THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 6, July–August 1977
Printed in U.S.A.

Faulty-Trunk Detection Algorithms Using EADAS/ICUR Traffic Data

By J. S. KAUFMAN

(Manuscript received November 1, 1976)

A class of algorithms for detecting abnormally short-holding-time trunks has been developed that utilizes individual trunk data available in EADAS/ICUR (Engineering and Administrative Data Acquisition System/Individual Circuit Usage Recorder). This data consists of a two-dimensional statistic that compresses the raw trunk measurements—the state of the trunk (busy or idle) sampled every 100 or 200 seconds—into a manageable form. Because this data is essentially a sufficient statistic for the stochastic process used to model the (unobservable) trunk state measurements, one of the algorithms developed is Wald's sequential probability ratio test. Two of the algorithms developed have been implemented in ICAN (Individual Circuit Analysis Program), and are currently being used to test trunks associated with the No. 1 crossbar, No. 5 crossbar, crossbar tandem (1XB, 5XB, XBT), and step-by-step switching machines. The focus in this paper, however, is on the modeling and analysis aspects of the problem, and only slight attention is paid to the various trade-offs and real-world constraints encountered in implementing the algorithms.

I. INTRODUCTION

A message trunk, the basic connecting link in the switched telephone network, provides the communication path between switching machines as well as certain call setup capabilities, such as supervision, signaling, and ringing. For an important class of trunk faults that cause call failure, the trunk is released by the switching system upon customer abandonment and is again available to fail another call. As a result, a single undetected faulty trunk of this type can fail a significant fraction of the offered attempts to a group and will characteristically have an abnormally short holding time.

Because of their potential service impact, significant efforts have been made to understand and quantify the impact that such abnormally short-holding-time trunks have on central office and network service.[1–4] It is now widely understood as a result of these studies that this generic trunk fault results in a fraction of service attempts "killed," which is out of all proportion to their number in the trunk population. Consequently, traffic data available from new and existing traffic data-acquisition systems has been viewed in the light of increasing interest in trunk-fault detection. In particular, with the advent of the Bell System EADAS/ICUR (Engineering and Administrative Data Acquistion System/Individual Circuit Usage Recorder) system,[5] it was natural to ask whether the new individual trunk data available could be used to detect such "killer" trunks.*

This paper discusses the theoretical aspects of a class of killer-trunk detection algorithms that utilize the individual trunk traffic data available in EADAS/ICUR. These algorithms were designed for, and practical versions of them are presently implemented in, the ICAN† portion of the EADAS/ICUR system. We focus here, however, on the problem formulation, modeling, and analysis aspects of the algorithms without bringing in many of the diverse factors and trade-offs encountered in the actual implementation.

Because the holding time of a trunk affects the statistical properties of the trunk data in EADAS/ICUR, it is natural to formulate the killer-trunk detection problem as a problem in the testing of statistical hypotheses. In this context our modeling effort is basically an attempt to precisely define the state of a trunk (normal or killer) and expose the relevant underlying distributions. Well-known aspects of the theory of hypothesis testing are then applicable and immediately suggest a number of different tests. Sequential tests are naturally considered since the EADAS/ICUR data evolve sequentially in time. Questions about the robustness of the models, and the structure and performance of statistical tests, are addressed using standard analytic tools.

The material in this paper has been organized into six major sections, whose content we briefly describe. After considering the data available in EADAS/ICUR (Section II), we proceed to model a trunk (Section 3.1), motivate an appropriate set of statistical hypotheses suitable to our problem (Section 3.2), and briefly review several classical tests for deciding between statistical hypotheses (Section 3.3). With these preliminaries out of the way, we develop individual trunk algorithms based solely on individual trunk data. Proceeding in a heuristic manner, we use the individual trunk data to "derive" an *ad hoc* killer-trunk-detection

* The term killer trunk has been widely adopted in referring to a faulty switching-machine-accessible trunk in a group whose average holding time is substantially smaller than the average group holding time.

† Individual circuit analysis program—a software program that analyzes much of the EADAS/ICUR traffic data and maintains the system data base.

algorithm (Section 4.1). Although the insight gained in proceeding in a heuristic fashion is significant, we shift our emphasis in Section 4.2 and rigorously derive an optimal test statistic. It is interesting to find that the ad hoc statistic is essentially one of two symmetric statistics which comprises the optimal test statistic. The relationship between these individual trunk statistics is further explored in Section 4.3.

In Section V we factor grouping information (which essentially identifies all trunks common to a trunk group) into the picture, and develop detection algorithms tailored to trunks associated with the No. 5 crossbar switching machine. This development necessitates modeling the 5XB trunk-group selection procedure, and several results due to Forys and Messerli[2] are utilized. In Section VI we shift our discussion to the performance of the 5XB group algorithms, deriving approximate expressions for the mean statistic update and mean detection time in Sections 6.1 and 6.2, respectively. The paper concludes in Section 6.3 with an approximate analysis of the false-alarm probability of the 5XB group algorithms.

II. EADAS/ICUR DATA

The structure of a killer-trunk detection algorithm is largely dependent on the type of individual trunk measurements available.* In EADAS/ICUR, the raw (unobservable) data consists of the state of each trunk (busy or idle) every 100 or 200 seconds. Fortunately, the data accumulations available essentially summarize all the relevent information in the raw data.

2.1 Switch-count and transition data

The data available from the EADAS/ICUR system, which can be used to distinguish between normal and killer trunks, is obtained by sampling individual trunks at 100 or 200 second intervals. This data consists of periodic accumulations (typically hourly, two-hourly, or three-hourly) of both the Busy states, and the State transitions. The busy state accumulation is usually referred to as the switch count. For the 200-second sampling option with a one-hour accumulation period, the switch-count is an integer between 0 and 18. A state transition occurs whenever the state of a trunk (busy or idle) is different at two successive scans. For the 200-second sampling option with a one-hour accumulation period, the number of state transitions is an integer between 0 and 17.

If we denote the ith scan during an accumulation period in which m scans occur by x_i, and let 0 and 1 correspond to trunk idle and trunk

* Until very recently, almost all trunk-traffic measurements were obtained on a group rather than on an individual trunk basis.

