/1$. Units arrive at the first counter, the arrival following an Erlangian distribution with two phases with a mean rate of arrival 10. The counter serves the units one by one in the order of arrival, the service time following a Normal distribution with mean 8 and variance 2. The units after leaving the first counter, queue up before a second counter, after a time lag of one unit. The service time distribution at the second counter is a Normal distribution with mean 7 and variance 3. We are interested in the waiting time distribution at the second counter.

Let,

$a_n(t)\mathrm{d}t$ represent the probability of the number of arrivals at the first counter in time $(t, t + \mathrm{d}t) = n$,

$b_i(t)$ represent the density function of the service time at the i'th counter, $i = 1, 2$.

λ_1 represent the parameter of the input distribution at the first counter,

μ_i^p, σ_i represent the p'th moment about the origin and the variance respectively of the service time distribution at the i'th counter, $i = 1, 2$. We denote μ_i^1 by μ_i.

Using similar arguments to that of Finch [8], we have, for a queue, let l_r represent the interdeparture time between the r'th and the $(r + 1)$'*th* departure,

n_r represent the number of customers left behind by the r'th departure and $H_{r+1}(t, j)$ be the joint frequency distribution of n_{r+1} and l_r;

then :

$$H_{r+1}(t) = P_r\{n_r = 0\} . P_r\{t < l_r < t + \delta t \mid n_r = 0\} + P_r\{n_r > 0\}.$$
$$P_r\{t < l_r < t + \delta t \mid n_r > 0\} . \text{ Where } H_{r+1}(t) = P\{t < l_r < t + \delta t\}.$$

When $n_r > 0$, the distribution of l_r is just the distribution of service time. Hence, the second term on the right hand side reduces to

$$\{1 - P_r(0)\}\, \mathrm{d}B(t)$$

where $B(.)$ is the service time distribution and

$$P_r(0) = P_r\{n_r = 0\};$$

when $n_r = 0$, the first term simplifies to

$$\mathrm{Pr}(0) \int_0^t a_1(t - x)\mathrm{d}B(x)$$

so that we have, as $r \to \infty$:

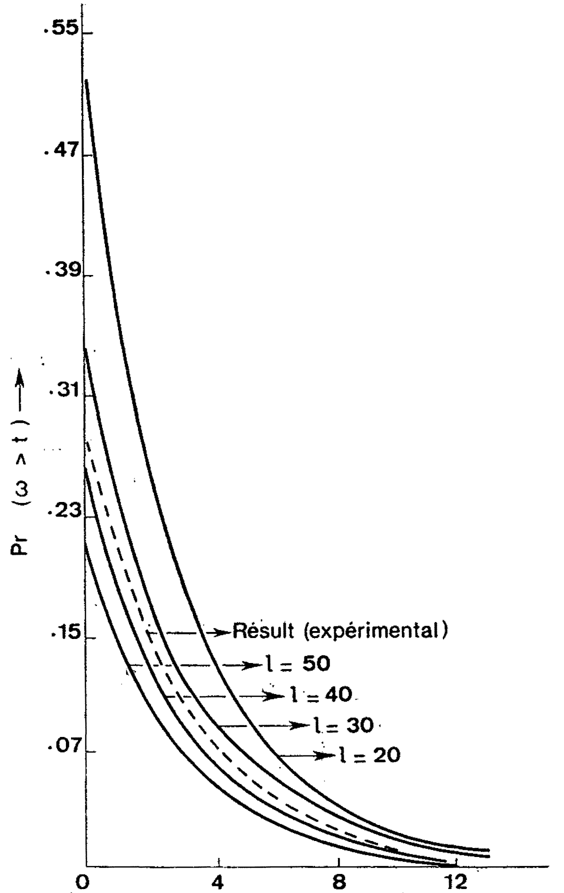
$$H(t)\delta(t) = P(0) \int_0^t a_1(t - x)\mathrm{d}B(x) + \{1 - P(0)\}\, \mathrm{d}B(t).$$

Let $g^*(s)$ represent the Laplace transform of $g(t)$ with $\mathrm{Re}(s) \geqslant 0$. The Laplace transform of the distribution of the output from the first counter is therefore,

$$H^*(S) = P_0 \left\{ \frac{1}{25} \frac{1}{(S + \frac{1}{5})^2} - 1 \right\} b^*(S) + b^*(S) \tag{1}$$

The waiting time distribution at the second queue is not affected by the constant time lag occuring in the cascade process which is being considered, since all the units are being delayed by the same time. Thus the arrival at the second counter is the same as the departure from the first, and can be given as,

$$H^*(S) = \left\{ (1 - \rho_1) \frac{1}{25} \frac{1}{(S + \frac{1}{5})^2} + \rho_1 \right\} b^*(S) \tag{2}$$



g1

where $P_0 = (1 - \rho_1)$, ρ_1 being the traffic-intensity at the first counter. Now, if we consider a sequence X_n, $n = 1, 2, \ldots$ of mutually independent random variables with distribution function $K(x) = \Pr\{X_n \leqslant x\} - \infty \leqslant x \leqslant + \infty$ and denote

$$S_0 = 0,$$
$$S_n = X_1 + X_2 + \ldots X_n$$
$$K_n(x) = \Pr\{S_n \leqslant x\}, n \geqslant 1.$$

the Laplace transform of the waiting time distribution $w(t)$ at the second counter is given by (see Prabhu [9]) :

$$\int_{0-}^{\infty} e^{\theta x} dw(x) = \exp\left[-\sum_{n=1}^{\infty}\frac{1}{n}\int_{0+}^{\infty}(1-e^{\theta x})dK_n(x)\right]$$

$$\text{for } \mathrm{Re}(\theta) < 0 \text{ where } K(x) = \int_0^{\infty} B_2(x+u)h(u)du \quad (3)$$

$B_2(.)$ is the distribution of the service time at the second counter and $h(u)$ is the density function of the interarrival time distribution at the second counter.

