Some Inequalities in the Theory of Telephone Traffic

By V. E. BENEŠ

(Manuscript received May 26, 1965)

The dynamical theory of telephone traffic in connecting networks, initiated by A. K. Erlang, has long lacked satisfactory ways of making approximations and deriving inequalities. These would reduce the fantastic computational burden implicit in the "statistical equilibrium" equations while still controlling accuracy. It is the aim of this paper to present a start in such a direction, in the form of inequalities (valid for wide classes of networks) for moments, probabilities, and ratios of expectations, among these last being the loss. The bounds in one series of these inequalities all depend on the known distribution of the number of calls in progress in a nonblocking network associated with the network under study. In a second series of cognate, simpler, but weaker inequalities, these bounds depend on Erlang's loss function or more generally on the terms of the Poisson distribution.

I. INTRODUCTION

Determining the grade of service of a telephone connecting network, as measured, for example, by the probability of blocking, continues to be a major outstanding problem of telephone traffic theory. Two principal methods are available for solving this problem. The first, simulation of a mathematical model of the operating system, has, with the advent of large high-speed computers, become very much less arduous than it used to be. The second, *calculation* of desired probabilities and expectations from [the] statistical equilibrium equations [of a mathematical model for network operation], is still hampered by the astronomical order of the equations, and in spite of its apparent promise, and its success with trunking problems early in the century, cannot be said to have reached fruition as far as connecting networks are concerned.

Indeed, it is taking so long for the strictly analytical approach to develop beyond its trunking and delay applications that in practical engineering circles its value is in serious question. The problem is not so much the lack of a suitable basic theory, for the models provided by the "statistical equilibrium" approach have been available since the time of Erlang. The real problem is a lack of approximate methods for collecting and reducing the information available in these models in a manageable way to desired quantities *without losing track of the accuracy* of the approximations along the way. It is no trick to dream up approximate ways of calculating loss. But who can meet a challenge to show *theoretically* that his approximate method is not off by more than fifty per cent?

Some basic studies of the combinatorial and probabilistic features of connecting systems have been undertaken in previous work.[1] From these emerged a broad class of Markov stochastic processes suitable as mathematical descriptions of operating connecting networks. The statistical equilibrium equations for these models have been solved in principle with complete rigor, and the probability of blocking defined and calculated in principle. The results obtained were valid for arbitrary networks, and so were of necessity rather complex. Subsequent effort has been concentrated on reducing the rigorous results to practice by finding bounds and inequalities, and by making suitable approximations.

It is the aim of this paper to present, as the first step in such a program, a number of inequalities involving such quantities of interest to the traffic engineer as the probability of blocking, the mean and variance of the number of calls in progress, and the probability of more than k calls in progress. These inequalities have several noteworthy features:

(*i*) They are simple.

(*ii*) Most of them are consequences of one analytical "basic lemma".

(*iii*) The bounds they give are couched in terms of the distribution of the number of calls in progress in a corresponding nonblocking network of comparable size, or in terms of the Poisson or truncated Poisson (Erlang) distribution. (These distributions are familiar in traffic theory, but they have not been exploited systematically to give rigorous bounds for large classes of connecting networks.)

(*iv*) They afford ways of directly converting combinatorial information about network structure into probabilistic information about the chance of loss, the load carried, the attempt rate, etc.

In casting about for approximations and inequalities in a subject such as the present one, it is reasonable to collect first those that are valid for wide classes of connecting networks, and then those that depend on special combinatorial features of certain connecting networks. Only the first task has been attempted here; a start on the second appears in a later paper.[2]

II. THE THEORETICAL MODEL AND ITS PRINCIPAL PROPERTIES

Before summarizing results, we describe the assumptions on which they are based, the notations in which they are couched, and the salient properties of the theoretical model used to represent network operation. The results themselves will be given and derived only for the important case of "one-sided" networks,[3] for which all inlets are also outlets; it will be easy to see that analogous results (with similar proofs) are valid in the "two-sided", and other, cases.

With S the set of permitted (i.e., physically meaningful) states of the network ν (of T terminals) under study, we recall[1] that S is partially ordered by inclusion $\leqq$, where $x \leqq y$ means that state x can be obtained from state y by removing zero or more calls. If x is a state, the notation $|x|$ will denote the number of calls in progress in state x.

The Markov stochastic process x_t (taking values on S) studied in previous work[1,3] is used as a mathematical description of an operating connecting network subject to random traffic. This process is based on two simple probabilistic assumptions:

(*i*) Holding-times of calls are mutually independent variates, each with the negative exponential distribution of unit mean.

(*ii*) If u is an inlet idle in state x, and $v \neq u$ is any outlet, there is a (conditional) probability

$$\lambda h + o(h), \qquad \lambda > 0$$

that u attempt a call to v in $(t,t+h)$ if $x_t = x$, as $h \to 0$. All terminals have the same traffic characteristics.

The choice of unit mean for the holding-times merely means that the mean holding-time is being used as the unit of time, so that only the traffic parameter λ needs to be specified.

It is assumed that attempted calls to busy terminals are rejected, and have no effect on the state of the system; similarly, blocked attempts to call an idle terminal are refused, with no change of state. Successful attempts to place a call are completed instantly with some choice of route.

To describe how routes are assigned to calls, we introduce a *routing matrix* $R = (r_{xy})$, with the following properties: For each x, with A_x the set of states accessible from x by new calls, let Π_x be the partition of A_x induced by the equivalence relation of "having the same calls up", or satisfying the same "assignment" of inlets to outlets; then for each $Y \in \Pi_x$, r_{xy} for $y \in Y$ is a probability distribution over Y; in all other cases $r_{xy} = 0$.

The interpretation of the routing matrix R is this: Any $Y \in \Pi_x$ represents all the ways in which a particular call c not blocked in x (between an inlet idle in x and an outlet idle in x) could be completed when the network is in state x; for $y \in Y$, r_{xy} is the chance that if this call c is attempted, it will be routed through the network so as to take the system to state y. That is, we assume that if c is attempted in x, then a state y is drawn at random from Y with probability r_{xy}, independently each time c is attempted in x; the state y so chosen indicates the route c is assigned. The distribution of probability $\{r_{xy}, y \in Y\}$ thus indicates how the calling-rate λ due to the call c is to be spread over the possible ways of putting up the call c. It is apparent that

$$\begin{aligned}\sum_{y \in A_x} r_{xy} &= \text{number of calls each of which can} \\ &\quad\ \text{actually be put up in state } x \\ &= s(x), \ (\text{"successes" in } x),\end{aligned}$$

the second equality defining $s(\cdot)$ on S. This account of the method of routing completes the description of the traffic models to be studied.

The "statistical equilibrium" equations for the stationary probabilities $\{p_x, x \in S\}$ have the simple form

$$[\,|x| + \lambda s(x)]p_x = \sum_{y \in A_x} p_y + \lambda \sum_{y \in B_x} p_y r_{yx}, \qquad x \in S$$

where

A_x = set of states accessible from x by placing a new call,
B_x = set of states accessible from x by a hangup.

The probability of blocking, or call-congestion, written in the mnemonic form Pr{bl} is just

$$\Pr\{\text{bl}\} = \frac{\sum_{x \in S} p_x \beta_x}{\sum_{x \in S} p_x \alpha_x} \tag{1}$$

where

β_x = number of idle inlet-outlet pairs that are blocked in state x.
α_x = number of idle inlet-outlet pairs in state x.

The mean of the number of calls in progress is

$$m = \sum_{x \in S} |x| p_x$$

and its variance is

$$\sigma^2 = \sum_{x \in S} (|x| - m)^2 p_x .$$

It has been shown[4] that for a "one-sided" network ν of T terminals we have

$$1 - \Pr\{\text{bl}\} = \frac{1}{\lambda} \frac{2m}{(T - 2m)^2 - T + 2m + 4\sigma^2} . \tag{2}$$

This formula relates the important parameters of the system, limiting their possible values to a surface in five dimensions.

III. TRAFFIC IN NETWORKS

It is important to equip a reader with intuitive motivation for mathematical procedures and results, and to do this at an early enough stage in the exposition of work for it to be helpful. With this motivation in mind we proceed with a discussion of certain traffic theory topics, to which the ensuing mathematics is most directly relevant.

The best-known and most widely used results in telephone traffic theory are undoubtedly those deduced in Erlang's classical model for a finite trunk group: c trunks, Poisson arrivals at rate $a > 0$, negative exponential holding-times, and blocked calls cleared without retrials. As is familiar, the probability of k calls in progress in equilibrium in this model is

$$p_k = \frac{\frac{a^k}{k!}}{\sum_{j=0}^{c} \frac{a^j}{j!}}, \quad k = 0, \cdots, c,$$

the probability of blocking is just $E(c,a) = p_c$, the load offered is a, the load carried is $m = a(1 - p_c)$, and the load variance is

$$\sigma^2 = m - ap_c(c - m).$$

It is important to note precisely just what is given, and what is calculated, in this model. The attempt rate and the number of trunks are given, and all else is calculated from a and c. This is because there is no "finite-source effect" here, no diminution of the instantaneous calling rate when many calls are in progress.