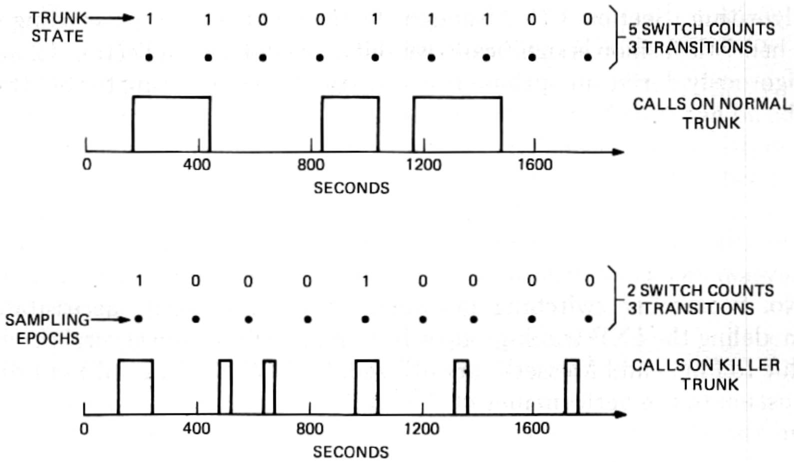


Fig. 1—KCUR data.

busy, respectively, the available data may be written

$$(i)\quad n(m) = \sum_{i=1}^{m} x_i \text{ (switch count)}$$

$$(ii)\quad t(m) = \sum_{i=2}^{m} |x_i - x_{i-1}| \text{ (state transitions).}$$

Thus, the raw (unavailable) data in the form of a binary sequence,

$$\mathbf{x}_m = (x_1, \cdots, x_m),$$

is compressed into the two statistics $n(m)$ and $t(m)$.

Because the holding time of a killer trunk is generally on the order of a few tens of seconds, it should have substantially more state transitions than a normal trunk, for a given switch count. Figure 1 illustrates the sampling process on both a normal and killer trunk.* For the purposes of this figure individual calls are represented by rectangles, call durations correspond to the width of the rectangles, and a half-hour accumulation period with the 200-second sampling option is used.

We note in passing that for the 200-second sampling option, very little information is lost by "compressing" the raw data $\mathbf{x}_m = (x_1, \cdots, x_m)$ into the two statistics $n(m)$ and $t(m)$. Thus, normal conversation lengths tend to be in the vicinity of 3 to 4 minutes and, hence, with the 200-second sampling option, we expect that only adjacent samples are significantly

* The realizations shown in Fig. 1 are more or less typical for a 5XB trunk group with a mean group holding time of approximately 4 minutes operating at about 40-percent occupancy, and having a killer trunk with a mean holding time of approximately 1 minute.

correlated. If $\{x_i\}_{i=1}^m$ is Markovian, then $[n(m),t(m)]$ is almost a sufficient statistic ($[n(m),t(m),x_1,x_m]$ is sufficient) for $\mathbf{x}_m$ (see Section 4.2). Note also that for a trunk in the killer state, succesive samples should be essentially independent (for both sampling options).

2.2 Grouping information

In addition to the switch-count and transition data available from the EADAS/ICUR system, we are also able to utilize a system map to identify (*i*) all trunks common to a trunk group, and (*ii*) the trunk-selection procedure* associated with the trunk group. It turns out that using this grouping information,† in addition to the switch-count and transition data, enhances the detection potential considerably.

Thus, we divide the class of algorithms into two types according to whether or not grouping information is utilized. The first type, which uses only switch-count and transition data is referred to as an individual trunk algorithm. These individual trunk algorithms are applicable to all trunks—including two-way trunks—independent of the type of switching machine they are associated with. They do however assume knowledge of the trunks nominal holding time. The second type of algorithm uses the grouping information in addition to the switch-count and transition data and is referred to as a group algorithm. Group algorithms are "tailored" to a specific kind of trunk-selection procedure and, hence, apply to trunk groups associated with specific switching machines. For the purposes of this paper, the trunk-selection procedure considered is random selection of idle trunks. This procedure models the selection procedure of trunk groups associated with the 5XB switching machine. Group algorithms generally apply only to one-way trunk groups.

III. PRELIMINARY CONSIDERATIONS

In attempting to quantify the intuitive notion that a killer trunk will exhibit more transitions than a normal trunk (see Fig. 1), for a given switch count, it is natural to consider the transition probabilities:

$$P_{1,0} = P\{x_{t+\tau} = 0/x_t = 1\}$$

and

$$P_{0,1} = P\{x_{t+\tau} = 1/x_t = 0\},$$

---

* The map in EADAS/ICUR indicates the type of switching machine that the trunk group is associated with, and this allows us to model the trunk-selection procedures (see Ref. 2).

† We will consistently use "grouping information" to refer to both the identification of all trunks common to the group and the trunk-selection procedure associated with the group.

where x_t denotes the state of a trunk at epoch t and τ denotes the sampling interval. Of course, to evaluate these transition probabilities, we must be concrete about how we model a trunk.

Before tying ourselves down to any specific model, however, it is useful to view these conditional probabilities in a canonical form. Thus, suppose we begin by assuming only that the binary valued process x_t is stationary. We have then the following simple result:

Lemma 1: Let x_t be a binary valued, stationary random process and let ρ and $R(\cdot)$ denote its mean and covariance function, respectively. Then,

$$(i)\ \ P_{1,0}(\rho,\tau) = (1-\rho)\left\{1 - \frac{R(\tau)}{R(0)}\right\} \tag{1}$$

$$(ii)\ \ \rho P_{1,0}(\rho,\tau) = (1-\rho)P_{0,1}(\rho,\tau). \tag{2}$$

Proof: Part (i) is a consequence of the definition of $R(\cdot)$. That is,

$$\begin{aligned} R(\tau) &= E(x_t x_{t+\tau}) - \rho^2 \\ &= P(x_t = 1, x_{t+\tau} = 1) - \rho^2, \end{aligned}$$

where

$$\rho = E(x_t) = P(x_t = 1).$$

Part (ii) follows from the two identities:

$$\rho = \rho P_{1,1}(\rho,\tau) + (1-\rho)P_{0,1}(\rho,\tau)$$

and

$$1 = P_{1,1}(\rho,\tau) + P_{1,0}(\rho,\tau).$$

A consequence of this result is that uncorrelatedness and independence are equivalent:

Corollary 1: For the process in lemma 1, $x_t, x_{t+\tau}$ *are independent if and only if* $R(\tau) = 0$.

