

# Analysis of First-Come First-Served Queuing Systems With Peaked Inputs

By H. HEFFES

(Manuscript received March 8, 1973)

*This paper treats the problem of analyzing a first-come first-served queuing system, in equilibrium, when subjected to a peaked input (e.g., traffic overflowing a trunk group with Poisson input). The basic GI/M/N (renewal input to N exponential servers) queuing result is used, together with each of two models for representing peaked traffic, the Equivalent Random (E-R) model and the Interrupted Poisson Process (IPP) model. The equilibrium virtual delay distribution is derived and compared with the equilibrium distribution of delays seen by arriving calls. Numerical examples are presented, along with comparisons of results using both the above models. The results show that delays can be quite sensitive to peakedness.*

## I. INTRODUCTION

It is well known, from analysis of blocking in trunk groups, that the blocking seen by peaked traffic (e.g., traffic overflowing a first-choice trunk group with Poisson input) can be significantly larger than blocking seen by Poisson traffic with the same intensity. In this paper, we are interested in determining the effect peaked traffic has on delays in queuing systems. The analysis was motivated by a study of sender attachment delay in Crossbar Tandem switching machines receiving alternate routed (peaked) traffic.

We treat the problem of analyzing a first-come first-served queuing system with peaked input for the situation where there is no idle server if there is a waiting customer. The basic tool is the GI/M/N queuing result which requires a characterization of the input process in terms of the Laplace-Stieltjes transform of the interarrival time distribution. This characterization is provided using Wilkinson's Equivalent Random<sup>1</sup> (E-R) model where the peaked input is modeled as an overflow process from a finite trunk group with Poisson input.

The size of the finite trunk group and intensity of the Poisson input are chosen so that the overflow process produces the desired mean and peakedness (variance to mean ratio of trunks up on an infinite trunk group). The E-R method has been widely used in analyzing blocking in trunking networks and is considered our basic model when the source of the traffic peakedness is from alternate routing.

A second model, which gives a much simpler characterization of the peaked input (both computationally and analytically), is the Interrupted Poisson Process.<sup>2</sup> Here the peaked traffic is considered to be the output process of a switch, with Poisson input, where the switch is opened and closed for independent, exponentially distributed time intervals. The parameters of the switch are chosen to match either the first two or three moments of trunks up on an infinite trunk group with the corresponding moments obtained for the E-R model.

Comparisons between results using each of the above models are presented along with a set of numerical results which show that delays can be quite sensitive to peakedness.

For Poisson traffic, the virtual delay distribution (time congestion)\* is identical with the delay distribution seen by arriving calls (call congestion). This is not the case for peaked traffic. Since in some applications measurements form estimates of time congestion (e.g., SADR measurements), it is of interest to relate the time and call congestion quantities. Numerical results show the sensitivity of the relationship to peakedness.

## II. GI/M/N QUEUING RESULTS<sup>†</sup>

Consider a recurrent input<sup>‡</sup> to an  $N$  server queuing system. The service times are independent, exponentially distributed with the mean service time given by  $1/\mu_2$ . The customers are served in their order of arrival, and there is no idle server if there is a waiting customer.

Let  $F(x)$  denote the distribution function of the interarrival times. The mean interarrival time,  $1/\lambda_2$ , is given by

$$\frac{1}{\lambda_2} = \int_0^{\infty} x dF(x) \quad (1)$$

\* Time and call congestion are commonly used in trunking analysis (BCC system; i.e., blocked calls cleared). In the delay case (BCD system), we use call congestion for the delays seen by arriving calls and time congestion for the virtual delay. See appendix for precise definitions.

<sup>†</sup> The reader is referred to chapter 2 of Ref. 3 for a more detailed mathematical description of these results.

