THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 10, December, 1972
Printed in U.S.A.

The Accuracy of Call-Congestion Measurements for Loss Systems with Renewal Input

By A. KUCZURA and S. R. NEAL

(Manuscript received June 23, 1972)

The concept of a generalized renewal process is used to derive an asymptotic approximation for the variance of the observed proportion of unsuccessful attempts on a trunk group during a given time-interval. Calls are assumed to arrive according to a general renewal process, and those which are blocked leave the system and do not return (loss system).

As an application of our result we examine the special case of an overflow input–an important example from telephone networks with alternate routing. Comparison of our results with values obtained from simulation indicates that the approximation is quite accurate for telephone traffic-engineering purposes.

I. INTRODUCTION

In a communication network, the proportion of unsuccessful attempts on a trunk group during a specified interval of time is called the measured call-congestion, and is used to estimate the single-hour blocking probability for many of the trunk groups in the Bell System network. In order to determine how many measurements should be taken to properly assess system performance, one needs to know the statistical accuracy of the estimated blocking probability.

In the context of telephone traffic-engineering, the measured call-congestion is an unbiased estimate of the blocking probability, and hence we use its variance as an indicator of the precision of the measurements. For loss systems with exponentially-distributed service times, the variance has previously been studied under the assumption that calls originate according to a Poisson process.[1] However, attempts on a trunk group are well approximated by a Poisson process only for those groups which do not serve overflow traffic from subtending groups, so that earlier results do not cover intermediate high-usage and final groups.

Assuming that the arrival process is of the renewal type, we derive an asymptotic approximation for the variance of the measured call-congestion for loss systems with exponentially-distributed service times. Since a single stream of overflow traffic is a renewal process, our results provide an estimate of the accuracy of call-congestion measurements made on intermediate high-usage and final trunk-groups.*

In Section II we describe the mathematical model used to solve our problem. The asymptotic approximation is derived in Section III. We also consider the "number of calls carried" as an unbiased estimate of the carried load and derive an asymptotic estimate of its variance. Section IV contains numerical results and Section V consists of a summary and our conclusions.

II. MATHEMATICAL MODEL

We consider a system of c servers serving customers, the arrival epochs of which constitute a renewal process. We assume that the interarrival times are independent and identically distributed according to the distribution function F, and that the service times are also independent and identically distributed according to an exponential distribution with unit mean. If all servers are occupied when a customer arrives, he leaves and has no further effect on the system. If an idle server is available when a customer arrives, service begins immediately.

Let $(0, t]$ denote a time interval of length t which commences at a point chosen at random on the time axis. Let $N(t)$ be the number of arrivals and $O(t)$ the corresponding number of blocked attempts in $(0, t]$. The ratio $O(t)/N(t)$ is the measured call-congestion. In Section III we show that the variance of $O(t)/N(t)$ can be approximated in terms of the covariance between $O(t)$ and $N(t)$ and of the individual first two moments of $O(t)$ and $N(t)$. We now describe the mathematical model used to obtain the required moments.

2.1 *A Multi-Dimensional Renewal Process*

Let t_n, $n = 0, 1, 2, \cdots$, be the instant of time at which the nth overflow occurs, $t_0 < 0 < t_1 < \cdots$, and set $X_n = t_n - t_{n-1}$. The interoverflow times X_n, $n = 1, 2, \cdots$, form a sequence of independent (because holding times are exponential) and identically distributed

* For engineering purposes, the total overflow traffic offered to such groups is adequately described by a single overflow process.[2]

random variables. Let K_n, $n = 1, 2, \cdots$, be the number of arrivals occurring in $(t_{n-1}, t_n]$ and define the row vector

$$\chi_n = (1, K_n), \qquad n = 1, 2, \cdots .$$

Since K_n, $n = 1, 2, \cdots$, is also a sequence of independent and identically distributed random variables, the components of the vector sequence χ_n, $n = 1, 2, \cdots$, are independent and identically distributed. Now, set

$$\eta(t) = \sum \chi_n ,$$

where the sum is taken over all n such that $0 < t_n \leqq t$. With these definitions it follows that for large t,

$$\eta(t) \approx (O(t), N(t))$$

and that (χ_n, X_n), $n = 1, 2, \cdots$ is a multi-dimensional renewal process.[3] Consequently, the results communicated by J. M. Hammersley[3] in the discussion of W. L. Smith's paper apply directly to our model. In particular, we shall use his eqs. (25) and (26) to compute the moments of $\eta(t)$.