Note that the dependence of $R(\cdot)$ on ρ has been suppressed for convenience.

3.1 Modeling an individual trunk

A particularly simple way to model a trunk is as the server in a single server loss* system with a Poisson arrival process and an exponential service time distribution. This model is commonly denoted by M/M/1-loss.[6] Let x_t denote the state of the server:

* In a loss system, customers who are blocked depart without waiting.

$$x_t = \begin{cases} 1 & \text{if server is busy at epoch } t \\ 0 & \text{if server is idle at epoch } t \end{cases}$$

and $R(\cdot)$ the covariance function of x_t. It is easily shown[7] that for the M/M/1-loss system,

$$R(\tau) = R(0) \exp\{-(\lambda + \mu)\tau\},$$

where λ and μ are the mean arrival and service rates, respectively. Thus, $P_{1,0}(\rho,\tau)$ may be written as:

$$P_{1,0}(\rho,\tau) = (1 - p)\left\{1 - \exp\left(\frac{-\mu\tau}{1-\rho}\right)\right\}, \tag{3}$$

where the trunk occupancy ρ is equal to $\lambda/(\lambda + \mu)$. Throughout this paper we will be concerned with $\tau = 100$ or 200 seconds and a nominal holding time $1/\mu$ in the vicinity of 3 minutes. The mean holding time of a killer trunk $1/\mu^*$ will always be expressed as $1/r\mu$ with r typically in the range 5 to 15. Thus, if we denote $\mu\tau$ by S, we may write the transition probability $P\{x_{t+\tau} = 0/x_t = 1\}$ for a trunk with mean occupancy ρ as

$$P_{1,0}(\rho,r) = (1 - \rho)\left[1 - \exp\left(\frac{-rS}{1-\rho}\right)\right], \tag{4}$$

where $r = 1$ corresponds to a normal trunk. (Since $P_{1,0} = 1 - \rho$ implies that $x_\tau, x_{2\tau}, \cdots$ are independent, we will assume independence for r sufficiently large in subsequent sections.)

Figure 2 is a plot of $P_{1,0}$ vs ρ corresponding to $S = 10/9$ (200-second sampling and a 3-minute mean holding time) for several values of r. $P_{1,0}$ is essentially equal to $1 - \rho$ for $r \geqq 5$. Figure 3 is a similar plot of $P_{1,0}$ vs ρ corresponding to 100-second sampling and a 3-minute mean holding time ($S = 5/9$). In this figure $P_{1,0}$ is essentially equal to $1 - \rho$ for r $\geqq$ 7.5.

Before putting too much emphasis on the transition probabilities in Figs. 2 and 3, it is prudent to consider the effect of factoring more realistic assumptions into the single server loss model. Thus, while the Poisson arrival process assumption is probably a reasonable assumption for a trunk in a 5XB trunk group (random selection of idle servers), it poorly models the overflow nature of the traffic offered to trunks in a 1XB/XBT trunk group.† In the latter case, it is more appropriate to model the input stream to a trunk as a peaked process.[6] Figure 4 is a plot of $P_{1,0}$ vs ρ parameterized by the peakedness (z) of the input stream. This figure is based on an expression for $P_{1,0}$ derived for a GI/M/1-loss model‡ with

† The trunk-selection procedure for 1XB and XBT trunk groups is essentially a two-sided ordered hunt.[2]

‡ GI/M/1-loss denotes a single server loss system with a renewal process input stream (GI) and an exponential (M) service time distribution.

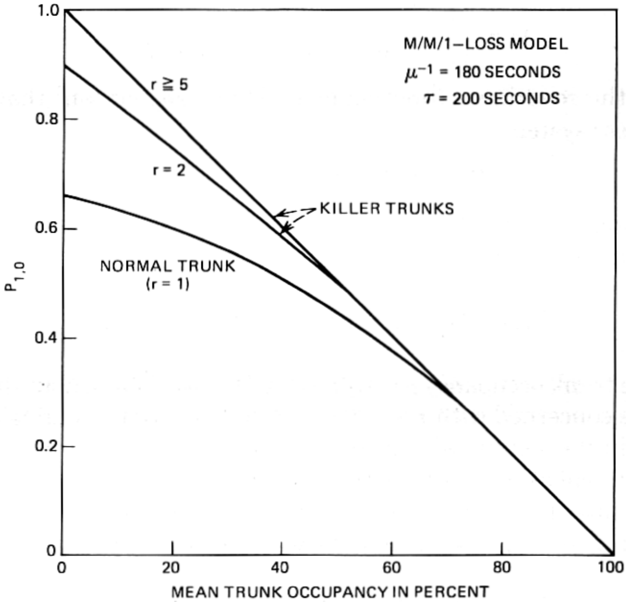


Fig. 2— 1 → 0 transition probability for the 200-second sampling option.

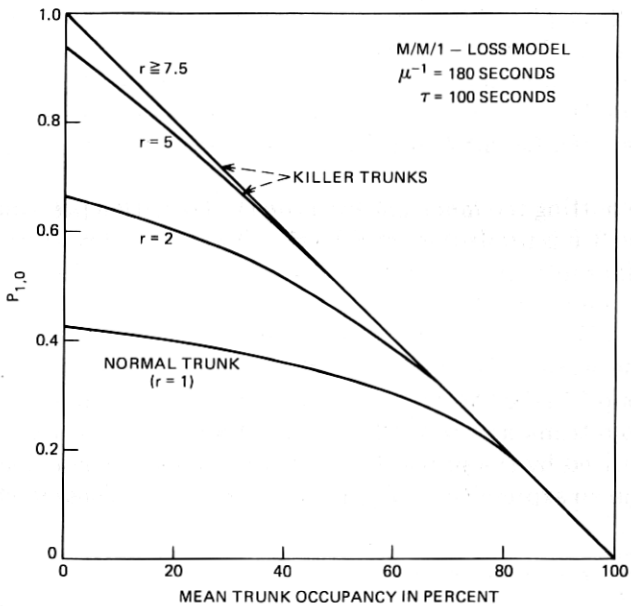


Fig. 3— 1 → 0 transition probability for the 100-second sampling option.