In a telephone connecting network model with a finite number of terminals, however, the finite source effect is inescapable. The attempt rate (or offered load, if the mean holding time is the unit of time) is not given *a priori*, but must itself be determined from the statistical equilibrium equations. This fact is sometimes overlooked. The same circumstance applies to the carried loads, whether the total load or simply the loads on particular parts (e.g., junctors or links, or groups thereof) within the networks; all these loads are functions of network structure and operation (e.g., routing) and are not given *a priori*. It is, nevertheless, a common practice to assume that such loads are known.[4,5]

For the reasons cited in the foregoing paragraph, the fact that the basic relationship (2) obtains between m, $\Pr\{\text{bl}\}$, σ^2, T, and λ assumes additional importance over and above its value as an aid to rough calculation.

Also, while it is inescapable, the finite source effect may nevertheless be demonstrably negligible. For example, T may be so large and λ so small that the finite source effect is virtually absent and can be neglected: almost everyone is idle almost all of the time. On the other hand, this may not happen, and thus it is important to be able to foretell to some extent when it does. Our analyses provide (among other results) some upper bounds on how large the finite source effect in a particular model actually is, and thus are of use in deciding whether or not it can be ignored.

It is known that there are respects in which either the Poisson distribution $e^{-a}(a^j/j!)$, $j = 0, 1, 2, \cdots$, or sometimes the truncated Poisson or Erlang distribution

$$\frac{\dfrac{a^j}{j!}}{\displaystyle\sum_{i=0}^{c} \frac{a^i}{i!}}\, j = 0,1, \cdots, c,$$

plays a boundary or limiting role for the equilibrium distribution of the number of calls in progress in various stochastic telephone traffic models. Examples of this phenomenon abound. In Palm's "infinite trunk" model[1] this equilibrium distribution is exactly the Poisson; in Erlang's classical model[1] for c trunks, Poisson arrivals, and lost calls cleared, it is exactly the truncated Poisson.

Further, it has been shown[6] that if the present model is used to describe the operation of a nonblocking network, then as $\lambda \to 0$ and $T \to \infty$ with $\frac{1}{2}\lambda T^2 = a =$ constant, the distribution of the number of calls in progress approaches the Poisson with mean a. Finally, it is suggestive but perhaps less directly relevant that in the present model the expansion of $\Pr\{|x_t| = k\}$ in powers of λ has the form

$$\Pr\{|x_t| = k\} = p_0(\lambda^k/k!)u_k + o(\lambda), \qquad \lambda \to 0$$

where p_0 is the probability that no calls are in progress (p_0^{-1} = normalization constant) and u_k is a constant depending only on the structure of the network and on the routing rule R used.[3]

All these facts suggest that the relationships of the distribution $\{p_k\}$ of the number of calls in progress to various possible distributions similar in algebraic character to the truncated Poisson should be explored in a systematic way, with special efforts to establish rigorous inequalities for quantities of interest in terms of truncated Poisson distributions. Such inequalities are obtained in the sequel.

Applications of the inequalities are numerous. A particularly important one provides a precise form of the following natural approximation procedure: In most telephone systems, the chance of having many more calls in progress than the average will be small; hence little error will be incurred in the calculation of loss if the states with many more calls in progress than the average are omitted from the sums defining [cf. (1)] the loss.

IV. SUMMARY AND CONCLUSIONS

Some of the problems, ideas, and observations that motivated the inequalities to be presented here were considered informally in Section III. The problem of estimating the extent of the "finite source effect", and the fact that distributions related to the Poisson or Erlang distributions give rigorous and useful bounds in traffic theory, were both mentioned.

In Section V we discuss the distribution of the number of calls in progress, and remark on the basic inequality

$$\Pr\{k \text{ calls in progress}\} \leqq \Pr\{\text{no calls in progress}\} \cdot \frac{\lambda^k}{k!} \prod_{j=0}^{k-1} \binom{T-2j}{2}. \tag{3}$$

The distribution of the number of calls in progress, it is to be recalled,[1] entirely determines the load carried and the load offered in a "one-sided" network; thus it also determines the probability of loss, by (2).

Section VI contains two analytical lemmas on which all the ensuing inequalities are based. The first merely observes that all extrema of a bilinear functional on a polyhedron must be achieved at the vertices. The second lemma is used over and over again in the sequel and for this reason it is called the "basic lemma." For certain special convex polyhedra and bilinear functionals, it pinpoints that extreme point of the polyhedron at which the functional assumes its maximum. Many problems of traffic theory lead to polyhedra and functionals of just these special types, whence their relevance.

A network ν, together with a routing rule R for ν, is called a *system*. There is a natural map μ which takes a system (ν, R) into the distribution of the number of calls in progress induced by (ν,R) (for a stochastic process x_t describing the operation of ν under R and under the traffic assumptions of Section II). Section V is devoted to proving a basic preliminary result to the effect that if ν carries at most w calls then for any R the induced distribution of the number of calls in progress belongs (if normalized so that $\Pr\{x_t = 0\} = 1$) to a special convex set of $(w + 1)$ dimensions, describable in terms of w, λ, and T, and closely related to the factors

$$b_k = \frac{\lambda^k}{k!} \prod_{j=0}^{k-1} \binom{T - 2j}{2}, \qquad k = 1, \cdots, w, \tag{4}$$

appearing in (3) and in the theory of traffic in nonblocking networks.[6]

All the preceding preliminary results are combined in Section VIII to prove a principal inequality for ratios of expectations: For nondecreasing nonnegative $f(\cdot)$ and positive nonincreasing $g(\cdot)$,

$$\max_{\nu,R} \frac{E\{f(|x_t|)\}}{E\{g(|x_t|)\}} = \frac{\sum_{j=0}^{w} f(j) b_j}{\sum_{j=0}^{w} g(j) b_j},$$

where b_k are as in (4), and the maximum is over ν and R appropriate to ν such that ν has T terminals and carries at most w calls. In Section IX we make direct applications of this result to the mean load carried and to the attempt rate.

The extent of the finite source effect is estimated in Section X in terms of the quantities b_k of (4), or, more roughly, in terms of the Erlang loss function $E(c,a)$, with $a = \frac{1}{2}\lambda T^2$. Section XI next considers the problem of estimating the equilibrium chance that more than k calls are in progress; again, this is done in terms of the b_k, and also by means of Erlang's function, using the basic lemma. Estimates of this probability have important applications to studying the error incurred in omitting states with more than k calls in progress in the sums in (1), defining loss.

It is natural to expect that in most telephone systems the probability that an immoderately large number (about twice the average number) of calls be in progress is small. This expectation suggests omitting states with more than k calls in progress from the sums defining loss, for some suitable k, as an approximation. The next two sections, XII and XIII, are concerned with the magnitude and the sign, respectively, of the error in this approximation. Two of the results are simple enough to paraphrase: Theorem 5: In virtually all cases of practical interest, if $\Pr\{|x_t| > k\} \leqq [p/(1 + p)]$, then omitting states with more than k

calls in progress in calculating Pr{bl} [by (1)] will not result in an error of more than $100p$ per cent. Theorem 6: If $\Pr\{|x_t| > k\} \leqq \epsilon$, then omitting states ... [etc., as above] will not result in an absolute error of more than ϵ.

Because the loss, Pr{bl}, is a bilinear (or linear fractional) functional of the state probabilities, determining the sign of the error incurred in the approximation under discussion is usually not simple. This sign depends (Theorems 7, 8, 9) on whether or not the fraction of *hangups* made with more than $k + 1$ calls in progress exceeds the fraction of *attempts* made with more than k calls in progress. In particular, if the fraction of *attempts* made with at most k calls in progress is not more than a certain expression (15) involving k and Erlang's loss function the approximation is an underestimate; whereas if the fraction of *hangups* made with at most $k + 1$ calls in progress exceeds another similar expression (16), then the approximation overestimates loss.

Other approximations are considered in Section XIV. A natural one is omission of states with more than k calls in progress in (2) for loss in terms of the mean and variance of the load. The basic lemma implies that this approximation is *always* an upper bound.

The final section, XV, exhibits a simple upper bound on the loss in terms of a bound on the number of blocked idle terminal-pairs in a state with k calls up, i.e., a bound of the form

$$\beta_x \leqq f_{|x|}, \qquad f(\cdot) \text{ increasing}. \tag{5}$$

This result, to be developed in a later paper,[2] provides a reasonably manageable way of converting combinatorial information about network structure directly into probabilistic inequalities about loss. The search for bounds of the form (5) for various classes and kinds of network is now one of the next most important tasks of congestion theory.

Some of the conclusions to be drawn from the present work are set down in the following list; many others will occur to those skilled in the art.