We could also have used Smith's results on cumulative processes to obtain, in an indirect way, the covariance between $O(t)$ and $N(t)$. However, the concept of a cumulative process is not as naturally suited to our problem as is Hammersley's generalization of a renewal process.

Let $\mu_n(c) = E[X_1^n]$ be the nth moment of the interoverflow times from a group of c servers and

$$\nu_n = \int_0^\infty \xi^n \, dF(\xi).$$

For brevity we shall denote the arrival intensity ν_1^{-1} by λ. Equation (25) of Ref. 3 states that

$$E[O(t)] = \frac{t}{\mu_1(c)} \tag{1}$$

and

$$E[N(t)] = \frac{t}{\mu_1(c)} E[K_1].$$

But since $E[N(t)]$ is also equal to λt, we have (as is clear intuitively)

$$E[K_1] = \lambda \mu_1(c). \tag{2}$$

Now, using eq. (26) of Ref. 3 we have for large t

$$\text{Cov}\,[O(t),\,N(t)] \sim \frac{t}{\mu_1^2(c)}\,\{\lambda\mu_2(c) - E[K_1X_1]\}. \tag{3}$$

In this expression, as in others below, we omit terms which behave as $o(t)$ for large t. The results of Section IV indicate their contribution to be negligible.

Since the overflow epochs constitute a renewal process, and since we have a renewal input, we can also use eq. (26) of Ref. 3 to obtain

$$\text{Var}\,[O(t)] \sim \frac{t}{\mu_1^3(c)}\,[\mu_2(c) - \mu_1^2(c)] \tag{4}$$

and, with a change in definitions,

$$\text{Var}\,[N(t)] \sim \frac{t}{\nu_1^3}\,[\nu_2 - \nu_1^2] \tag{5}$$

for large values of t. Since ν_1 and ν_2 can be computed directly from F, we only need $\mu_1(c)$, $\mu_2(c)$ and $E[K_1X_1]$ in order to evaluate the asymptotic expressions (3), (4) and (5).

2.2 *The Joint Distribution of K_1 and X_1*

Let

$$g_c(k,\,t)\,dt = dP\{K_1 = k,\,X_1 \leqq t\},$$

where the differential on the right-hand side is to be taken with respect to the variable t. Using the same arguments as those presented in Ref. 4, pages 388–389, we obtain

$$g_c(k,\,t) = e^{-t}g_{c-1}(k,\,t) + \sum_{m=1}^{k-1}\int_0^t (1 - e^{-u})g_c(k - m,\,t - u)g_{c-1}(m,\,u)\,du. \tag{6}$$

If we define

$$\gamma_c(w,\,s) = \int_0^\infty \sum_{k=1}^\infty w^k g_c(k,\,t)e^{-st}\,dt, \tag{7}$$

then it follows from (6) that

$$\gamma_c(w,\,s) = \frac{\gamma_{c-1}(w,\,s+1)}{1 - \gamma_{c-1}(w,\,s) + \gamma_{c-1}(w,\,s+1)}. \tag{8}$$

This relation is identical to relation (7) of Ref. 4 derived for the case of

Poisson input, and (contrary to the author's comment) is also valid for any arrival process of the renewal type.

Relation (8) is of the same form as the recurrence relation for the Laplace–Stieltjes transform of the interoverflow distribution for loss systems with renewal input (see Ref. 5, page 37). Consequently, we can follow the outline of the analysis in Ref. 5 to obtain γ_c.