<sup>‡</sup> The peaked input will be modeled by recurrent processes. For these processes,  $F(0^+) = 0$ .

and the Laplace-Stieltjes transform of  $F(x)$  is given by

$$\Phi(s) = \int_0^{\infty} e^{-sx} dF(x). \quad (2)$$

We define the load/server (sometimes called occupancy) as

$$\rho = \frac{\lambda_2}{N\mu_2}. \quad (3)$$

The following result is available to us:<sup>3</sup> If  $\rho < 1$ , then the equilibrium delay distribution (as seen by arriving calls, i.e., call congestion) exists and is given by

$$Pr[\text{delay} > T] = \frac{A}{1 - \omega} \exp[-N\mu_2(1 - \omega)T]. \quad (4)$$

The exponential delay distribution is seen to be a function of two parameters,  $\omega$  and  $A$ . The parameter  $\omega$  is the solution, in  $(0, 1)$ , of the equation

$$\omega = \Phi[N\mu_2(1 - \omega)]. \quad (5)^*$$

For  $\rho < 1$ , this equation is known to have a unique solution in  $(0, 1)$ ; furthermore, the solution can be found by successively iterating on (5); i.e.,

$$\omega_{i+1} = \Phi[N\mu_2(1 - \omega_i)] \quad (6)$$

with  $\omega_0 \in [0, 1)$ . If we define

$$\Phi_j = \Phi(j\mu_2) \quad (7a)$$

and

$$C_j = \prod_{\nu=1}^j \frac{\Phi_{\nu}}{1 - \Phi_{\nu}}, \quad (7b)$$

then the parameter  $A$  is given by

$$A = \frac{1}{\left[ \frac{1}{1 - \omega} + \sum_{j=1}^N \frac{\binom{N}{j} [N(1 - \Phi_j) - j]}{C_j(1 - \Phi_j)[N(1 - \omega) - j]} \right]}. \quad (8)$$

Thus, given the characterization of the input in terms of  $\Phi(s)$ , eqs. (6), (7), and (8) provide the means to compute the equilibrium delay distribution as seen by arriving calls (4). The mean of the equilibrium

\* In Poisson traffic, the solution of (5) is  $\omega = \rho$ . In some cases, to be treated later, (5) can be solved in closed form.

delay distribution is given by

$$E[\text{delay}] = \frac{A}{N\mu_2(1-\omega)^2}. \quad (9)$$

In the next two sections we discuss the Equivalent Random and Interrupted Poisson models for generating peaked traffic and characterizing  $\Phi(s)$ .

### III. EQUIVALENT RANDOM MODEL

The E-R model treats peaked traffic as an overflow process from a finite trunk group with Poisson input. The holding time for the trunk group is assumed to be exponential with mean  $1/\mu_1$ .<sup>\*</sup> The number of trunks and the intensity of the Poisson traffic are chosen so the mean and variance to mean ratio (peakedness) of trunks up on an infinite overflow group closely match the desired mean and peakedness. If we denote the desired mean and peakedness by  $m_d$  and  $z_d$  respectively, then Rapp's formulas<sup>4</sup> for the Poisson load ( $a_{eq}$ ) and number of trunks ( $c$ ),

$$a_{eq} = \frac{\lambda_{eq}}{\mu_1} = m_d z_d + 3z_d(z_d - 1) \quad (10)$$

$$c = a_{eq} \left( \frac{m_d + z_d}{m_d + z_d - 1} \right) - m_d - 1, \quad (11)$$

yield an overflow process which approximates the desired process. It should be noted that use of (10) and (11), or truncation of  $c$  from (11), will often lead to overflow traffic with mean and peakedness different from (but usually close to) the desired values. To quantify this effect, the actual mean and peakedness should be computed using

$$m = a_{eq} B(c, a_{eq}) \quad (12)$$

$$z = \left[ 1 - m + \frac{a_{eq}}{c + 1 + m - a_{eq}} \right], \quad (13)$$

where  $B$  is the Erlang B function.

Takács<sup>3</sup>, Chapter 4, shows that the Laplace-Stieltjes transform of the interarrival time distribution of the overflow traffic is given by

$$\Phi(s) = \sum_{j=0}^c \binom{c}{j} \frac{1}{\lambda_{eq}^j} \prod_{i=0}^{j-1} (s + i\mu_1) / \sum_{j=0}^{c+1} \binom{c+j}{j} \frac{1}{\lambda_{eq}^j} \prod_{i=0}^{j-1} (s + i\mu_1), \quad (14)^\dagger$$

<sup>\*</sup> Note that we can have  $\mu_1 \neq \mu_2$ . For example, if overflow traffic from trunks is offered to a group of senders,  $1/\mu_1$  is of the order of minutes, whereas  $1/\mu_2$  is of the order of seconds.