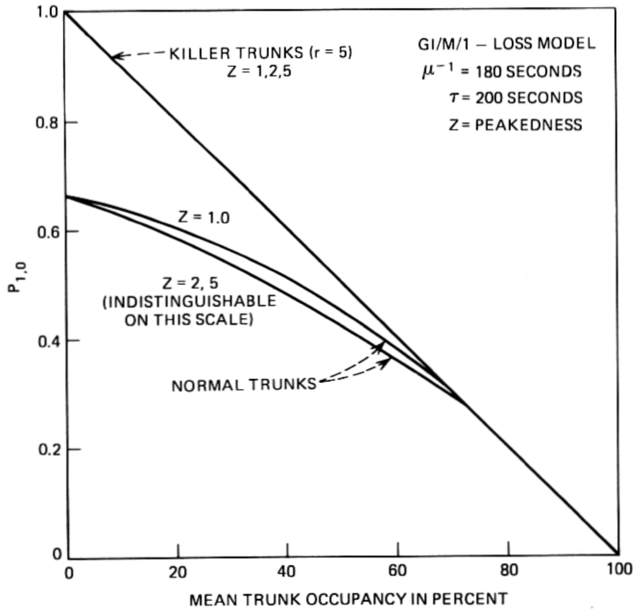


Fig. 4— 1 → 0 transition probability for the GI/M/1-loss model with a switched Poisson arrival process—200-second sampling option.

a switched Poisson input stream (commonly used to model overflow traffic).[8] Appendix A contains several details on the model and derivation. It is clear that the effect of peaked traffic on the transition probability is very small ($z = 1$ corresponds to a Poisson stream).

Recent data[9] indicates that the service time distribution of a normal trunk has a coefficient of variation significantly greater than 1 (the exponential case). Thus, in Appendix A we derive the covariance function of the server process x_t for an M/G/1-loss‡ model with a mixed exponential type of service distribution. Figure 5 is a plot of $P_{1,0}$ vs ρ parameterized by the coefficient of variation of the mixed exponential service distribution. We see that increasing the coefficient of variation has a noticeable effect on the transition probabilities, but the effect is to increase the discrimination between the normal and killer-trunk transition probabilities.

Thus, it would appear that the transition probabilities based on the M/M/1-loss model are reasonably robust to perturbations in the trunk model. In addition, one suspects that using these transition probabilities in a detection scheme, which exploits the basic differences between killer and normal trunk transitions, might lead to a conservative design.

‡ M/G/1-loss denotes a single server loss system with a Poisson (M) input stream and a general (G) service time distribution.

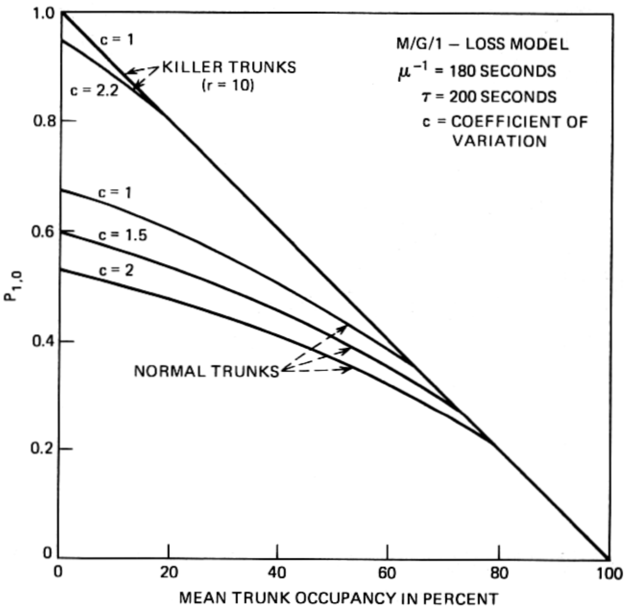


Fig. 5— 1 → 0 transition probability for the M/G/1-loss model with a mixed exponential service time distribution—200-second sampling option.

3.2 Testing statistical hypotheses—a basic idea

Suppose a trunk has constant mean occupancy ρ ($Ex_t = \rho$) and we observe it for h seconds during which n switch counts and t_{10} $1 \rightarrow 0$ state transitions accumulate. We may write

$$n = \sum_{k=1}^{m} x_{k\tau}$$

and

$$t_{10} = \sum_{k=2}^{m} [x_{k\tau} - x_{(k-1)\tau}]^{-},$$

where

$$m = \frac{h}{\tau} \text{ and } z^{-} = \begin{cases} 0 & z \geqq 0 \\ 1 & z < 0 \end{cases}.$$

Hence, we have

$$E(n) = \rho \left(\frac{h}{\tau}\right) \tag{5a}$$

and

$$E(t_{10}) = (m-1)E([x_{k\tau} - x_{(k-1)\tau}]^{-}) = \left(\frac{h}{\tau} - 1\right) \rho P_{1,0}. \tag{5b}$$

Thus, for example, a trunk at 20-percent occupancy (sampled at the 100-second rate) would accumulate 108 switch counts, on the average, in 15 hours. The mean number of $1 \rightarrow 0$ state transitions in this time interval corresponding to a normal trunk is 43, but the corresponding mean number of $1 \rightarrow 0$ state transitions for a killer trunk (during the same time interval) is 86. (Referring to Fig. 3, we see that $P_{1,0}(\rho = 0.2)$ is 0.40 and 0.80 for a normal and killer ($r \geqq 7.5$) trunk, respectively.)

With this example in mind, it is natural to consider a detection scheme of the following type:

(*i*) Wait until we accumulate n_0 switch counts on a trunk.

(*ii*) When $n = n_0$, compare the accumulated $1 \rightarrow 0$ state transitions t_{10} to some threshold T_f.

(*iii*) If $t_{10} > T_f$, decide that the trunk is a killer, otherwise decide that the trunk is normal. [t_{10} is not directly available (see Section 2.1), but it may be estimated by $t/2$.]