First, Riordan's results imply that $\gamma_c(w, s)$ can be written in the following form:

$$\gamma_c(w, s) = \frac{D_c(w, s)}{D_{c+1}(w, s)}, \tag{9}$$

where $D_0(w, s) = 1$, and, as can be seen by setting $c = 0$ in (7),

$$D_1(w, s) = \frac{1}{w\alpha(s)}, \tag{10}$$

where

$$\alpha(s) = \int_0^\infty e^{-s\xi}\, dF(\xi).$$

Furthermore, for $r \geqq 1$,

$$D_{r+1}(w, s) = D_r(w, s) + \left[\frac{1}{w\alpha(s)} - 1\right] D_r(w, s + 1). \tag{11}$$

Following Riordan, we define

$$\lambda_j = \lambda_j(w, s) = 1 - \frac{1}{w\alpha(s + j)}. \tag{12}$$

Now using (10) and (11) and mathematical induction (as noted by Riordan) one can show that

$$D_r(w, s) = 1 + \sum_{j=1}^{r} (-1)^j \binom{r}{j} \lambda_0 \lambda_1 \cdots \lambda_{j-1}. \tag{13}$$

Finally, since

$$E[K_1^m X_1^n] = (-1)^n \frac{\partial^{m+n}}{\partial w^m\, \partial s^n} \gamma_c(w, s) \bigg|_{\substack{w=1\\ s=0}}$$

for $m = 0, 1$ and $n \geqq 0$, we can compute the required moments directly from (8).

First, differentiation of (8) with respect to s yields the following results:

$$\mu_1(c) = \nu_1 D_c$$

and

$$\mu_2(c) = \frac{\nu_2}{\nu_1}\mu_1(c) + 2\mu_1(c)\sum_{k=1}^{c}\mu_1(k) - 2\nu_1 D_c^{(01)}, \tag{14}$$

where

$$D_c = D_c(1, 1) = 1 + \sum_{j=1}^{c}(-1)^j\binom{c}{j}\Lambda_0\Lambda_1 \cdots \Lambda_{j-1} \tag{15}$$

is the reciprocal of the generalized Erlang B blocking,

$$\Lambda_k = 1 - \frac{1}{\alpha(k+1)},$$

and

$$D_c^{(01)} = \frac{\partial}{\partial s} D_c(w, s)\bigg|_{\substack{w=1\\s=1}}.$$

Performing the last operation, we have

$$D_c^{(01)} = \sum_{j=1}^{c}(-1)^j\binom{c}{j}\Lambda_0\Lambda_1 \cdots \Lambda_{j-1}\left[\frac{\Omega_0'}{\Lambda_0} + \frac{\Omega_1'}{\Lambda_1} + \cdots + \frac{\Omega_{j-1}'}{\Lambda_{j-1}}\right] \tag{16}$$

where Ω_k' is the derivative of $\lambda_k(1, s)$ evaluated at $s = 1$, i.e.,

$$\Omega_k' = \frac{\alpha'(k+1)}{\alpha^2(k+1)}.$$

Similarly, the joint expectation of K_1 and X_1 is given by

$$E[K_1X_1] = \mu_1(c) + \frac{2}{\nu_1}\mu_1(c)\sum_{k=1}^{c}\mu_1(k) + \nu_1 D_c^{(10)} - D_c^{(01)}, \tag{17}$$

where

$$D_c^{(10)} = \frac{\partial}{\partial w} D_c(w, s)\bigg|_{\substack{w=1\\s=1}}$$

is given by

$$D_c^{(10)} = \sum_{j=1}^{c}(-1)^j\binom{c}{j}\Lambda_0\Lambda_1 \cdots \Lambda_{j-1}\left[\frac{1}{\Lambda_0} + \frac{1}{\Lambda_1} + \cdots + \frac{1}{\Lambda_{j-1}} - j\right]$$

and $D_c^{(01)}$ is given by (16).