<sup>†</sup> The value  $c$  is considered an integer. In practice, we round  $c$  given by (11) up and down and choose the one that gives an actual  $z$  closest to  $z_d$ . The actual  $m$  and  $z$  are computed from (12) and (13) using the rounded values of  $c$ .

where the empty product is unity. Note that  $\Phi(s)$ , given by (14), has to be repeatedly evaluated in (6) for the solution of (5) and in (8). In its present form, (14) is unsuitable for computation (for large  $c$ ) because of numerical problems. In order to avoid these problems, we use a recursive method for the evaluation of  $\Phi(s)$  developed by A. Descloux.<sup>5</sup> If we use the notation  $\Phi^c(s)$  to denote the dependence of  $\Phi(s)$  on  $c$ , then  $\Phi^c(s)$  satisfies

$$[\Phi^k(s)]^{-1} = \frac{s}{\lambda_{eq}} + 1 + \frac{\mu_1}{\lambda_{eq}} k - \frac{\mu_1}{\lambda_{eq}} k \Phi^{k-1}(s) \quad (15)$$

with initial condition

$$\Phi^0(s) = \frac{\lambda_{eq}}{s + \lambda_{eq}}. \quad (16)$$

The solution is thus obtained as follows: Using (15) to evaluate  $\Phi(s)$ , we iterate on (6) to find the  $\omega$  parameter and subsequently the  $A$  parameter (8). Having evaluated the  $\omega$  and  $A$  parameters we can compute  $Pr[\text{delay} > T]$  from (4).

We now turn our attention to the interrupted Poisson process model for generating peaked traffic, the resulting simplifications, and comparisons with the E-R model.

#### IV. INTERRUPTED POISSON PROCESS MODEL

This model, suggested by W. S. Hayward and analyzed by A. Kuczura,<sup>2</sup> treats peaked traffic as a Poisson process modulated by a random switch where the switch is opened and closed for independent, exponentially distributed time intervals. The importance of this process is that it can provide a simple and accurate approximation to overflow traffic.

This model contains three parameters, the intensity of the Poisson process into the switch,  $\lambda_s$ , calls per second; the mean open time of the switch,  $1/\bar{\omega}_s$ , seconds; and the mean closed time of the switch,  $1/\bar{\gamma}_s$ , seconds. If we choose

$$\frac{\lambda_s}{\mu_1} = A_s = mz + 3z(z - 1), \quad (17)$$

$$\frac{\bar{\omega}_s}{\mu_1} = \omega_s = \frac{m}{A_s} [m + 3z - 1], \quad (18)$$

and

$$\frac{\bar{\gamma}_s}{\mu_1} = \gamma_s = \left[ \frac{A_s}{m} - 1 \right] \omega_s, \quad (19)$$

then the mean and variance to mean ratio of trunks up on an infinite

trunk group, with mean holding time  $1/\mu_1$ , will be  $m$  and  $z$  respectively.<sup>2</sup> This corresponds to the two-parameter match of Ref. 2. The so-called three-parameter match is obtained by matching the first three moments of trunks up on an infinite trunk group with the corresponding moments that would be obtained using the E-R model. In this case,  $c$  and  $a_{\text{eq}}$  are computed from (10) and (11) and the switch parameters are obtained from

$$\frac{\lambda_s}{\mu_1} = A_s = a_{\text{eq}} \left[ \frac{\delta_2(\delta_1 - \delta_0) - \delta_0(\delta_2 - \delta_1)}{(\delta_1 - \delta_0) - (\delta_2 - \delta_1)} \right], \quad (20)$$

$$\frac{\bar{\omega}_s}{\mu_1} = \omega_s = \frac{\delta_0}{A_s} \left[ \frac{A_s - a_{\text{eq}}\delta_1}{\delta_1 - \delta_0} \right], \quad (21)$$

and

$$\frac{\bar{\gamma}_s}{\mu_1} = \gamma_s = \frac{\omega_s}{a_{\text{eq}}} \left[ \frac{A_s - a_{\text{eq}}\delta_0}{\delta_0} \right], \quad (22)$$

where the  $\delta_k$ , defined in Ref. 2, are given by

$$\delta_0 = B(c, a_{\text{eq}}) \quad (23)^*$$

and

$$\delta_k^{-1} = \frac{c + k - a_{\text{eq}}}{k} + \frac{a_{\text{eq}}}{k} \delta_{k-1}. \quad (24)^\dagger$$