(*i*) The terms b_k, given by (4), of the distribution of the number of calls in progress in nonblocking networks can be used to give inequalities for the mean load carried, the attempt rate, the loss, and other quantities of interest arising in the study of traffic in blocking networks.

(*ii*) Terms of the Poisson or Erlang distribution, long used in trunking theory and in certain limiting cases of no congestion, can be used to give inequalities similar to, but always weaker and simpler than, those of (*i*), for the same quantities of interest.

(*iii*) The inequalities of (*i*) become those of (*ii*) in the "infinite source" limit $\lambda \to 0$, $T \to \infty$, $\frac{1}{2}\lambda T^2 = a =$ constant, with λ the calling rate per pair of idle lines, and T the total number of lines. (This limit is interesting and relevant to practical matters.)

(*iv*) Of the networks with T terminals, the nonblocking ones carry the most load and have the smallest attempt rate; among those that can carry at most w calls, the networks which are nonblocking up to w calls in progress, and block completely at w calls in progress, carry the most load and have the smallest attempt rate.

(*v*) If a network carries at most w calls, its load per line is at most

$$\lambda T[1 - E(w,a)],$$

and the equilibrium chance that it have more than k calls in progress is at most

$$1 - a^{w-k} \frac{k!}{w!} \frac{E(k,a)}{E(w,a)},$$

where $E(\cdot,a)$ is Erlang's loss function and $a = \frac{1}{2}\lambda T^2$.

(*vi*) In almost all cases, omitting states with more than three times the average number of calls in progress from the sums in (1) defining loss will result in at most a 50 per cent error in the loss.

(*vii*) Omission, in calculating loss by (1), of all states with more than k calls in progress will result in an underestimate if k is low enough. If (2) is used, this omission always overestimates loss.

(*viii*) If $\Pr\{x_t > k\} < \epsilon$, the above omission makes an absolute error of at most ϵ, if formula (1) is used.

(*ix*) Any bound $\beta_x \leqq f_{|x|}$, $f_{|}$, on the number of blocked idle terminal pairs in a state x at once yields the inequality

$$\Pr\{\text{bl}\} \leqq \frac{\sum_{j=0}^{w} f_j \frac{a_j}{j!}}{\sum_{j=0}^{w} \alpha_j \frac{a^j}{j!}}.$$

When the right-hand side is within an order of magnitude of the left, this result puts a large premium on combinatorial studies in networks of the rate at which blocking goes up with number of calls in progress.

V. THE DISTRIBUTION OF THE NUMBER OF CALLS IN PROGRESS

The calculation of the call-congestion, or probability of blocking $\Pr\{\text{bl}\}$, reduces in general to that of the stationary state-probabilities

$\{p_x, x \in S\}$. In the case of one-sided connecting networks, however, the basic formula (2) shows that a knowledge of the equilibrium mean and variance of the number of calls in progress is sufficient to determine the call-congestion. It is particularly important, then, to study the distribution of the number of calls in progress very carefully, because: (a) it contains all the information necessary to calculate congestion; (b) being a distribution of probability over a finite subset of the integers, it is a much simpler object than the distribution $\{p_x\}$ over S; (c) without doubt, it is much easier to approximate than $\{p_x\}$ itself, and so is much more likely to be useful. Various properties, inequalities, etc., pertaining to this distribution are studied in this section.

We use the notation

$$p_k = \sum_{|x|=k} p_x$$

for the probability that k calls are in progress in equilibrium. We know from Lemma 1 of Ref. 3 that for $1 \leqq k \leqq w = \max_{x \in S} |x|$,

$$kp_k = \lambda \sum_{|x|=k-1} p_x s(x). \tag{6}$$

This formula expresses the fact that in equilibrium the average rate of entrances into the set $\{x: |x| \geqq k\}$ must equal the average rate of exits from this set.

Unfortunately, (6) does not in general permit an actual calculation of $\{p_k\}$, because it depends, on the right, on the actual distribution of probability over $\{x: |x| = k - 1\}$, and not merely on p_{k-1}. However, let us observe (i) that if it takes more than $k - 1$ calls in progress to block any call at all then

$$s(x) = \binom{T - 2|x|}{2}, \qquad \text{for} \qquad |x| = \leqq k - 1,$$

and (ii) that in any case

$$s(x) \leqq \binom{T - 2|x|}{2}.$$

Thus, if n is the minimum number of calls which must be in progress in order that there by any blocked calls at all, we find that

$$kp_k = \lambda p_{k-1} \binom{T - 2k + 2}{2}, \qquad 1 \leqq k \leqq n$$

$$kp_k \leqq \lambda p_{k-1} \binom{T - 2k + 2}{2}, \qquad 1 \leqq k \leqq w.$$

Iteration of these relations then gives

$$
\begin{aligned}
p_k &= p_0 \frac{\lambda^k}{k!} \prod_{j=0}^{k-1} \binom{T-2j}{2}, \qquad 1 \leqq k \leqq n \\
p_k &\leqq p_0 \frac{\lambda^k}{k!} \prod_{j=0}^{k-1} \binom{T-2j}{2}, \qquad 1 \leqq k \leqq w.
\end{aligned}
\tag{7}
$$

The bound on the right of this last inequality has the same form as the exact formula for p_k for a nonblocking network.[6] This implies that for λ fixed the maximum possible value of the ratios

$$(p_k/p_0) \qquad k = 1, \cdots, w$$

is achieved by nonblocking networks, and it is achieved by a blocking network at a particular value of k only if

$$\frac{1}{p_{k-1}} \sum_{|x|=k-1} s(x) p_x = \binom{T-2k+2}{2}$$

i.e., only if the conditional expectation of $s(\cdot)$ given that $k-1$ calls are in progress is the number

$$\binom{T-2k+2}{2}.$$

Since this is an upper bound for $s(\cdot)$ over all x with $|x| = k-1$, this means that all the probability is concentrated on the nonblocking states, so that the bound (7) is also achieved for $k-1$. (This observation will be fundamental in the proof of Lemma 3.)

Reasoning from (6) leads to the inequalities

$$\lambda p_{k-1} \min_{|y|=k-1} s(y) \leqq k p_k \leqq \lambda p_{k-1} \max_{|y|=k-1} s(y)$$

and thence by iteration to the

Remark:

$$\frac{\lambda^k}{k!} \prod_{j=0}^{k-1} \min_{|y|=j} s(y) \leqq \frac{p_k}{p_0} \leqq \frac{\lambda^k}{k!} \prod_{j=0}^{k-1} \max_{|y|=j} s(y).$$

This result indicates (to a first approximation) how the values assumed by the "success" function $s(\cdot)$ on S affect the distribution of the number of calls in progress, and through it, the congestion or probability of blocking. Obviously, the nearer the network is to being nonblocking, i.e., the nearer $s(\cdot)$ comes to assuming the value

$$\binom{T - 2|x|}{2}$$

for state x, the closer p_k will be to its upper bound (7), and the less will be the congestion.

VI. TWO PRELIMINARY RESULTS

Lemma 1: Let P be a polyhedron in n-dimensional Euclidean space, and let

$$F(x) = \frac{c_1 + (a,x)}{c_1 + (b,x)}, \qquad (\cdot\,,\cdot) = \textit{inner product}$$

be a bilinear (or linear fractional) function of the n-vector x, such that the plane $c_2 + (b,x) = 0$ does not intersect P. Then the extreme values of $F(\cdot)$ on P are assumed at the vertices of P.

Proof: Let x be a point interior to P. Since the sign of

$$\frac{\partial F}{\partial x_i} = \frac{a_i[c_2 + (b,x)] - b_1[c_1 + (a,x)]}{[c_2 + (b,x)]^2}$$

does not depend on x_i, we can find another point $y \in \partial P$ (the boundary of P) such that $F(x) \leqq F(y)$. The point y will be on a face P_1 of P determined by a linear condition $(c,x) = \alpha$ which can be used to eliminate one of the variables from $F(\cdot)$ to get a new bilinear function $F_1(\cdot)$ of $(n - 1)$ variables agreeing with $F(\cdot)$ on P_1. Except for dimension, the problem of maximizing $F_1(\cdot)$ over P_1 is of exactly the same form as that maximizing $F(\cdot)$ over P. The result is true for $n = 1$, and hence for all $n \geqq 1$. The argument for minima is dual.