Before concluding this section, we show that the covariance function

(3) reduces to the expression given by Descloux[1] for the case of Poisson input. Substituting (14) and (17) into (3), we have

$$\mathrm{Cov}\,[O(t), N(t)] \sim \frac{t}{\mu_1(c)}\left[\frac{\nu_2 - \nu_1^2}{\nu_1^2} - \frac{D_c^{(01)} + \nu_1 D_c^{(10)}}{\mu_1(c)}\right]. \tag{18}$$

Since $\mu_1^{-1}(c) = \lambda B(c, \lambda)$, where $B(c, \lambda) = D_c^{-1}$ is the generalized Erlang B blocking, we can write

$$\mathrm{Cov}\,[O(t), N(t)] \sim \lambda t B(c, \lambda)\left\{\frac{\nu_2 - \nu_1^2}{\nu_1^2} - B(c, \lambda)[D_c^{(10)} + \lambda D_c^{(01)}]\right\}.$$

When the input is Poisson, we have the following simplifications:

$$\frac{\nu_2 - \nu_1^2}{\nu_1^2} = 1,$$

$$\Lambda_k = -\frac{k+1}{\lambda},$$

$$\Omega_k' = -\frac{1}{\lambda},$$

$$\begin{aligned} D_c^{(10)} + \lambda D_c^{(01)} &= -\sum_{j=1}^{c} j\binom{c}{j} j!\, \lambda^{-j} \\ &= \lambda \frac{\partial}{\partial \lambda} E_{1,c}^{-1}(\lambda) \end{aligned}$$

where $E_{1,c}(\lambda)$ is the Erlang B blocking probability

$$E_{1,c}(\lambda) = \frac{\lambda^c}{c! \displaystyle\sum_{j=0}^{c} \frac{\lambda^j}{j!}}.$$

With these simplifications the covariance function becomes

$$\mathrm{Cov}\,[O(t), N(t)] = \lambda t \frac{\partial}{\partial \lambda}[\lambda E_{1,c}(\lambda)].$$

Substituting for the derivative (see Ref. 6, page 1)

$$\frac{\partial}{\partial \lambda} E_{1,c}(\lambda) = \frac{1}{\lambda}[c - \lambda + \lambda E_{1,c}(\lambda)]E_{1,c}(\lambda),$$

we obtain Descloux's result

$$\mathrm{Cov}\,[O(t), N(t)] \sim \lambda t E_{1,c}(\lambda)[1 + c - \lambda + \lambda E_{1,c}(\lambda)].$$

III. CALL CONGESTION AND CARRIED LOAD VARIANCES

In Section II we derived asymptotic approximations for the variances of $N(t)$ and $O(t)$ and the covariance between the two. These expressions can now be combined to obtain an asymptotic approximation for the variance of $O(t)/N(t)$. Moreover, without additional effort we can also approximate the variance of

$$\mathcal{L}(t) = \frac{N(t) - O(t)}{t}, \tag{19}$$

which is the number of calls carried per mean holding time, i.e., an estimate of carried load.

3.1 *Variance of Call Congestion*

From the theory of standard errors the variance of the call congestion is given approximately by[7]

$$\mathrm{Var}\left[\frac{O(t)}{N(t)}\right] \approx \left\{\frac{\mathrm{Var}\,[O(t)]}{E^2[O(t)]} + \frac{\mathrm{Var}\,[N(t)]}{E^2[N(t)]} - \frac{2\,\mathrm{Cov}\,[O(t),\,N(t)]}{E[O(t)]E[N(t)]}\right\}\frac{E^2[O(t)]}{E^2[N(t)]}. \tag{20}$$