For a given mean  $m$  and peakedness  $z$ , it has been shown<sup>2</sup> that the Laplace-Stieltjes transform of the interarrival time distribution of the output process of the switch is given by

$$\Phi(s) = \frac{k_1 r_1 \mu_1}{s + r_1 \mu_1} + \frac{k_2 r_2 \mu_1}{s + r_2 \mu_1}, \quad (25)$$

where the parameters  $r_1$ ,  $r_2$ ,  $k_1$ , and  $k_2$  are given by

$$r_1 = \frac{1}{2} [A_s + \omega_s + \gamma_s + \sqrt{(A_s + \omega_s + \gamma_s)^2 - 4A_s\omega_s}] \quad (26)$$

$$r_2 = \frac{1}{2} [A_s + \omega_s + \gamma_s - \sqrt{(A_s + \omega_s + \gamma_s)^2 - 4A_s\omega_s}] \quad (27)$$

$$k_1 = \frac{A_s - r_2}{r_1 - r_2} \quad (28)$$

$$k_2 = 1 - k_1 \quad (29)$$

and  $A_s$ ,  $\omega_s$ , and  $\gamma_s$  are given by (17), (18), and (19) for the two-parameter match and by (20), (21), and (22) for the three-parameter match.

\*  $B$  is the Erlang B function.

† This equation can be simply obtained from equation (1.15) in the appendix to Ref. 1.

With (25) defining  $\Phi(s)$ , eq. (5) becomes a cubic in the  $\omega$  parameter of the delay distribution. Dividing the cubic by the known root at unity gives the desired root in  $(0, 1)$

$$\omega = \left( \frac{1 + \alpha_1 + \alpha_2}{2} \right) - \sqrt{\left( \frac{1 + \alpha_1 + \alpha_2}{2} \right)^2 - \frac{\lambda_s}{N\mu_2} \left( 1 + \frac{\bar{\omega}_s}{N\mu_2} \right)} \quad (30)$$

with

$$\alpha_1 = \frac{r_1\mu_1}{N\mu_2} \quad (31)$$

and

$$\alpha_2 = \frac{r_2\mu_1}{N\mu_2}. \quad (32)$$

Thus  $\omega$ , given by (30), together with (8) specify the equilibrium delay distribution. Note that the iteration procedure (6) has been eliminated and that  $\Phi_j$ , defined by (7a), is simple to compute using (25).

#### V. TIME CONGESTION

For Poisson traffic, the virtual delay distribution (time congestion) is identical with the delay distribution as seen by arriving calls (call congestion). This is not the case for peaked traffic. Since in some applications measurements form estimates of time congestion (e.g., SADR measurements), it is of interest to relate the time and call congestion quantities.

We define time congestion as the delay in being serviced experienced by a fictitious call arriving at an arbitrary time  $t$  when the system is in equilibrium.\* It is shown in the appendix that, in equilibrium, the relationship between call congestion (CC) and time congestion (TC) is given by

$$Pr[TC > T] = \frac{\rho}{\omega} Pr[CC > T] \quad (33)^\dagger$$

if  $\rho < 1$  and the interarrival time distribution,  $F(x)$ , is not a lattice distribution.†  $Pr[CC > T]$  is the delay distribution seen by arriving calls and is given by (4).

#### VI. NUMERICAL RESULTS AND DISCUSSION

One measure of system performance of possible interest is the mean delay experienced by arriving calls, given by (9), versus CCS offered

\* See appendix for a more precise definition of time congestion.

† For Poisson traffic,  $\omega = \rho$  giving  $Pr[TC > T] = Pr[CC > T]$ .

‡ The E-R model and Interrupted Poisson model clearly satisfy the nonlattice hypothesis.

to the servers where

$$\# \text{ CCS} = 36 \frac{\mu_1}{\mu_2} m \quad (34)$$

with  $m$  given by (12) in the E-R model and

$$m = A_s \frac{\omega_s}{\omega_s + \gamma_s} \quad (35)$$

for the interrupted Poisson model. Here  $A_s$ ,  $\gamma_s$ , and  $\omega_s$  are given by (17), (18), and (19) for the two-parameter match and by (20), (21), and (22) for the three-parameter match. Figures 1(a) and 1(b) are plots of the mean delay characteristics for each of the three models of interest. The values of peakedness  $z$  ranging from 1 to 3 are presented on Fig. 1(a), and the  $z = 4$  results are plotted on Fig. 1(b). The parameters of this example are  $N = 18$ ,  $1/\mu_1 = 180$  seconds, and  $1/\mu_2 = 7.6$  seconds.