Basic Lemma (Lemma 2): Let $\lambda = (\lambda_0, \lambda_1, \cdots, \lambda_n)$ be a vector of $(n + 1)$ positive numbers, and let Λ be the closed convex hull of the points

$$\begin{array}{cccccc}
\lambda_0, & 0, & 0, & 0, & \cdots, & 0 \\
\lambda_0, & \lambda_1, & 0, & 0, & \cdots, & 0 \\
\lambda_0, & \lambda_1, & \lambda_2, & 0, & \cdots, & 0 \\
 & & \vdots & & & \\
\lambda_0, & \lambda_1, & \lambda_2, & & \cdots, & \lambda_n .
\end{array}$$

Let $f(\cdot)$ be nondecreasing and nonnegative, and let $g(\cdot)$ be nonincreas-

ing and positive, on $\{0,1, \cdots, n\}$. *Then with* $(\cdot\,,\cdot) =$ *inner product,*

$$\max_{v \in \Lambda} \frac{(f,v)}{(g,v)} = \frac{(f,\lambda)}{(g,\lambda)}.$$

Proof: It follows from Lemma 1 that the maximum is assumed at a vertex. However, as can be verified,

$$\frac{\sum_{j=0}^{k} f_j \lambda_j}{\sum_{j=0}^{k} g_j \lambda_j} \leqq \frac{\sum_{j=0}^{k+1} f_j \lambda_j}{\sum_{j=0}^{k+1} g_j \lambda_j}, \quad k = 0, 1, \cdots n - 1.$$

VII. BASIC INCLUSION

Let I be the set of inlets, and Ω that of outlets, of a possible or intended connecting network to be used for making calls from I to Ω. By a *network* ν *for* I *and* Ω we mean a quadruple

$$\nu = (G,I,\Omega,S)$$

where G is a linear graph indicating network structure, I and Ω are respectively the inlets and outlets of the network, and S is the set of permitted states.[1] The letter w is used to stand for the largest possible number of calls in progress; thus

$$w = \max_{x \in S} |x|.$$

If ν is one-sided, $I = \Omega$ and $w \leqq [\frac{1}{2} |I|]$. If ν is two-sided, $I \cap \Omega = \theta$ and $w \leqq \min\{|I|, |\Omega|\}$.

By a *system for* I *and* Ω we mean a pair (ν,R) with ν a network for I and Ω and R a routing rule defined on the states $x \in S = S(\nu)$ and satisfying the conditions of Section II. It follows from the theoretical assumptions made in Section II that, together with a value of the traffic parameter $\lambda > 0$, ν and R determine a stochastic process x_t taking values on S with a stationary distribution

$$\{p_x = p_x(\nu,R,\lambda),\ x \in S(\nu)\}$$

determined by the equilibrium condition (cf. Section II).

We shall assume that I, Ω, and λ are fixed, and shall omit indications of dependence on these notions or numbers.

Let $\mathcal{S}_n$ denote the set of systems for I and Ω such that $w \leqq n$. With

$$p_k = p_k(\nu,R) = \sum_{\substack{|x|=k \\ x \in S(\nu)}} p_x(\nu,R)$$

the map $\mu(\cdot\,,\,\cdot)$ is defined on $\mathcal{S}_n$ for each value of $\lambda > 0$ by

$$\mu\colon (\nu,R) \rightarrow \frac{(p_0\,, p_1\,, \cdots\,, p_n)}{p_0}, \qquad p = p(\nu,R),$$

i.e., its value for (ν,R) is the distribution of the number of calls in progress in the associated stochastic process x_t, normalized so that $p_0 = 1$.

Lemma 3: Let $n \leqq [\tfrac{1}{2}T]$, *and let*

$$b_0 = 1$$

$$b_k = \frac{(\tfrac{1}{2}\lambda)^k}{k!}\,\frac{T!}{(T-2k)!}, \qquad k = 1, \cdots, n. \tag{8}$$

Let C be the closed convex hull of the $(n + 1)$*-dimensional points*

$$\begin{aligned} c_0 &= (1,0,0, \cdots, 0) \\ c_1 &= (1\,, b_1\,, 0\,, \cdots, 0) \\ c_2 &= (1\,, b_1\,, b_2\,, \cdots, 0) \\ &\;\;\vdots \\ c_n &= (1\,, b_1\,, b_2\,, \cdots, b_n). \end{aligned}$$

Then

$$\mu(\mathcal{S}_n) \subseteq C,$$

i.e., C includes the μ*-image of* $\mathcal{S}_n$.

Proof: We show first that each c_i, $i = 0, \cdots, n$ is in fact in the image of $\mathcal{S}_n$ under $\mu(\cdot)$. Let ν_0 be the trivial network containing no crosspoints, and let R_0 be the trivial rule that says nothing. Then

$$\mu\colon(\nu_0,R_0) \rightarrow c_0\,,$$

and $c_0 \in \mu(\mathcal{S}_n)$. Now let ν_k, $k = 1, \cdots, n$, be a "one-sided" network consisting (*i.*) of a concentrator taking T terminals to $2k$ in a nonblocking manner, and (*ii.*) of a nonblocking "one-sided" network on those $2k$ terminals. In such a network, obviously, a state is nonblocking if fewer than k calls are in progress; the network blocks up completely as soon as k calls are in progress. It follows from the arguments for Theorem 1 of Ref. 6 that for any routing rule R_k appropriate to ν_k,

$$\mu\colon(\nu_k,R_k) \rightarrow c_k\,.$$

From formula (7) of Section V we know that

$$\frac{p_k(\nu,R)}{p_0(\nu,R)} = \frac{p_k}{p_0} \leqq b_k\,.$$

Hence, to show that $\mu(\mathcal{S}_n)$ is contained in C, it suffices to show that if for some $(\nu,R) \in \mathcal{S}_n$ and some $1 \leqq k \leqq n$,

$$(p_k/p_0) = b_k, \qquad p = p(\nu,R)$$

then for all $1 \leqq j \leqq k$,

$$(p_j/p_0) = b_j .$$

Suppose then that $p_k/p_0 = b_k$. Using (6) we find

$$\begin{aligned} kp_k = \lambda \sum_{|y|=k-1} p_y s(y) &= \lambda p_{k-1} E\{s(x_t) \mid |x_t| = k-1\} \\ &= p_0 k b_k \\ &= p_0 \lambda \binom{T-2k+2}{2} b_{k-1} . \end{aligned}$$

Hence,

$$\frac{p_{k-1}}{p_0} = b_{k-1} \frac{\binom{T-2k+2}{2}}{E\{s(x_t) \mid |x_t| = k-1\}} .$$

But,

$$\max_{|y|=k-1} s(y) \leqq \binom{T-2k+2}{2}$$

so the ratio is $\geqq 1$. But we know from (7) that $p_{k-1}/p_0 \leqq b_{k-1}$. Hence, equality holds, and by iteration,

$$p_j/p_0 = b_j, \qquad 1 \leqq j \leqq k.$$

VIII. PRINCIPAL INEQUALITY FOR RATIOS OF EXPECTATIONS

We now combine Lemmas 2 and 3 to obtain a basic inequality for ratios of expectations. Applications of this result to the quantities of interest in traffic engineering appear in the following Sections IX through XIV.

Theorem 1: If $f(\cdot)$ is nondecreasing and nonnegative, and $g(\cdot)$ is nonincreasing and positive, on $\{0,1, \cdots, n\}$, then

$$\max_{(\nu,R) \in \mathcal{S}_n} \frac{E\{f(|x_t|)\}}{E\{g(|x_t|)\}} = \frac{\sum_{j=0}^{n} f(j) b_j}{\sum_{j=0}^{n} g(j) b_j} ,$$

(the expectation being calculated with respect to the stationary probabilities associated with ν and R.)

Proof: By Lemma 3, the image of $\mathcal{S}_n$ under $\mu(\cdot\ ,\ \cdot)$ is contained in the closed convex hull C of the points $c_0, c_1, \cdots, c_n$. By the basic lemma, the maximum of the functional

$$\xi(r) = \frac{(f,r)}{(g,r)} = \frac{\sum_{j=0}^{n} f(j)r_j}{\sum_{j=0}^{n} g(j)r_j}$$

for $r \in C$ is assumed at $r = c_n$.

IX. INEQUALITIES FOR THE MEAN AND THE ATTEMPT RATE

Let ν and ν' be two connecting networks with the same number of terminals, and the same offered traffic λ per idle pair. If ν is nonblocking, it is our intuitive expectation that it will carry at least as great a load as ν', and that (since more lines of ν are busy on the average than of ν') the attempt rate for ν' will be at least as great as that for ν. It is being assumed here, of course, that in each case the operation of the network is being represented by a stochastic process x_t of the type described in Section II, with

$$m = \text{carried load} = \frac{\sum_{k=1}^{w} kp_k/p_0}{1 + \sum_{k=1}^{w} p_k/p_0}$$

$$\lambda E\{\alpha_{|x_t|}\} = \text{attempt rate} = \lambda \frac{\sum_{k=0}^{w} p_k/p_0 \binom{T-2k}{2}}{1 + \sum_{k=1}^{w} p_k/p_0}.$$

Arguments ν,R are used in the next three results to indicate dependence on the network ν and the routing rule R under discussion.