The derivation of this expression is based on squaring the first-order terms of a Taylor series expansion of $O(t)/N(t)$ about the means of $O(t)$ and $N(t)$; the higher central moments of $O(t)$ and $N(t)$ are omitted. The accuracy of the approximation is discussed with the numerical results in Section IV. Using eqs. (1), (2), (4), (5), and (18) to substitute for the various quantities on the right-hand side of (20), we obtain

$$\mathrm{Var}\left[\frac{O(t)}{N(t)}\right] \approx \frac{1}{t}\left\{\frac{\mu_2(c) - \mu_1^2(c)}{\mu_1(c)} - \frac{\nu_2 - \nu_1^2}{\nu_1} + \frac{2\nu_1}{\mu_1(c)}\left[D_c^{(01)} + \nu_1\, D_c^{(10)}\right]\right\}\left(\frac{\nu_1}{\mu_1}\right)^2. \tag{21}$$

3.2 *Variance of Carried Load*

The expectation of (19) is

$$E[\mathcal{L}(t)] = \lambda\left[1 - \frac{1}{\lambda\mu_1(c)}\right].$$

Since $1/[\lambda\mu_1(c)]$ is the blocking probability, we have

$$E[\mathfrak{L}(t)] = \lambda[1 - B(c, \lambda)]$$

which shows that (19) is an unbiased estimate of the carried load. The variance of this estimate is given by

$$\text{Var}\,[\mathfrak{L}(t)] = \frac{1}{t^2}\{\text{Var}\,[N(t)] + \text{Var}\,[O(t)] - 2\,\text{Cov}\,[O(t), N(t)]\}. \qquad (22)$$

Substituting for the various terms on the right-hand side of (22), we obtain for large t

$$\text{Var}\,[\mathfrak{L}(t)] \sim \frac{1}{t}\left\{\frac{\mu_2(c) - \mu_1^2(c)}{\mu_1^3} + \frac{\nu_2 - \nu_1^2}{\nu_1^3}\left[1 - \frac{2\nu_1}{\mu_1(c)}\right] + \frac{2}{\mu_1^2(c)}\,[D_c^{(01)} + \nu_1\, D_c^{(10)}]\right\}. \qquad (23)$$

IV. NUMERICAL RESULTS

We are primarily interested in the accuracy of the measured call-congestion when the input is overflow traffic. Of course, one expects that the variance of the measured call-congestion will increase as the variance-to-mean ratio (peakedness) z of the offered traffic increases. We must also test the accuracy of the various analytical approximations. In particular, we must determine whether the standard measurement period of one hour (about 20 mean holding-times) is long enough for the asymptotic expressions (3) through (5) and the approximation (20) to be accurate enough for engineering purposes.

We computed the estimate (21) of call-congestion variance for trunk groups of 6, 10, 20, 30 and 40 trunks serving overflow traffic with various values of peakedness ranging from one to ten and a measurement interval of $t = 20$ mean holding-times. In each case, the offered load was varied over a range from 0.05 erlangs/trunk to 2 erlangs/trunk. The interarrival-time distribution of the originating traffic was obtained by using the Interrupted Poisson process[8] with a three-moment match.

To check the accuracy of the various approximations, estimates of the variance were obtained by simulation for each of the cases mentioned above. The results for the five different trunk groups were of the same general form as for the ten-trunk case shown in Figs. 1 and 2.

Figure 1 is a graph of the standard deviation of the measured call-congestion $\sigma_B = \{\text{Var}\,[O(t)/N(t)]\}^{\frac{1}{2}}$ vs the offered load $\alpha = \lambda = \nu_1^{-1}$ for several values of z and $c = 10$ trunks. Notice that for a fixed value