It is seen that, while the two-parameter results tend to overestimate both the E-R and three-parameter results, the differences are indistinguishable (for the entire CCS range shown) up to  $z = 1.5$  and small

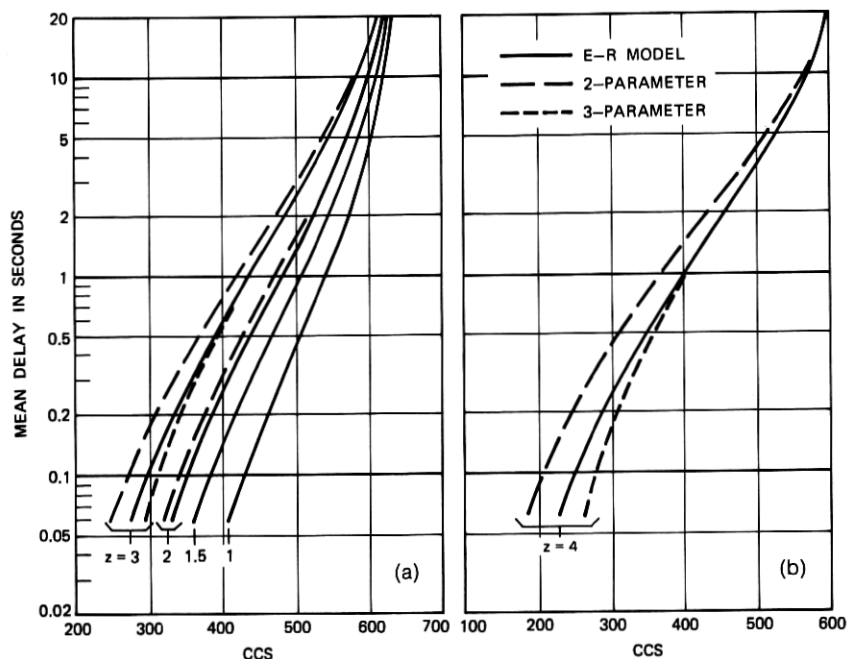


Fig. 1—Mean delay characteristics.



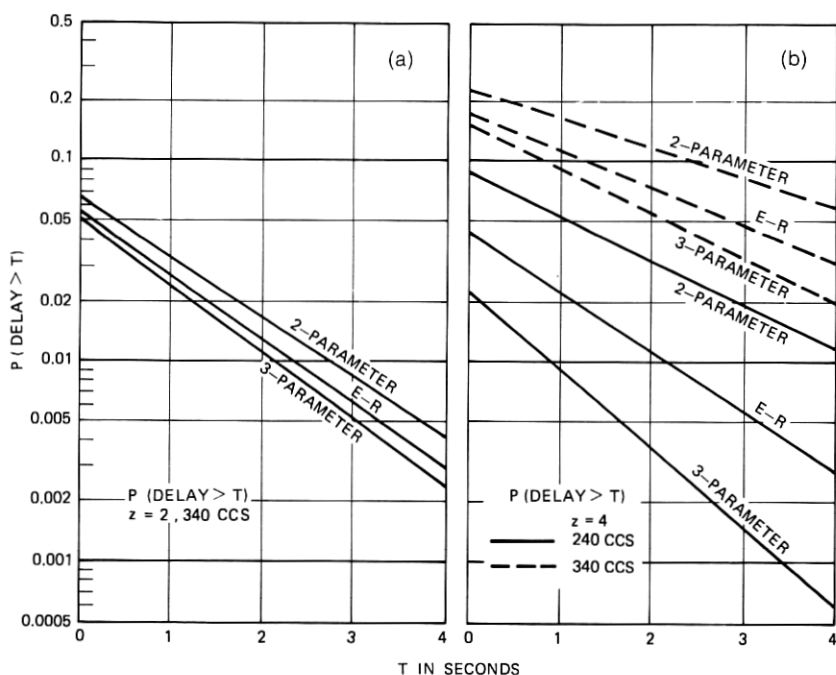


Fig. 2—Delay distribution.