Theorem 2: Let ν be a one-sided network of T terminals, and let $w = w(\nu) = \max_{x \in S(\nu)} |x|$, $a = \frac{1}{2}\lambda T.^2$ Then for any R such that $(\nu,R) \in \mathcal{S}_w$

$$m(\nu,R) \leq \frac{\sum_{j=1}^{w} jb_j}{1 + \sum_{j=1}^{w} b_j} \leq a[1 - E(w,a)].$$

Proof: For the first inequality let $f(j) = j$ and $g \equiv 1$ in Theorem 1; for the second, use the basic lemma and $b_j < a^j/j!$.

Corollary 1: Let (ν,R) *and* (ν',R') *belong to* $\mathcal{S}_w$ *for some integer* w, *with both* ν *and* ν' *one-sided. If* ν' *is nonblocking until* w *calls are in progress, then with* $a = \frac{1}{2}\lambda T^2$

$$m(\nu,R) \leqq m(\nu',R') = \frac{\sum_{j=1}^{w} jb_j}{1 + \sum_{j=1}^{w} b_j} \leqq a[1 - E(w,a)].$$

The following properties of the Erlang loss function

$$E(c,a) = \frac{\frac{a^c}{c!}}{\sum_{j=0}^{c} \frac{a^j}{j!}}$$

are used:

$$\frac{\sum_{j=1}^{c} j \frac{a^j}{j!}}{\sum_{j=0}^{c} \frac{a^j}{j!}} = a[1 - E(c,a)],$$

$$\frac{\sum_{j=1}^{c} j^2 \frac{a^j}{j!}}{\sum_{j=0} \frac{a^j}{j!}} = a[1 - E(c,a)] - a^2 E(c,a)\left[\frac{c}{a} - 1 + E(c,a)\right] + a^2[1 - E(c,a)]^2,$$

$$= (a + a^2)[1 - E(c,a)] - acE(c,a).$$

Theorem 3: Let ν *be a one-sided network of* T *terminals. Then for* $w = w(\nu)$, $a = \frac{1}{2}\lambda T^2$, *and any* R *such that* $(\nu,R) \in \mathcal{S}_w$,

$$E\{\alpha_{|x_t|}\}_{\nu,R} \geqq \frac{\sum_{j=0}^{w} \binom{T - 2j}{2} b_j}{1 + \sum_{j=1}^{w} b_j} \geqq \frac{\sum_{j=0}^{w} \binom{T - 2j}{2} \frac{a^j}{j!}}{\sum_{j=0}^{w} \frac{a^j}{j!}}.$$

Proof: The first inequality follows from taking $f \equiv 1$ and $g(j) = \alpha_j$ in Theorem 1; the second, from the basic lemma. The last term on the right is expressible in terms of Erlang's loss function as

$$\binom{T}{2} - a(2T - 3 + 2a)[1 - E(w,a)] - 2awE(w,a).$$

X. APPLICATIONS TO ESTIMATING THE "FINITE-SOURCE EFFECT"

It is reasonable to expect that in a telephone system with a large number T of terminals, each one contributing only a small amount of traffic, the "finite-source effect" will be small. Since the finite-source effect is a diminution of the instantaneous calling rate due to busy terminals, it is properly measured by the *fraction of busy terminals*, i.e., in our model, the quantity

$$q = 2m/T = \text{load per customer's line in erlangs.}$$

When T is so large and λ so small that q is very small we might with justifiable confidence replace our finite-source model with an infinite source model. One way of doing this is to consider a sequence of connecting networks which concentrate traffic from more and more terminals into a sub-connecting network of fixed structure. However, we do not here digress into a detailed consideration of this transition; the basic idea has been at the heart of applications of the Poisson arrival process in telephone traffic theory since its beginning. Instead, we obtain an upper bound on q in terms of T and λ; this bound provides a conservative estimate of the negligibility of the finite source effect.

Corollary 2: With $a = \frac{1}{2}\lambda T^2$, $E(c,a)$ *the (first) Erlang loss function, and* $b_0, b_1, \cdots, b_w$ *as in formula (8),*

$$q \leqq \frac{2}{T}\frac{\sum_{j=1}^{w} jb_j}{\sum_{j=0}^{w} b_j} \leqq \frac{2a}{T}\{1 - E(w,a)\} \qquad (w = \max_{x\varepsilon S} |x|)$$

$$\leqq \frac{2a}{T}\left(1 - \frac{e^{-a}a^w}{w!}\right)$$

$$\leqq \frac{2a}{T} = \lambda T.$$

Proof: The first inequality follows from Theorem 1, the second from the basic lemma, and the third from

$$E(w,a) = \frac{\frac{a^w}{w!}}{\sum_{j=0}^{w}\frac{a^j}{j!}} \geqq \frac{e^{-a}a^w}{w!}.$$

Alternatively, since the finite-source effect is a diminution of the calling rate due to busy terminals, one can also estimate it in terms of the difference between the *maximal* calling rate $\lambda\binom{T}{2}$ when no terminals are busy and the *average* calling rate

$$\lambda E\{\alpha_{|x_t|}\} = \lambda \sum_{j=0}^{w} p_j \binom{T-2j}{2}.$$

This estimate is covered in the

Corollary 3: The average diminution

$$D = \lambda \binom{T}{2} - \lambda E\{\alpha_{|x_t|}\}$$

in calling rate due to busy terminals satisfies the inequality

$$D \leqq \lambda a(2T - 3 + 2a)[1 - E(w,a)] + \lambda awE(w,a)$$
$$\leqq O(T^{-1}) \quad \text{as} \quad \lambda \to 0, \qquad T \to \infty, \qquad a = \tfrac{1}{2}\lambda T^2.$$

Proof: Theorem 3 and the known properties of $E(\cdot,a)$.

XI. ESTIMATE OF THE CHANCE OF MORE THAN k CALLS IN PROGRESS

The chance $\Pr\{|x_t| > k\}$ is a quantity that is useful in estimating the extent of the finite source effect, and the error incurred in ignoring states with more than k calls in progress in calculating loss. (See Section XII.) Upper bounds for it are given in

Theorem 4: If ν is one-sided, $w = w(\nu)$, and $a = \frac{1}{2}\lambda T^2$, then

$$\Pr\{|x_t| > k\} \leqq \frac{\sum_{j=k+1}^{w} b_j}{1 + \sum_{j=1}^{w} b_j}$$

$$\leqq 1 - \frac{a^{w-k}k!}{w!}\frac{E(k,a)}{E(w,a)}.$$

Proof: For the first inequality, choose

$$f(j) = \begin{cases} 0 & j \leqq k \\ 1 & j > k \end{cases}$$

and $g(\cdot) \equiv 1$ in Theorem 1; the second follows from the basic lemma.

XII. APPROXIMATION THEOREMS FOR THE PROBABILITY OF BLOCKING

In any telephone system that provides adequate service the probability of a substantially larger than average number of calls in progress will be small. Thus, in using the formula for probability of blocking,

$$\Pr\{\text{bl}\} = \frac{\sum_{x \in S} p_x \beta_x}{\sum_{x \in S} p_x \alpha_x},$$

it should be possible to omit states with more than k calls in progress from the sums, without incurring too much error. It is the purpose of this section to examine this possibility rigorously. In particular, we wish to answer the following very important question: If p is a given positive number, how large must k be so that the omission of states with more than k calls in progress, i.e., the approximation

$$\Pr\{\text{bl}\} \approx \frac{\sum_{|x| \leqq k} p_x \beta_x}{\sum_{|x| \leqq k} p_x \alpha_x}$$

results in an error of at most $100p$ per cent?

In what follows we shall make systematic use of the following abbreviations:

$$r = \sum_{|x| \leqq k} p_x s(x) \tag{9}$$

$$s = \sum_{|x| \leqq k} p_x \alpha_x \tag{10}$$

$$u = \sum_{|x| > k} p_x s(x) \tag{11}$$

$$v = \sum_{|x| > k} p_x \alpha_x .$$

(The notation b for the probability of blocking, used in Ref. 3, e.g., is being avoided in favor of $\Pr\{\text{bl}\}$.) It can be seen that

$$r + u = \frac{\text{success rate}}{\lambda} = \frac{m}{\lambda}$$

$$s + v = \frac{\text{attempt rate}}{\lambda}$$

$$1 - \Pr\{\text{bl}\} = \frac{r + u}{s + v}.$$

Thus, omitting states with more than k calls in progress in calculating $1 - \Pr\{\text{bl}\}$ is equivalent to approximating it by

$$r/s. \tag{12}$$

We note that

$$v/(s + v)$$

is the fraction of attempts made with more than k calls in progress.

Lemma 4: $v/(s + v) \leqq Pr\{|x| > k\}$.