If the trunk occupancy were known and fixed, this scheme would appear to be very reasonable. The analogy to the usual scheme suggested for deciding between a fair and a biased coin is clear: n_0 is the number of (hopefully) independent experiments (analogous to the number of coin tosses), with each experiment having just two possible outcomes: the scan which follows the switch count is either 0 or 1. Thus, each switch count is associated with a $1 \rightarrow 0$ state transition ("heads") or a $1 \rightarrow 1$ state transition ("tails").

From the point of view of statistical hypothesis testing, we are thinking of two underlying states:

$$\text{Null hypothesis } H_0\text{: } P_{1,0}(\rho) = P \text{ trunk normal}$$

$$\text{Alternate hypothesis } H_1\text{: } P_{1,0}(\rho) = P^* \text{ trunk killer}$$

Thus, our intuition suggests that a threshold test of the type sketched above is natural for distinguishing between H_0 and H_1. We will see (Section 4.1) that a (nonoptimal) test of this form arises naturally from pursuing the coin tossing analogy further.

3.3 Problem formulation—sharpening the focus

To simplify matters, assume to begin with that

(*i*) The nominal mean holding time $1/\mu$ is known.

(*ii*) The trunk occupancy ρ is known and constant.

(*iii*) The switch-count and transition data accumulations $n(m)$ and $t(m)$ are continuously available (scan by scan).

With these assumptions, it is an easy matter to conceptually describe the "optimum" scheme for deciding between the two simple hypotheses,

$$H_0\text{: Trunk normal (mean holding time } 1/\mu)$$

$$H_1\text{: Trunk killer (mean holding time } 1/r\mu),$$

with it understood that "trunk" refers to one of the specific models described in Section 3.1 (for concreteness assume the M/M/1-loss model).

Let $\mathbf{x}_m = (x_1, \cdots, x_m)$ be the sequence of trunk states up to and including the mth scan, and let the available data be (as before)

$$n(m) = \sum_{i=1}^{m} x_i, \; t(m) = \sum_{i=2}^{m} |x_i - x_{i-1}|.$$

Let

$$P_{im}(n,t) = P(n(m) = n, t(m) = t/H_i) \; i = 0 \text{ or } 1$$

and let

$$\ell_m(n,t) = \frac{P_{1m}(n,t)}{P_{0m}(n,t)}.$$

The joint probability distributions $P_{im}(n,t)$, $i = 0,1$, are well defined for any specific trunk model, but they may be nontrivial to derive. $\ell_m(.,.)$ viewed as a function of the vector random variable $[n(m),t(m)]$ is referred to as the likelihood ratio statistic and plays a central role in the theory of statistical hypothesis testing. More specifically, the optimum test (in a variety of senses) for deciding between two simple hypotheses involves suitably comparing ℓ_m to a threshold (or thresholds) in order to make a decision.

We briefly review two optimum tests, the Neyman-Pearson (fixed sample) test[10] and Wald's sequential probability ratio test (SPRT), using notation appropriate to our (discrete) problem.

3.3.1 The Neyman-Pearson test

Suppose α and β denote the type 1 and 2 errors* of the test,

$$\text{Choose } H_1 \text{ if } \ell_m \geqq T$$

$$\text{Choose } H_0 \text{ if } \ell_m < T,$$

and suppose α' and β' denote the type 1 and 2 errors of any other test (requiring m samples) for deciding between H_0 and H_1. Neyman and

* The type 1 and 2 errors, α and β, are often referred to as the probability of false alarm and the probability of miss, respectively (α = probability of choosing H_1 given H_0 is the true state, β = probability of choosing H_0, given H_1 is the true state.)

Pearson's classical result is: if $\alpha' \leqq \alpha$, then $\beta' \geqq \beta$. Thus, of all tests requiring m samples and having a false-alarm probability not exceeding α, the likelihood ratio test achieves the minimum probability of miss (maximum probability of detection). Since $\alpha = P(\ell_m > T/H_0)$, choosing a sample size m and threshold T to achieve $\alpha \leqq \alpha_0$ requires knowledge of the (conditional) distribution of ℓ_m. Similarly, having chosen m and T, calculating $\beta = P(\ell_m < T/H_1)$ requires the distribution of ℓ_m (conditioned on H_1). Note also that with such a fixed sample test, we decide in advance to accumulate exactly m samples before making a decision. In many contexts, data accumulates sequentially in time, and rigidly requiring m samples—independent of the particular realization that is unfolding—is not an optimal strategy.

3.3.2 Wald's sequential probability ratio test

Using Wald's SPRT,[10,11] we continue to update ℓ_k, $k = 1,2,\cdots$ and defer a decision as long as $\ell_k \epsilon (T_0,T_1)$. We make a decision the *first* time ℓ_k falls outside the interval (T_0,T_1). Thus,

$$\text{if } T_0 < \ell_k < T_1 \quad k = 1,2,\cdots,m-1$$

$$\text{and } \ell_m \notin (T_0,T_1),$$

$$\text{then choose } H_1 \text{ if } \ell_m \geqq T_1$$

$$\text{and choose } H_0 \text{ if } \ell_m \leqq T_0.$$

Clearly the stopping time m of the SPRT is a random variable, and the mean of m (given either hypothesis) is a measure of the time it takes to reach a decision. (Under a wide variety of circumstances, the SPRT terminates with probability 1.) Let $E_i(m)$ $(i = 0 \text{ or } 1)$ denote the mean stopping time, given that hypothesis i is in effect. Given a SPRT with type 1 and 2 errors α and β, and with mean stopping times $E_0(m)$ and $E_1(m)$, consider any other test (*sequential or not*) with type 1 and 2 errors α' and β', and with mean stopping times $E_0'(m)$ and $E_1'(m)$. The SPRT has the following optimal character[10]

$$\text{if } \alpha' \leqq \alpha \text{ and } \beta' \leqq \beta,$$

$$\text{then } E_0'(m) \geqq E_0(m) \text{ and } E_1'(m) \geqq E_1(m).$$

Thus a SPRT is superior to a fixed sample test, if both tests have the same type 1 and 2 errors, in the sense that on the average it reaches a decision more quickly (under either hypotheses).

In sharp distinction to the fixed sample test, the thresholds T_0 and T_1 required to approximately achieve specified type 1 and 2 errors are trivially determined.[11] On the other hand, even determining the mean and variance of the stopping time is often a difficult chore.