up to about  $z = 2$ . The three-parameter results\* are seen to be indistinguishable from the E-R results up to  $z = 2.0$  and close up to about  $z = 3$ . In all cases, the E-R results tend to lie between the two- and three-parameter results. We see in Fig. 1(b) that the results differ greatly for  $z = 4$ . The results of Ref. 6 show that fixing the equivalent-random mean and variance for a renewal process does not necessarily tightly tie down the blocking in the BCC case. We are observing the same phenomenon here. In fact, nonnegligible discrepancies between the E-R and three-parameter results for larger  $z$ 's can occur despite a matching of three moments.

Specific delay distributions are shown in Figs. 2(a) and 2(b). An observed property of the results are that the  $\omega$  parameter (5) for the two-parameter case exceeded the  $\omega$  parameter of the E-R model which, in turn, exceeded the  $\omega$  parameter for the three-parameter case. This explains the slope differences. This also tends to order the  $T = 0$  results as shown. In many cases, the  $A$  parameter of the distribution

\* In order to avoid severe numerical problems in computing the three-switch parameters, double precision was used in eqs. (20) through (24).

(8) for the two-moment match exceeded the  $A$  parameter for the E-R model which, in turn, exceeded the  $A$  parameter for the three-parameter match. From these figures, we again observe the closeness of results for  $z = 2$  and the large differences for  $z = 4$ .

Figures 3(a) and 3(b) plot the load service relationship  $P(\text{delay} > 2.5 \text{ seconds})$  versus offered load for each of the three models. The comparisons again exhibit the same characteristics that were seen for the mean delay results [Figs. 1(a) and 1(b)].

These results show the extreme sensitivity of the queuing system performance to the peakedness of the input process. They also give some insight into the region where two- and three-parameter results are expected to be closest, i.e., low peakedness and high congestion. When seeing the discrepancies between the two- and three-parameter matches and the E-R model, we may question which is the bench mark. If the peakedness arises from alternate routing, the E-R model seems basic since it is an overflow model which has been shown to accurately describe superposition of overflows.<sup>1</sup> This has led to its wide use in analyzing blocking in trunking networks. It should also be noted that the parameters of the interrupted Poisson process have been chosen to match the moments of the E-R model.

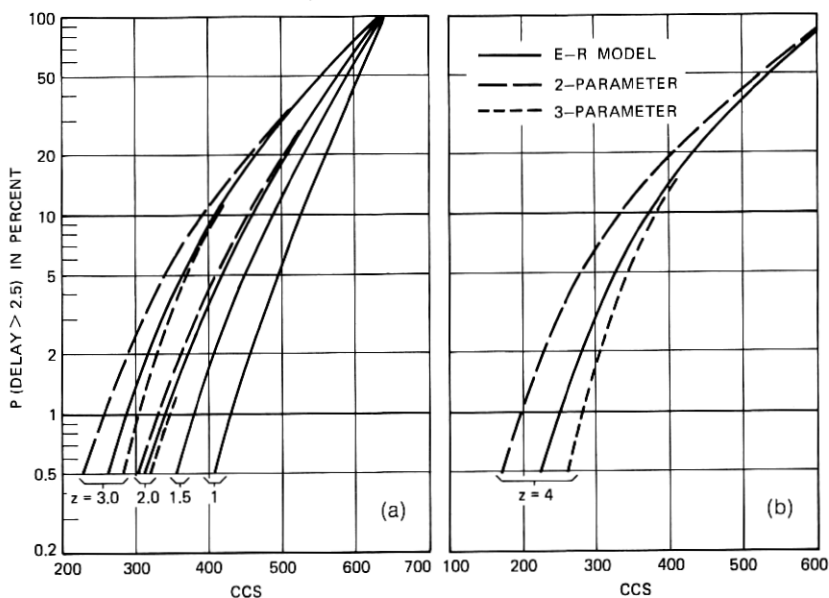


Fig. 3—Load service characteristics.