Proof: We have

$$\frac{v}{s + v} = \frac{\sum_{|x|>k} p_x\alpha_x}{\sum_{|x|\leqq k} p_x\alpha_x + \sum_{|x|>k} p_x\alpha_x}.$$

Let $\alpha = \max\{\alpha_x : |x| = k\}$. Then because $\alpha_{(\cdot)}$ is antitone on S

$$\frac{v}{s + v} \leqq \frac{v}{\alpha \Pr\{|x| \leqq k\} + v}.$$

Since for $\mu > 0$

$$\frac{d}{dt}\frac{t}{\mu + t} = \frac{\mu}{(\mu + t)^2} > 0$$

we can replace v in the last inequality by its majorant

$$\alpha \Pr\{|x| > k\},$$

which proves the lemma.

Also, it is seen that

$$\frac{v - u}{s + v - r - u} = \frac{\sum_{|x|>k} p_x\beta_x}{\sum_{x\epsilon S} p_x\beta_x}$$

is the fraction of blocked attempts that occur when more than k calls are in progress.

Theorem 5: If, simultaneously, the fraction of attempts that are blocked with more than k calls in progress is at most $p/(1 + p)$, *and* $Pr\{|x| > k\} \leqq p/(1 + p)$, *then omitting states with more than k calls in progress when using* (1) *for* $Pr\{bl\}$ *will not result in an error of more than* 100p *per cent.*

Proof: It suffices to show that

$$\left|\frac{r + u}{s + v} - \frac{r}{s}\right| \leqq p\left(1 - \frac{r + u}{s + v}\right).$$

The hypothesis and Lemma 4 imply that

$$\frac{v}{s+v} \leqq \frac{p}{1+p}$$

and hence that

$$v \leqq ps,$$

$$v\left(1 - \frac{r}{s}\right) \leqq p\,(s - r)$$

$$(1 + p)\,(r + v) \leqq p\,(s + v) + \frac{r}{s}\,(s + v).$$

Since $u \leqq v$, we have

$$(1 + p)\,(r + u) \leqq \left(p + \frac{r}{s}\right)(s + v)$$

or

$$\frac{r+u}{s+v} - \frac{r}{s} \leqq p\left(1 - \frac{r+u}{s+v}\right),$$

which is one half of the requisite inequality. For the other half, the hypothesis gives

$$\frac{v-u}{s-r} \leqq p$$

$$v - u \leqq p(s - r)$$

$$\frac{r}{s}v \leqq u + p\,(s - r)$$

$$\leqq pv + (1 - p)u + p(s - r)$$

$$= u + p(s + v - r - u)$$

so that

$$\frac{v\frac{r}{s} - u}{s+v} \leqq p\left(1 - \frac{r+u}{s+v}\right)$$

or

$$\frac{r+u}{s+v} - \frac{r}{s} \geqq -p\left(1 - \frac{r+u}{s+v}\right).$$

Since $\beta \leqq \alpha$,

$$\frac{v - u}{s + v - r - u} \leqq \frac{\alpha_{k+1} \Pr\{| x_t | > k\}}{s + v - r - u}.$$

Hence, the hypotheses of Theorem 5 are satisfied if both

$$\Pr\{| x_t | > k \} \leqq \frac{p}{1 + p}$$

and

$$\Pr\{| x_t | > k\} \leqq \frac{p}{1 + p} \frac{s + v - r - u}{\alpha_{k+1}} = \frac{p}{p + 1} \frac{\Pr\{\text{bl}\}}{\alpha_{k+1}/(s + v)}$$

or if both

$$\left.\begin{aligned} \Pr\{| x_t | > k\} &\leqq \frac{p}{1 + p} \\ \Pr\{\text{bl}\} &\leqq \frac{\binom{T - 2k - 2}{2}}{s + v}. \end{aligned}\right\} \qquad (13)$$

We now show that the first inequality in (13) is easily met by a choice of k that depends very simply on p and on the carried load m, and that with this choice of k, the second inequality holds for virtually all cases of interest. By Chebyshev's inequality, the first inequality is satisfied if

$$k \geqq \frac{p + 1}{p} m - 1,$$

for then

$$\Pr\{| x_t | > k\} \leqq \frac{m}{k + 1} \leqq \frac{p}{1 + p}.$$

As for the second, we have

$$s + v = \binom{T - 2m}{2} + 2\sigma^2,$$

$$0 \leqq 4\sigma^2/T^2 \leqq 1.$$

Thus,

$$\frac{\binom{T - 2k - 2}{2}}{s + v} \geqq \frac{\binom{T - 2k - 2}{2}}{\binom{T - 2m}{2}} \cdot \frac{\binom{T - 2m}{2}}{\binom{T - 2m}{2} + \frac{1}{2}T^2}$$

Since $m \gg 2p$, choosing $k = m + (m/p) - 1$ makes the first factor greater than unity. The second factor is, with $q = 2m/T =$ line usage

$$\frac{1}{1 + \dfrac{\frac{1}{2}T^2}{\left(\dfrac{T - 2m}{2}\right)}} = \frac{1}{1 + \dfrac{T^2}{(T - 2m)^2 - T + 2m}} = \frac{1}{1 + \dfrac{1}{(1 - q)^2} + o(1)}$$

as T becomes large. As q assumes values in the representative range 0 to 0.2, the second factor varies between 0.5 and about 0.39, if the $o(1)$ term is ignored. The strongest form of the second inequality in (13) is then roughly

$$\Pr\{\mathrm{bl}\} \leqq 0.4,$$

and is virtually always fulfilled in cases of practical interest. Thus, for example, to obtain an error of at most 25 per cent, it is sufficient to consider only states with at most

$$5m - 1$$

calls in progress. For a 50 per cent error, only states with at most

$$3m - 1$$

calls in progress need be considered.

In many cases, especially in those in which very little is known about the actual value of the probability $\Pr\{\mathrm{bl}\}$ of blocking, it may be desirable to assess the effect of neglecting states with more than k calls in progress on the *absolute* error rather than the *percentage* error. This situation is covered by the following simple result:

Theorem 6: Let $\epsilon > 0$ *be any positive real number. If* $Pr\{|x| > k\} \leqq \epsilon$, *then omitting states with more than k calls in progress in calculating* $\Pr\{\mathrm{bl}\}$ *by* (I) *will not result in an absolute error of more than* ϵ.

Proof: It is sufficient to establish that

$$\left|\frac{r + u}{s + v} - \frac{r}{s}\right| \leqq \epsilon.$$

By Lemma 4, we have

$$\frac{v}{s + v} \leqq \Pr\{|x| > k\} \leqq \epsilon$$

and hence, using $u \leqq v$,

$$\frac{v}{s+v} \leqq \frac{r}{s} + \epsilon - \frac{r}{s+v}$$

$$\frac{r+u}{s+v} - \frac{r}{s} \leqq \epsilon.$$

For the other half of the requisite inequality, we observe that

$$\epsilon + \frac{u}{s+v} \geqq \epsilon \frac{r}{s}$$

$$\frac{u}{s+v} - \frac{r}{s} + \frac{r}{s}(1-\epsilon) \geqq -\epsilon.$$

From the hypothesis we have

$$\frac{v}{s+v} \leqq \epsilon$$

$$\frac{s}{s+v} \geqq 1 - \epsilon$$

$$\frac{r}{s+v} \geqq \frac{r}{s}(1-\epsilon),$$

and hence,

$$\frac{r+u}{s+v} - \frac{r}{s} \geqq -\epsilon,$$

which proves the result.

XIII. THE SIGN OF THE ERROR

In the two preceding theorems we have studied the approximation $1 - (r/s)$ to $\Pr\{\mathrm{bl}\}$ (obtained by omitting from the sums in (1) states with more than k calls in progress) without considering whether this approximation will tend to overestimate or underestimate $\Pr\{\mathrm{bl}\}$. This question is now taken up.

Various intuitive arguments why $1 - (r/s)$ should lie on one side or the other of $\Pr\{\mathrm{bl}\}$ come readily to mind. The number β_x of blocked idle inlet-outlet pairs in state x tends first to grow with $|x|$, but then as $|x|$ becomes large enough it must again decrease to zero, because $\beta_x \leqq \alpha_x =$ number of idle inlet-outlet pairs in state x. However, if the network cannot carry more than w calls with $w \ll \frac{1}{2}T$, it is possible that β_x is actually monotone increasing (or isotone) with respect to the partial ordering $\leqq$ of S; since α_x is definitely monotone decreasing (or antitone) on $(S,\leqq)$, one might in this case expect that omitting the states where β_x is largest and α_x is smallest would tend to make the congestion seem to be *less* than it actually is.

Similarly, viewing (r/s) as an approximation to

$$\frac{\sum_{x \in S} p_x s(x)}{\sum_{x \in S} p_x \alpha_x} = 1 - \Pr\{\text{bl}\} \tag{14}$$

and noting that $s(\cdot)$ is antitone on $(S,\leqq)$, one might expect that omitting the states where $s(\cdot)$ is smallest would tend to make $1 - \Pr\{\text{bl}\}$ larger than it is.