In Section 4.2, we explicitly calculate the SPRT* for the simple hypothesis testing problem described at the beginning of this section. Before looking at this optimum test, however, we describe an ad hoc algorithm which is very robust and consequently attractive from a practical point of view.

IV. INDIVIDUAL TRUNK ALGORITHMS

A basic underlying assumption in this section is that the normal mean holding time of a trunk is known. Thus, if the algorithms in this section are designed relative to a normal mean holding time of 3 minutes, they will not discriminate between normal trunks having a mean holding time in the vicinity of 40 seconds,† and an actual killer trunk with the same mean holding time—both of these trunks will be detected as killer trunks.

The rationale for studying this type of detection problem is two-fold: from the practical point of view the simplicity of implementation and general applicability‡ of these algorithms is attractive, and EADAS/ICUR can flag trunk groups which should not be studied by the killer trunk-detection algorithms (thus preventing false alarms on normal short-holding-time trunks). From the theoretical point of view, it was natural to consider this problem before factoring group information into the picture.

Another modeling assumption used in this section (as well as in subsequent ones) is that the arrival process is stationary within data accumulation intervals, but the mean arrival rate may change arbitrarily from one accumulation period to another. Since we use equilibrium analysis (e.g., in calculating $P_{1,0}$) we assume, in effect, that equilibrium is achieved instantaneously.

4.1 An ad hoc algorithm

The essential idea of the test suggested in Section 3.2, is to decide on the state of a trunk by comparing the number of $1 \rightarrow 0$ state transitions (t_{10}) to some threshold T_f, *conditional* on having accumulated a fixed number of switch-counts. We heuristically* proceed to derive such a test, using a standard likelihood ratio formulation, and explicitly take into account the time-variability of traffic.

Let $\mathbf{x}_m = (x_1, \cdots, x_m)$ correspond to the (unobservable) binary se-

* Based on the M/M/1-loss model for a trunk.

† Trunks in special-purpose trunk groups (credit checking, weather, etc.) will typically have mean holding times in the vicinity of 40 seconds.

‡ The individual trunk algorithms can be used to test *any* trunk—regardless of the type of switching machine the trunk is associated with.

* The distributional assumptions made in this section are intuitively motivated, but cannot be rigorously justified. We examine these assumptions carefully in Section 4.3.

quence of trunk states during an accumulation period in which m scans occur. Let $t_{10}(m)$ and $n(m)$ be the number of $1 \rightarrow 0$ state transitions and switch counts associated with $\mathbf{x}_m$. Denote the conditional probability

$$P(t_{10}(m) = t/n(m) = n)$$

for a normal and killer trunk by $P(t/n)$ and $P^*(t/n)$, respectively. These conditional distributions depend, of course, on the trunk's occupancy and on the particular trunk model we have in mind. [$P^*(t/n)$ also depends on the killer parameter r.] However, for the purposes of the heuristic development of this section, we do not precisely define which trunk model we have in mind.

Since each switch count is associated with either a $1 \rightarrow 0$ or a $1 \rightarrow 1$ state transition with probabilities $P_{1,0}$ and $P_{1,1} = 1 - P_{1,0}$, respectively for a normal trunk, and since we expect successive transition events on a trunk to be essentially independent,* it seems reasonable to assume that $P[t_{10}(m) = t/n(m) = n]$ for a normal trunk is binomially distributed with parameters n and $P_{1,0}$. This same argument applies to a killer trunk. Denote the binomial distribution with parameters n and p by $b(k;n,p)$ $k = 0,\cdots,n$, where

$$b(k;n,p) = \binom{n}{k} p^k(1-p)^{n-k}.$$

Thus, we may think of a trunk with occupancy ρ during an accumulation period as having a conditional distribution

$$P[t_{10}(m) = t/n(m) = n] = b[t;n,P_{1,0}(\rho,r)], \tag{6}$$

with $r = 1$ and $r = r_0$ corresponding to the normal and killer states of the trunk. (Recall that $P_{1,0}(\rho,r)$ is essentially independent of r for $r \geqq r_0$ with $r_0 = 7.5$ and 5.0 for 100- and 200-second sampling, respectively.) With these assumptions, we may think of testing the two simple hypotheses:

$$H_0\colon P(t/n) = b(t;n,P_{1,0}) \quad P_{1,0} = P_{1,0}(\rho,1)$$

$$H_1\colon P(t/n) = b(t;n,P^*_{1,0}) \quad P^*_{1,0} = P_{1,0}(\rho,r_0).$$

If the $1 \rightarrow 0$ transition and switch-count accumulations for two successive and contiguous accumulation periods are (t_1,n_1) and (t_2,n_2) respectively, we assume that

$$P(t_1,t_2/n_1,n_2) = P(t_1/n_1)P(t_2/n_2).$$

The idea here is that the only dependence between the two successive

* The idea is that if significant correlation extends only one or two scans back, then successive transition events (events "triggered" by switch counts) should be essentially independent.

bit streams $\mathbf{x}_m^1 = (x_1,\cdots,x_m)$ and $\mathbf{x}_m^2 = (x_{m+1},\cdots,x_{2m})$ is essentially due to the dependence between x_m and x_{m+1}.

Thus, if we denote the transition and switch-court accumulations for the ith accumulation period (in which m_i scans occur) by $[t_{10}(m_i),n(m_i)]$ during which the trunk has occupancy ρ_i, we have

$$P[t_{10}(m_1) = t_1,\cdots,t_{10}(m_k) = t_k | n(m_1) = n_1,\cdots,n(m_k) = n_k] = \prod_{i=1}^{k} b[t_i;n_i,P_{1,0}(\rho_i,r)], \quad (7)$$

where ρ_i $i = 1,\cdots,k$ are the occupancies for the k accumulation periods. If $\boldsymbol{t}_k = (t_1,\cdots,t_k)$ and $\boldsymbol{n}_k = (n_1,\cdots,n_k)$ consider the likelihood ratio:

$$\ell(\boldsymbol{t}_k/\boldsymbol{n}_k) = \prod_{i=1}^{k} \frac{b[t_i;n_i,P_{1,0}(\rho_i,r_0)]}{b[t_i;n_i,P_{1,0}(\rho_i,1)]}. \quad (8)$$