3. THE ISOMORPHIC QUEUE

We can imagine that the waiting time distribution at the second counter is made up of two distributions : (*a*) exponential viz. $\frac{\lambda}{\mu} e^{-(\mu-\lambda)x}$ where λ and μ are parameters to be determined and (*b*) another distribution which we call an error distribution $\epsilon(.)$. When the waiting time distribution is of the type (*a*), we can always construct a queue, input to which occurs according to a Poisson process with parameter λ,. service is exponential with parameter μ, whose waiting time distribution is $\frac{\lambda}{\mu} e^{-(\mu-\lambda)x}$. Also our aim will be to make the error distribution as small as possible. We are therefore trying to find a markovian queue isomorphic to our original queueing system in series.

λ and μ are chosen in such a manner that, if (.) follows a distribution $\epsilon(.)$,

$$E[.] = 0.$$

and $\quad \mathrm{Var}[.] \leqslant 1$

where ' 1 ' is a preassigned small value.

Now : $w^{*i}(s) = w_i^*(s)\,\epsilon^*\,(s)$,

where $w_i^*(s)$ is the Laplace transform of the waiting time distribution supposed. i.e. $\frac{\lambda}{\mu} e^{-(\mu-\lambda)x}$.

or

$$\epsilon^*(s) = w^*(s) \mid w_1(s).$$

which is the Laplace transform of the error distribution.

On simplification we have

$$E[.] = \frac{\mu}{\lambda}[1 + (\mu - \lambda)\{\lambda_1(1 - \rho_1) + \mu_1 - \mu_2\}] = 0, \qquad (4)$$

$$\text{Var}\,[.] = 150(1 - \rho_1) + 20(1 - \rho_1)(\mu_1 - \mu_2) + \mu_1'' + \mu_2'' - 2\mu_1\mu_2 - [E(.)]^2. \qquad (5)$$

Expressions (4) and (5) give a relation between the parameters of the isomorphic queue, in terms of the parameters of the original in cascade. viz. λ_1, μ_1, μ_2. For any preassigned value of ' 1 ' one can by using (4) and (5) determine an isomorphic queueing system. Thus our aim is to find values of λ and μ which satisfy condition (4) along with minimising condition (5). The optimal value of λ and μ can be calculated either (*a*) analytically or (*b*) graphically. In the discussion, optimal λ and μ are found by graph *g*1. which gives waiting time as a function of time for various values of ' 1 '. The truncated line represents the simulated result.

In a study conducted earlier through simulation, the isomorphic queue was observed to have the following property :

Pr { waiting time = 0 }	= . 73
λ / μ	= . 2 678
λ	= . 4 088
μ	= . 1 095
$\mu - \lambda$	= . 2 993

The equation (4), on substituting known values (mentioned earlier in this discussion) gives

$$\mu - \lambda = .3\,333$$

The exact value of λ and μ can now be had from (5) as a function of ' 1 '.

The simulated values of λ and μ are quite similar to the value calculated from (4).

The above method can be applied to a cascade of more than two queues but, the approximations carried out at each stage may render the result i.e. the waiting time distribution calculated, inaccurate. An application of similar procedure to more complicated network of queues is being studied.

ACKNOWLEDGEMENTS

The author would like to thank Dr. A. Ghosal, for stimulating the interest in this field. The author would also like to thank Prof. R. Fortet for his comments.

REFERENCES

[1] JACKSON R. R. P., *Queueing Process with Phase Type Service*, J. Roy. Stat. Soc., 1956, B18, 129-132.

[2] REICH P. G., *Instability of some Flow Systems of Civil Aviation*, Opns. Res. Quart., 1957, 8, 126-132.

[3] GHOSAL A., *Elements of Operational Research*, Hindustan Publishing Corpn., Delhi, India.

[4] SAATY T. L., *Stochastic Network Flows; Advances in Network of Queues*, Proc. Symp. Congestion Theory, Univ. of North Carolina, 1966, 86-107.

[5] NELSON R. T., *Queueing Network Experiments with Varying Arrivals and Service Processes*, Naval Res. Log. Quart., 1966, 13, n° 3, 321-347.

[6] TOCHER K. D. *The Art of Simulation*, The English Univ. Press, London, 1963.

[7] GHOSAL A., *Some Aspects of Queueing and Storage Systems*, Lecture Notes in Opns. Res. and Math. Systems, Springer Verlag.

[8] FINCH P. D., *The Output Process of the Queueing System M/G/1* J. Roy. Stat. Soc., B21, 375-380, 1959.

[9] PRABHU N. U., *Queues and Inventory*, John Wiley, New York, 1965.