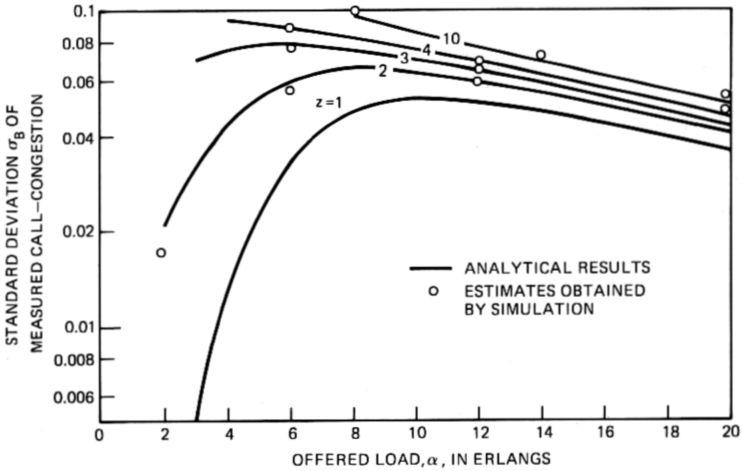

Fig. 1—Standard deviation of measured call-congestion vs offered load for $c = 10$ trunks and overflow input having peakedness z.

of α, σ_B increases as z increases (as was expected). The curves are terminated in the region of $z = \alpha$. In general, we found that our analytic results were in good agreement with the simulation for $\alpha > z$. This verifies the accuracy of our approximations when $\alpha > z$. However, a notable disparity occurred in several cases when $\alpha < z$. The latter inequality rarely arises in telephone traffic; but for other applications, where $\alpha < z$ might occur (e.g., data transmission), further work is required.

Figure 2 illustrates the behavior of the coefficient of variation σ_B/B of the measured call-congestion. For a large range of offered loads, the coefficient of variation decreases as z increases, indicating that the blocking probability increases faster as a function of z than does the standard deviation. Although the variance of the measured call-congestion decreases as α decreases for small α, the coefficient of variation (relative error) increases hyperbolically. Consequently, the relative accuracy of the measured call-congestion decreases as the blocking probability decreases, i.e., as the number $O(t)$ of observed overflows decreases.

Figure 3 displays the coefficient of variation as a function of $\alpha B = E[O(t)]/t$ for several trunk-group sizes. In each case we used both $z = 1$ and $z = 10$. The slope of the curves on the log paper is approximately $-1/2$. Hence, the coefficient of variation of the measured call-congestion is approximately inversely proportional to the square root of the number of blocked calls observed.

V. SUMMARY

For loss systems with renewal input and exponential holding times, we derived asymptotic approximations for the respective means and variances of $N(t)$, the number of arrivals, and $O(t)$, the number of overflows in the measurement period $(0, t]$. We also obtained an asymptotic estimate for the covariance between $O(t)$ and $N(t)$. Using these results, we obtained an estimate of the variance of the measured call-congestion $O(t)/N(t)$, as well as the variance of $\mathcal{L}(t) = [N(t) - O(t)]/t$ which is an unbiased estimate of carried load, provided the mean holding time is known.

Our analytical approximation for the variance of $O(t)/N(t)$ was checked by simulation for systems serving overflow traffic. In those cases which were tested, the simulation results were in excellent agreement with the analytical results for the range of system parameters (roughly $z < \alpha$) which normally arise in telephone-engineering applica-