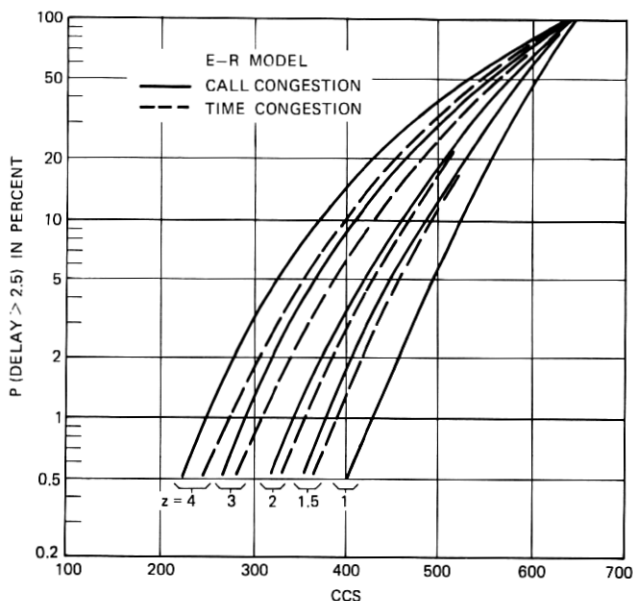


Fig. 4—Call and time congestion.

Although there are discrepancies between the two- and three-moment matches and the E-R model for high  $z$ 's, the IPP is a close approximation to the E-R model for a wide range of practical  $z$ 's. Furthermore, it should be emphasized that it provides a convenient method of analysis using birth and death equations (and simplified Laplace-Stieltjes transform) in many cases where the E-R model is intractable.

Figures 4 and 5 result from applying (33) to the example under consideration using the E-R model. Figure 4 shows the load service relationship for both call and time congestion. The call congestion results, taken from Figs. 3(a) and 3(b), are reproduced here for comparison. It is seen that the time congestion (TC) results consistently fall below the call congestion (CC) results.\* Figure 5 shows the CC-to-TC ratio ( $\omega/\rho$ ) as a function of peakedness. While for a given load the time and call congestion can differ substantially, the decrease in load that makes up the difference may be relatively small.

\*It has been shown by R. P. Marzec that for the Interrupted Poisson Process,  $Pr[TC > 0] \leq Pr[CC > 0]$ . This implies  $\rho/\omega \leq 1$ , which in turn implies  $Pr[TC > T] \leq Pr[CC > T]$ .

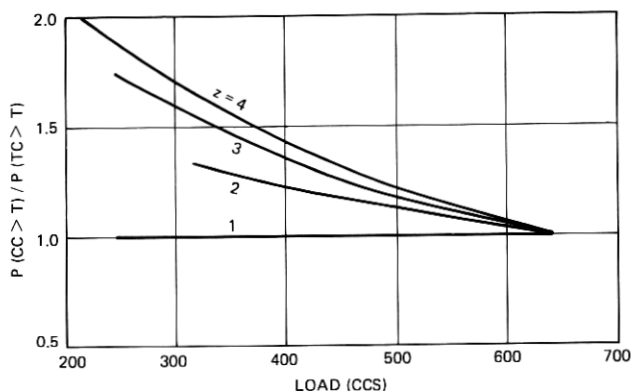


Fig. 5—Call-to-time-congestion ratio.

## VII. ACKNOWLEDGMENTS

The author expresses many thanks to J. M. Holtzman for his helpful suggestions and comments. He also thanks S. E. Miller for her excellent job in programming.

## APPENDIX

*Time Congestion*

Let  $\xi(t)$  denote the state of the system (number of customers waiting or being served at time  $t$ ) and let  $P_k(t) = Pr[\xi(t) = k]$ . Moreover, define  $\xi_n = \xi(t_n^-)$  (i.e., the number of customers in the system just prior to the arrival of the  $n$ th customer). Takács shows (Ref. 3, Theorem 1, p. 148) if  $\rho < 1$  then the limiting distribution

$$\lim_{n \rightarrow \infty} P[\xi_n = k] = P_k \quad (k = 0, 1, \dots) \quad (36)$$

exists and is independent of the initial distribution. Furthermore, he shows that

$$P_k = A\omega^{k-N} \quad (k = N, N+1, \dots), \quad (37)$$

where  $A$  is given by (8) and  $\omega$  by (5). It is also true that the above holds for  $k = N-1$ , i.e.,

$$P_k = A\omega^{k-N} \quad (k = N-1, N, N+1, \dots). \quad (38)$$

If we denote by  $\eta_n$  the waiting time of the  $n$ th customer, then the equilibrium delay distribution,