Denote the log likelihood ratio† log $\ell(\boldsymbol{t}_k/\boldsymbol{n}_k)$ by $\hat{\ell}(\boldsymbol{t}_k/\boldsymbol{n}_k)$ and note that

$$\hat{\ell}(\boldsymbol{t}_k/\boldsymbol{n}_k) = \sum_{i=1}^{k} \hat{\ell}(t_i/n_i),$$

where

$$\hat{\ell}(t_i/n_i) = \log \frac{b[t_i;n_i,P_{1,0}(\rho_i,r_0)]}{b[t_i;n_i,P_{1,0}(\rho_i,1)]}.$$

The expression $\hat{\ell}(t_i/n_i)$ can be written as

$$\hat{\ell}(t_i/n_i) = \alpha(\rho_i)t_i - a(\rho_i)n_i$$

with

$$a(\rho) = \log \frac{1 - P_{1,0}(\rho,1)}{1 - P_{1,0}(\rho,r_0)} \quad (9a)$$

and

$$\alpha(\rho) = a(p) + \log \frac{P_{1,0}(\rho,r_0)}{P_{1,0}(\rho,1)}. \quad (9b)$$

Thus, we have

$$\hat{\ell}(\boldsymbol{t}_k/\boldsymbol{n}_k) = \sum_{i=1}^{k} [\alpha(\rho_i)t_i - a(\rho_i)n_i]. \quad (9c)$$

Unfortunately, the occupancy in the ith accumulation period (ρ_i) is unknown and hence equation (9c) cannot be used as a test statistic. One obvious "fix" is to estimate ρ_i by $\hat{\rho}_i = n_i/m_i$, where n_i and m_i are the

† $l > T$ iff $g(\ell) > g(T)$ if g is monotone increasing, so the tests $\ell > T$ and $g(\ell) > g(T)$ are equivalent.

switch count and the number of scans, respectively, during the ith accumulation period. In stationary traffic, the estimate

$$\tilde{\rho}_i = \sum_{j=1}^{i} n_j \Big/ \sum_{j=1}^{i} m_j$$

would be used [$\sigma(\tilde{\rho}_i) \doteq 1/\sqrt{i}\,\sigma(\hat{\rho}_i)$ if $m_j = m$ for all j].

Corresponding to the sequence of accumulations, (t_1,n_i,m_i) $i = 1,2,\cdots$ we define r_i and R_i, $i = 1,2,\cdots$ by

$$r_i = \alpha(\hat{\rho}_i)t_i - a(\hat{\rho}_i)n_i,\ \hat{\rho}_i = n_i/m_i \tag{10a}$$

and

$$R_i = R_{i-1} + r_i \text{ with } R_0 = 0. \tag{10b}$$

Thus, we arrive at the sequential test:

(i) Compute R_i, $i = 1,2,\cdots$ and defer making a decision as long as $T_0 < R_i < T_1$.

(ii) If $i = k$ corresponds to the first accumulation period for which $R_i \notin (T_0,T_1)$, then

$$R_k \leqq T_0 \Rightarrow \text{Trunk normal}$$

$$R_k \geqq T_1 \Rightarrow \text{Trunk killer.}$$

If we ignore the fact that we are estimating ρ_i by $\hat{\rho}_i$, and by assuming that the various assumptions made are valid (see Section 4.3), we identify the above test as Wald's SPRT and as such T_0 and T_1 can be calculated as follows:[11] to approximately achieve type 1 and 2 errors, α and β, respectively, $\alpha + \beta < 1$, choose

$$T_0 = \log\left(\frac{\beta}{1-\alpha}\right) \tag{10c}$$

and

$$T_1 = \log\left(\frac{1-\beta}{\alpha}\right). \tag{10d}$$

Throughout this section, we have assumed that the $1 \rightarrow 0$ transitions (t_{10}) are available when, in fact, only the total transitions (t) are available. It should be clear that t_{10} can differ from $t/2$ by at most $\pm\frac{1}{2}$. To be precise, let $t_{10}(m)$, $t_{01}(m)$ be the number of $1 \rightarrow 0$ and $0 \rightarrow 1$ state transitions corresponding to a bit stream $\mathbf{x}_m = (x_1, \cdots, x_m)$. If $n(m)$ is the switch count corresponding to $\mathbf{x}_m$, then we have

$$n(m) = t_{10}(m) + t_{11}(m) + x_m \tag{11a}$$

and

$$n(m) = t_{01}(m) + t_{11}(m) + x_1, \tag{11b}$$

where $t_{11}(m)$ is the number of $1 \rightarrow 1$ state transitions. Therefore

$$t_{10}(m) + x_m = t_{01}(m) + x_1,$$

which together with $t(m) = t_{01}(m) + t_{10}(m)$ yields

$$t_{10}(m) = \frac{1}{2} t(m) + \left(\frac{x_1 - x_m}{2}\right) \tag{12a}$$

and

$$t_{01}(m) = \frac{1}{2} t(m) - \left(\frac{x_1 - x_m}{2}\right). \tag{12b}$$

Thus, we can write the statistic update (eq. 10a) as

$$\alpha \frac{t}{2} - \text{an} + \left(\frac{x_1 - x_m}{2}\right) \alpha.$$

[It is easy to show that $E[(x_1 - x_m)\alpha(\hat{\rho})] = 0$.]

We conclude this section with an interpretation of the statistic update. Rewriting the statistic update as

$$r = (\alpha - a)t_{10} - a(n - t_{10})$$

and using eq. (11a), we obtain

$$r = (\alpha - a)t_{10} - at_{11} - ax_m. \tag{13}$$

Now, from eqs. (9a) and (9b), it is clear that $\alpha > a > 0$. Thus, each $1 \rightarrow 0$ transition is weighted positively (evidence of a killer) while each $1 \rightarrow 1$ transition is weighted negatively (evidence of a normal trunk). This is an intuitive explanation of the fact that the random walk (eq. (10b))

$$R_k = R_{k-1} + r_k$$

has a positive drift if the trunk is a killer and a negative drift if the trunk is normal.