In fact, neither of the above intuitions is always correct; omission of states x with more than k calls in progress from the sums in the ratio (14) defining $1 - \Pr\{\text{bl}\}$ sometimes gives an underestimate, and at others gives an overestimate. Roughly, if k is large enough, $1 - (r/s)$ will be an overestimate of the loss, whereas if it is too small, it will be an underestimate.

Theorem 7: If the fraction of hangups made with more than $k + 1$ calls in progress exceeds the fraction of attempts made with more than k calls in progress, then omitting states with more than k calls in progress in the calculation of $\Pr\{\text{bl}\}$ *results in an overestimate; in the opposite case, the omission results in an underestimate.*

Proof: For $t \in [0,1]$, let

$$U(t) = \frac{r + ut}{s + vt}$$

so that $U(0) = r/s$ and $U(1) = 1 - \Pr\{\text{bl}\}$. It can be seen that

$$\frac{\sum_{j > k+1} j p_j}{\sum_{i=0}^{w} i p_i} = \frac{\sum_{|x| > k} s(x) p_x}{\sum_{x \in S} s(x) p_x} = \frac{u}{r + u},$$

$$\frac{\sum_{|x| > k} \alpha_x p_x}{\sum_{x \in S} \alpha_x p_x} = \frac{v}{s + v},$$

and that the following inequalities are all equivalent:

$$rv \geqq us$$

$$r(s + v) + ut(s + v) \geqq s(r + u) + tv(r + u)$$

$$U(t) \geqq U(1)$$

$$\frac{v}{s + v} \geqq \frac{u}{r + u}.$$

Theorem 8: If the fraction $s/(s + v)$ of attempts made with at most k calls in progress is at most

$$\frac{1 - E(k + 1,a)}{1 - E(w,a)} \frac{E(w,a)}{E(k + 1,a)} \frac{w!}{k + 1!} a^{k+1-w} \tag{15}$$

then omitting states with more than k calls in progress in calculating $Pr\{bl\}$ results in an underestimate:

$$Pr\{bl\} \geqq 1 - (r/s).$$

Proof: It can be verified that

$$1 - \frac{1 - E(k + 1,a)}{1 - E(w,a)} \frac{E(w,a)}{E(k + 1,a)} \frac{w!}{k + 1!} a^{k+1-w}$$

$$= \frac{\sum_{j=k+2}^{w} j \frac{a^j}{j!}}{\sum_{j=0}^{w} j \frac{a^j}{j!}} \geqq \frac{\sum_{j=k+2}^{w} jb_j}{\sum_{j=0}^{w} jb_j} \geqq \frac{\sum_{j>k+1}^{w} jp_j}{\sum_{j=0}^{w} jp_j} = \frac{\lambda \sum_{|x|>k} p_x s(x)}{m} = \frac{\lambda u}{m}.$$

The first equality follows from known properties of Erlang's function, the two inequalities follow from the basic lemma with $g \equiv 1$ and

$$f_j = \begin{cases} 0 & j \leqq k + 1 \\ 1 & j > k + 1, \end{cases}$$

and the last two equalities follow from (6) and the definition (11) of u, respectively. Thus the hypothesis gives

$$\frac{v}{s + v} > \lambda \frac{u}{m}$$

$$\frac{u}{v} < \frac{1}{\lambda} \frac{m}{s + v} = \frac{r + u}{s + v} = 1 - \Pr\{\text{bl}\},$$

and the argument now proceeds as in Theorem 7.

Remark:

$$\frac{s}{s+v} \leqq \frac{\sum_{j>k}^{} \alpha_j b_j}{\sum_{j=0}^{w} \alpha_j b_j} \leqq \frac{\sum_{j>k}^{w} \alpha_j \frac{a^j}{j!}}{\sum_{j=0}^{w} \alpha_j \frac{a^j}{j!}}.$$

Proof: Basic lemma, with $\lambda_j = \alpha_j b_j$.

It follows that

$$\frac{v}{s+v} = 1 - \frac{s}{s+v} \geqq 1 - \frac{\sum_{j>k} \alpha_j b_j}{\sum_{j=0}^{w} \alpha_j b_j} \geqq 1 - \frac{\sum_{j>k}^{w} \alpha_j \frac{a^j}{j!}}{\sum_{j=0}^{w} \alpha_j \frac{a^j}{j!}}.$$

Theorem 9: If the fraction $\lambda u/m$ of hangups made with $k + 1$ or fewer calls in progress is at least

$$1 - a^{k-w} \frac{w!}{k!} \frac{E(w,a)}{E(k,a)} \cdot \frac{\binom{T}{2} - a(2T - 3 + 2a)[1 - E(k,a)] - 2akE(k,a)}{\binom{T}{2} - a(2T - 3 + 2a)[1 - E(w,a)] - 2awE(w,a)}, \tag{16}$$

then omitting states with more than k calls in progress in calculating $Pr\{bl\}$ results in an overestimate:

$$Pr\{bl\} \leqq 1 - (r/s).$$

Proof: It can be verified, using the formula, for integers $c \leqq w$,

$$\sum_{j=0}^{c} \binom{T-2j}{2} \frac{a^j}{j!} = \frac{a^c}{c!} \frac{\binom{T}{2} - a(2T - 3 + 2a)[1 - E(c,a)] - 2acE(c,a)}{E(c,a)},$$

that

$$(16) = \frac{\sum_{j=k+1}^{w} \alpha_j \frac{a^j}{j!}}{\sum_{j=0}^{w} \alpha_j \frac{a^j}{j!}} \geqq \frac{\sum_{j=k+1}^{w} \alpha_j b_j}{\sum_{j=0}^{w} \alpha_j b_j}$$

$$\geqq \frac{\sum_{j=k+1}^{w} \alpha_j p_j}{\sum_{j=0}^{w} \alpha_j p_j} = \frac{v}{s+v},$$

where the first equality follows from the stated formula, the two inequalities follow from the basic lemma, and the last equality from the definitions of s and v. Hence, the hypothesis implies

$$\frac{\lambda u}{m} \geqq \frac{v}{s+v},$$

$$\frac{u}{r+u} \geqq \frac{v}{s+v},$$

and the argument again proceeds as in Theorem 7.

XIV. OTHER APPROXIMATIONS

Since

$$1 - \Pr\{\text{bl}\} = \frac{1}{\lambda} \frac{\sum_{j=1}^{w} j p_j}{\sum_{j=0}^{w} \alpha_j p_j}, \tag{17}$$

one can envisage an approximation

$$\Pr\{\text{bl}\} \approx 1 - \frac{1}{\lambda} \frac{\sum_{j=1}^{k} j p_j}{\sum_{j=0}^{k} \alpha_j p_j}, \tag{18}$$

obtained by omitting states with more than k calls in progress from the sums in (17). The basic lemma implies that this approximation is *always an overestimate.* We have

Theorem 10: For each $k = 1, \cdots, w$,

$$1 - Pr\{bl\} \geqq \frac{\lambda^{-1} \sum_{j=1}^{k} j p_j}{\sum_{j=0}^{k} \alpha_j p_j} = \frac{\sum_{|x| \leqq k-1} s(x) p_x}{\sum_{|x| \leqq k} \alpha_x p_x}.$$

Proof:

$$1 - \Pr\{\text{bl}\} = \frac{1}{\lambda} \frac{\sum_{j=1}^{w} j p_j}{\sum_{j=0}^{w} p_j \alpha_j}$$

$$\geqq \frac{1}{\lambda} \frac{\sum_{j=1}^{k} j p_j}{\sum_{j=0}^{k} p_j \alpha_j},$$

the inequality following from the basic lemma. We now use formula (6):

$$jp_j = \lambda \sum_{|y|=j-1} p_y s(y), \qquad j = 1, \cdots, w.$$

We note that the bound given in Theorem 10 is exactly the estimate r/s of (12) with the top term in the numerator sum omitted. This term is just $\lambda^{-1}(k+1)p_{k+1}$, and we have shown that

$$\begin{aligned}
\Pr\{\text{bl}\} - 1 + \frac{r}{s} &\leqq \frac{(k+1)p_{k+1}}{\lambda s} \\
&\leqq \frac{(k+1)b_{k+1}}{\lambda \sum_{j=0}^{k} b_j \alpha_j} \\
&\leqq a \frac{\frac{a^k}{k!}}{\lambda \sum_{j=0}^{k} \frac{a^j}{j!} \binom{T-2j}{2}}, \qquad \left(\text{with } a = \frac{\lambda T^2}{2}\right) \\
&\leqq \frac{aE(k,a)}{\lambda \binom{T}{2} + \lambda ak - \lambda a(2T - 2a^2 + 3a + k)[1 - E(k,a)]}.
\end{aligned}$$

The last bound goes to $E(k,a)$ as $\lambda \to 0$, $T \to \infty$, with $2a = \lambda T^2$. In this limit, then, if $1 - r/s$ underestimates $\Pr\{\text{bl}\}$ at all, it does so by at most $E(k,a)$.