$$W(x) = \lim_{n \rightarrow \infty} W_n(x) = \lim_{n \rightarrow \infty} P[\eta_n \leq x],$$

exists and is given by

$$W(x) = \sum_{j=0}^{N-1} P_j + \sum_{j=N}^{\infty} P_j \int_0^x e^{-N\mu_2 y} \frac{(N\mu_2 y)^{j-N}}{(j-N)!} N\mu_2 dy, \quad (39)$$

which reduces to (4). We have

$$W(x) = \sum_{j=0}^{N-1} P_j + \sum_{j=N}^{\infty} P_j I_j(x),$$

where  $I_j(x)$  represents the integral in (39). The complementary distribution,  $\bar{W}(x)$ , is given by

$$\bar{W}(x) = 1 - W(x) = \sum_{j=N}^{\infty} P_j [1 - I_j(x)]. \quad (40)$$

If  $\rho < 1$  and the interarrival distribution  $F(x)$  is nonlattice, then the equilibrium time congestion probabilities given by

$$P_k^* = \lim_{t \rightarrow \infty} P_k(t) \quad (k = 0, 1, \dots) \quad (41)$$

exist and are independent of the initial state. Furthermore,

$$P_k^* = \rho P_{k-1} \quad (k = N, N+1, \dots). \quad (42)^*$$

At this point, we are able to evaluate the equilibrium time congestion distribution under the hypothesis  $\rho < 1$ , and  $F(x)$  is not a lattice distribution.

Denote  $\eta(t)$  as the waiting time of a fictitious arrival at time  $t$  and let

$$W(t, x) = Pr[\eta(t) \leq x]. \quad (43)$$

We have

$$W(t, x) = \sum_{j=0}^{N-1} P_j(t) + \sum_{j=N}^{\infty} P_j(t) \int_0^x e^{-N\mu_2 y} \frac{(N\mu_2 y)^{j-N}}{(j-N)!} N\mu_2 dy \quad (44)$$

since  $\eta(t) = 0$  if  $\xi(t) < N$ . And if  $\xi(t) = j \geq N$ , then fictitious arrival must wait for  $j+1-N$  successive departures. These departures follow a Poisson process with intensity  $N\mu_2$ . Taking the limit and using (41) we have

$$W^*(x) = \lim_{t \rightarrow \infty} W(t, x) = \sum_{j=0}^{N-1} P_j^* + \sum_{j=N}^{\infty} P_j^* I_j(x). \quad (45)$$

Using (42), the complementary distribution is given by

$$\bar{W}^*(x) = \rho \sum_{j=N}^{\infty} P_{j-1} [1 - I_j(x)]. \quad (46)$$

\* See Theorem 2, p. 153 of Ref. 3.

From (38), (40), and (46):

$$\bar{W}^*(x) = \frac{\rho}{\omega} \bar{W}(x). \quad (47)^*$$

This result is given by Ref. 7, p. 229, for a single-server queue; the multiserver case is given here for completeness.

#### REFERENCES

1. Wilkinson, R. I., "Theories for Toll Traffic Engineering in the U. S. A." (Appendix, J. Riordan), B.S.T.J., 35, No. 2 (March 1956), pp. 421-514.
2. Kuczura, A., "The Interrupted Poisson Process As An Overflow Process," B.S.T.J., 52, No. 3 (March, 1973), pp. 437-448.
3. Takács, L., *Introduction to the Theory of Queues*, New York: Oxford University Press, 1962.
4. Rapp, Y., "Planning of Junction Network in a Multi-Exchange Area, I. General Principles," Ericsson Tech., 20 (1964), No. 1, pp. 77-130.
5. Descloux, A., "On Overflow Processes of Trunk Groups With Poisson Inputs and Exponential Service Times," B.S.T.J., 42, No. 2 (March, 1963), pp. 383-397.
6. Holtzman, J. M., "The Accuracy of the Equivalent Random Method with Renewal Inputs," Seventh International Teletraffic Congress, Stockholm, 1973.
7. Cohen, J. W., *The Single Server Queue*, New York: Wiley, 1969.

---

\* Note that, for Poisson traffic,  $\omega = \rho$ , which gives  $\bar{W}^*(x) = \bar{W}(x)$ .