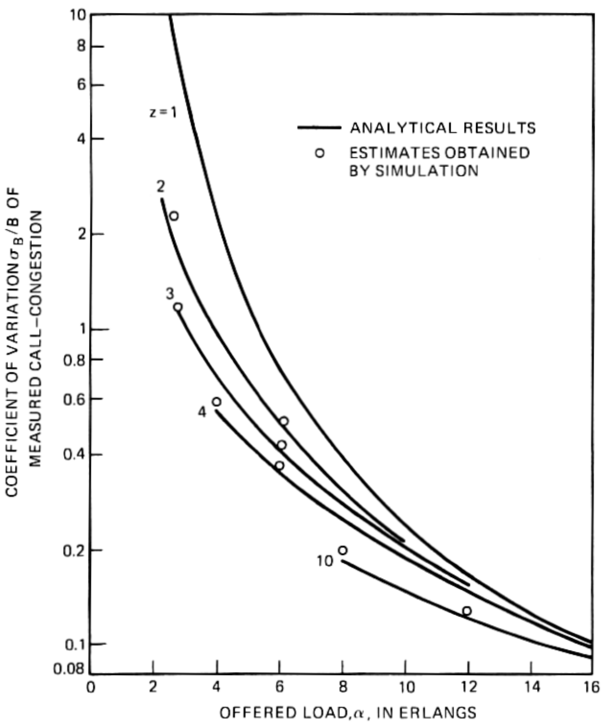

Fig. 2—Coefficient of variation of measured call-congestion vs offered load for $c = 10$ trunks and overflow input having peakedness z.

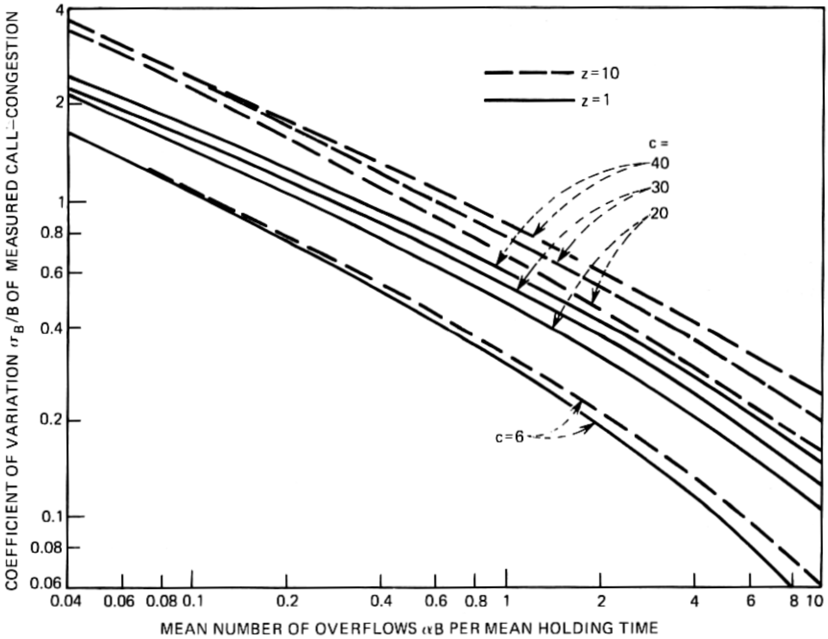

Fig. 3—Coefficient of variation of measured call-congestion vs overflow rate for $c = 6$, 20, 30, and 40 trunks, and input traffic with peakedness $z = 1$, 10.

tions. We also found empirically that the coefficient of variation of $O(t)/N(t)$ is approximately inversely proportional to the square root of the mean number of overflows in $(0, t]$.

REFERENCES

1. Descloux, A., "On the Accuracy of Loss Estimates," B.S.T.J., *44*, No. 6 (July–August 1965), pp. 1139–1164.
2. Wilkinson, R. I., "Theories of Toll Traffic Engineering in the U.S.A.," B.S.T.J., *35*, No. 2 (March 1956), pp. 421–514.
3. Hammersley, J. M., Discussion of paper by W. L. Smith, "Renewal Theory and Its Ramifications," J. Roy. Stat. Soc., Series B, *20*, No. 2, 1958, pp. 289–291.
4. 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.
5. Riordan, J., *Stochastic Service Systems*, New York: Wiley, 1962, pp. 36–40.
6. Nishimura, T., "The Derivatives of Erlang's B Formula," Elec. Commun. Lab. Tech. J., Tokyo, Extra Issue No. 22, 1966.
7. Kendall, M. G., and Stuart, A., *The Advanced Theory of Statistics*, *1*, New York: Hafner, 1958, pp. 231–232.
8. Kuczura, A., "The Interrupted Poisson Process as an Overflow Process," to appear in March 1973 issue of B.S.T.J.