Another result of the same character is

Theorem 11: With $a = \frac{1}{2}\lambda T^2$ and $k + 2 \leqq w$,

$$Pr\{bl\} \leqq 1 - \frac{r}{s} + \frac{k+2}{a} \frac{E(k+2,a)}{1 - E(k+2,a)}.$$

Proof:

$$\begin{aligned}
1 - \Pr\{\text{bl}\} &= \frac{1}{\lambda} \frac{\sum_{j=1}^{w} jp_j}{\sum_{j=0}^{w} \alpha_j p_j} \geqq \frac{1}{\lambda} \frac{\sum_{j=1}^{k+1} jp_j}{\sum_{j=0}^{k+1} \alpha_j p_j} \\
&= \frac{r}{s + \alpha_{k+1} p_{k+1}},
\end{aligned}$$

the inequality coming from the basic lemma, and the second identity from formula (6) and the definitions (10) and (11) of r and s respectively. Writing the last term on the right as

$$\frac{r}{s} \cdot \frac{s}{s + \alpha_{k+1}p_{k+1}} = \frac{r}{s} \cdot \left(1 - \frac{\alpha_{k+1}p_{k+1}}{s + \alpha_{k+1}p_{k+1}}\right)$$

we find, since $r/s < 1$,

$$1 - \Pr\{\mathrm{bl}\} - \frac{r}{s} \geqq - \frac{\alpha_{k+1}p_{k+1}}{\sum_{j=0}^{k+1} \alpha_j p_j}.$$

The basic lemma gives, using $jb_j = \lambda\alpha_{j-1}b_{j-1}$,

$$\frac{\alpha_{k+1}p_{k+1}}{\sum_{j=0}^{k+1} \alpha_j p_j} \leqq \frac{\alpha_{k+1}b_{k+1}}{\sum_{j=0}^{k+1} \alpha_j b_j} = \frac{(k+2)b_{k+2}}{\sum_{j=1}^{k+2} jb_j}$$

$$\leqq (k+2) \frac{\dfrac{a^{k+2}}{(k+2)!}}{\sum_{j=1}^{k+2} j \dfrac{a^j}{j!}} = \frac{k+2}{a} \frac{E(k+2,a)}{1 - E(k+2,a)}.$$

Theorem 12: Let K be a set of integers j all satisfying $j > \lambda\alpha_j[1 - Pr\{bl\}]$. Then omission of all the states in $\bigcup_{j \in K} L_j$ in the calculation of $Pr\{bl\}$ as defined by (2) results in an overestimate.

Proof: Let $\xi = (\xi_0, \cdots, \xi_w)$ be a $(w + 1)$ dimensional vector variable taking values in the positive orthant, and consider the function $V(\xi)$ defined by

$$V(\xi) = \frac{1}{\lambda} \frac{\sum_{j=1}^{w} j\xi_j}{\sum_{j=0}^{w} \alpha_j \xi_j}.$$

It is apparent that if $\xi = (p_0, \cdots, p_w) = p =$ distribution of the number of calls in progress, then

$$V(p) = 1 - \Pr\{\mathrm{bl}\}.$$

Now,

$$\frac{\partial V}{\partial \xi_j} = \frac{1}{\lambda} \frac{j - \lambda\alpha_j V(\xi)}{\sum_{j=0}^{w} \alpha_i \xi_j}.$$

Hence, $V(\xi) \leqq V(p)$ and $j \in K$ imply

$$\frac{\partial V}{\partial \xi_j} > 0.$$

Consider a path of integration Γ along which

$$\xi_j = p_j \qquad j \notin K$$

and which runs from the point ξ_s with coordinates

$$\begin{cases} 0, & \text{for } j \in K \\ p_j & \text{for } j \notin K \end{cases}$$

to the point $\xi = p$ in such a way that $d\xi_j/ds > 0$ for $j \in K$ along Γ. $V(\xi_s)$ is the approximation to $1 - \Pr\{\text{bl}\}$ resulting from omitting from (17) states having j calls in progress for $j \in K$.

It is apparent that there is a segment of Γ in the neighborhood of p on which $V(\xi) \leqq V(p)$. Since $V(\cdot)$ is continuous the set $A = \{\xi: V(\xi) \leqq V(p)\}$ is closed. If Γ first intersects ∂A at some point $q \neq p$ we have

$$V(p) = V(q) = V(p) - \int_q^p \sum_{j \in K} \frac{\partial V}{\partial \xi_j} \frac{d\xi_j}{ds} \, ds$$

which is impossible since the integral does not vanish. Thus

$$V(\xi_s) \leqq V(p).$$

It is easy to see that the condition $j > \lambda \alpha_j [1 - \Pr\{\text{bl}\}]$ in Theorem 12 occurs for relatively low values of j. For it is enough that

$$j > \frac{\lambda T^2}{2} [1 - \Pr\{\text{bl}\}] = \frac{m}{(1 - q)^2 - T^{-1}(1 - q) + 4\sigma^2 T^{-2}},$$

and thus it suffices that

$$j > \frac{m}{(1 - q)^2 + T^{-1}(1 - q)}.$$

The second term in the denominator is negligible for all but uninterestingly small values of T, so roughly j can be any integer larger than

$$\frac{m}{(1 - q)^2}.$$

With $q = 0.1$ erlang, a representative value, the condition is approximately $j > 1.22m$.

The method used in Theorem 12 also proves

Theorem 13: Let X be a set of states such that $x \in X$ implies

$$\frac{s(x)}{\alpha_x} > 1 - Pr\{bl\}.$$

Then the approximation

$$Pr\{bl\} \approx 1 - \frac{\sum_{x \in S-X} s(x) p_x}{\sum_{x \in S-X} \alpha_x p_x}$$

is an overestimate.

XV. INEQUALITY FOR PROBABILITY OF BLOCKING

Last, yet we hope not least, we give a basic inequality for the probability of blocking itself. The result to be given clearly shows how *combinatorial knowledge* about the connecting network of interest (in this case information about how fast the number of blocked pairs goes up with the number of calls in progress) can be used to give an upper bound on the loss.

Theorem 14: Let $\beta_x \leqq f_{|x|}$ for nondecreasing $f(\cdot)$. Then

$$Pr\{bl\} \leqq \frac{\sum_{j=0}^{w} f_j b_j}{\sum_{j=0}^{w} \alpha_j b_j} \leqq \frac{\sum_{j=0}^{w} f_j \frac{a^j}{j!}}{\sum_{j=0}^{w} \alpha_j \frac{a^j}{j!}}.$$

Proof:

$$\begin{aligned}
\Pr\{\mathrm{bl}\} = \frac{\sum_{x \in S} \beta_x p_x}{\sum_{x \in S} \alpha_x p_x} &\leqq \frac{\sum_{j=0}^{w} f_j p_j}{\sum_{j=0}^{w} \alpha_j p_j} \\
&\leqq \frac{\sum_{j=0}^{w} f_j b_j}{\sum_{j=0}^{w} \alpha_j b_j} \\
&\leqq \frac{\sum_{j=0}^{w} f_j \frac{a^j}{j!}}{\sum_{j=0}^{w} \alpha_j \frac{a^j}{j!}},
\end{aligned}$$

with $a = \frac{1}{2}\lambda T^2$. The first inequality follows from the hypothesis and the assumed one- or two-sided nature of the network; the ensuing two inequalities follow from the basic lemma.

The foregoing theorem makes it plain that much is to be learned about congestion in a connecting network from a study of the rate at which the special function $\{\beta_x, x \in S\}$ changes with $|x|$. The search for bounds of the form

$$\beta_x \leqq f_{|x|}, \qquad x \in S$$

(with f_j increasing) for various kinds or classes of connecting networks now becomes one of the next most important problems in the endeavor to bring, by purely analytical methods, A. K. Erlang's dynamical theory of telephone traffic to belated but final fruition. This problem is beyond the scope of this paper; some elementary phases of it are considered in a later paper.[2]

XVI. ACKNOWLEDGMENT

The author is indebted to W. S. Hayward, Jr., for a careful reading of the draft, and for pointing out an error in the proof of an earlier version of Theorem 5.

REFERENCES

1. Beneš, V. E., *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, New York, 1965.
2. Beneš, V. E., Some Bounds for Loss in Crossbar Networks, to be published.
3. Beneš, V. E., Markov Processes Representing Traffic in Connecting Networks, B.S.T.J., *42*, 1963, pp. 2795–2838.
4. Lee, C. Y., Analysis of Switching Networks, B.S.T.J., *34*, 1955, pp. 1287–1315.
5. Grantges, R. F., and Sinowitz, N. R., NEASIM: A General-Purpose Computer Simulation Program for Load-Loss Analysis of Multistage Central Office Switching Networks, B.S.T.J., *43*, 1964, pp. 965–1004.
6. Beneš, V. E., Properties of Random Traffic in Nonblocking Telephone Connecting Networks, B.S.T.J., *44*, 1965, pp. 509–525.