The fact that the update assigns a negative weight $(-a)$ whenever the last bit (x_m) is 1 uncovers a modeling deficiency. Recall that in eq. (6) we assumed

$$P(t_{10}(m) = t/n(m) = n) = b(t;n,P_{10}),$$

even though $x_m = 1$ can not contribute to an observable $1 \rightarrow 0$ transition. In this way we effectively modeled in a bias towards making "trunk normal" decisions. We can easily correct eq. (6) by conditioning on whether $x_m = 0$ or 1, obtaining:

$$P(t_{10}(m) = t/n(m) = n) = (1 - \rho)b(t;n,P_{10}) + \rho b(t;n - 1,P_{10}).$$

Now, proceeding as before in formulating the log likelihood ratio yields a statistic $\tilde{R}_k$, where

$$\tilde{R}_k = \tilde{R}_{k-1} + \tilde{r}_k$$

and

$$\tilde{r}_k = r_k + q_k,$$

where r_k is defined by eq. (10a),

$$q_k = \log \frac{1 + \left(\frac{\hat{\rho}_k}{1-\hat{\rho}_k}\right)\left(1 - \frac{t_k}{n_k}\right)\left(\frac{1}{1 - P_{1,0}(\hat{\rho}_k, r_0)}\right)}{1 + \left(\frac{\hat{\rho}_k}{1-\hat{\rho}_k}\right)\left(1 - \frac{t_k}{n_k}\right)\left(\frac{1}{1 - P_{1,0}(\hat{\rho}_k, 1)}\right)} \tag{14}$$

and $\hat{\rho}_k$, t_k, and n_k are the trunks occupancy estimate, $1 \rightarrow 0$ state transitions, and switch count, respectively, during the kth accumulation period.

Thus, we obtain our original test statistic with the correction term q_k added on. Note that $q_k \geqq 0$, $q_k \rightarrow 0$ as $\hat{\rho}_k \rightarrow 0$, and $q_k \rightarrow$ a as $\hat{\rho}_k \rightarrow 1$, which is just the type of behavior expected, to offset the bias term in r_k.

Having heuristically developed an ad hoc sequential algorithm that is intuitively appealing and easily implementable, it is natural to ask: how does it compare to the *optimum* sequential algorithm? In the following section, we rigorously develop an optimum sequential test.

4.2 An optimal algorithm

Consider the two simple hypotheses:

H_0: Trunk normal (mean occupancy ρ, mean holding time $1/\mu$)

H_1: Trunk killer (mean occupancy ρ, mean holding time $1/r_0\mu$).

The optimum test for deciding between the two hypotheses—in the sense of minimizing the mean decision time—for given type 1 and 2 errors, is Wald's SPRT (see Section 3.3), and it is based on the likelihood ratio statistic $\ell_m(t,n)$ given by

$$\ell_m(t,n) = \frac{P^*(t(m) = t, n(m) = n)}{P(t(m) = t, n(m) = n)}. \tag{15}$$

Thus, it is clear that the ad hoc test described in Section 4.1 is not optimal, based as it is on an assumed conditional distribution,

$$P(t_{10}(m) = t/n(m) = n).$$

Before proceeding to study eq. (15), we must define the trunk model precisely. In the developments that follow, we model a trunk as the server in an M/M/1-loss system (see Section 3.1). The model implies that the sequence of trunk states x_t, $t = k\tau$, $k = 1,2,\cdots$ is Markovian. Note that although this appears to be a reasonable model for a normal trunk with 200-second sampling, it ignores the conditional dependence "2 samples back," which is more important for 100-second sampling—e.g., x_t given $x_{t-\tau}$ is independent of $x_{t-2\tau}$ for the M/M/1-loss model. Taking this dependence into account in a trunk model would not be useful however, since the data needed to implement dependence "two scans back," is not available.

Since we are modeling the sequence of trunk states as a binary valued Markov process $x_{k\tau}$, k = 1,2,···, in equilibrium, it is clear that this process is characterized by $\boldsymbol{\theta} = (P_{1,0},P_{0,1})$, where $P_{1,0}$ and $P_{0,1}$ are the transition probabilities

$$P(x_{t+\tau} = 0/x_t = 1) \text{ and } P(x_{t+\tau} = 1/x_t = 0),$$

respectively. (In general, a binary valued Markov process $x_{k\tau}$, $k = 1,2,\cdots$ in equilibrium, can be characterized by any two of the three quantities ρ, $P_{1,0}$,$P_{0,1}$. For our special Markov process (based on the M/M/1-loss model), both $P_{1,0}$ and $P_{0,1}$ and hence the process itself is determined by ρ alone.) Now having observed any m-tuple of the samples, which we denote by $\mathbf{x}_m = (x_1,\cdots,x_m)$, it is trivial to show that the statistic

$$\boldsymbol{T}(\mathbf{x}_m) = (t(m),n(m),x_1,x_m)$$

is a sufficient statistic for $\boldsymbol{\theta}$. Thus, except for the initial and terminal states (x_1 and x_m), the transition and switch-count accumulations summarize all the "relevant information" in $\mathbf{x}_m$.

Our hypothesis-testing problem can now be formulated as follows: $x_1,x_2,\cdots$ is a binary-valued Markov chain in equilibrium with parameter $\boldsymbol{\theta} = (P_{0,1},P_{1,0})$ or $\boldsymbol{\theta}^* = (P^*_{0,1},P^*_{1,0})$. That is, our two states are

$$H_0\text{: } \{x_i\} \text{ Markovian, characterized by } \boldsymbol{\theta} = (P_{0,1},P_{1,0})$$

$$H_1\text{: } \{x_i\} \text{ Markovian, characterized by } \boldsymbol{\theta}^* = (P^*_{0,1},P^*_{1,0}).$$

Now, because $(t(m),n(m),x_1,x_m)$ is a sufficient statistic for $\boldsymbol{\theta}$, we know that the likelihood-test statistic based on the raw (unobservable) data $\mathbf{x}_m = (x_1,\cdots,x_m)$ will be expressible in terms of $t(m)$, $n(m)$, x_1 and x_m only. Thus, instead of studying eq. (15), we proceed (for simplicity) to study the likelihood-ratio statistic:

$$\hat{\ell}(\mathbf{x}_m) = \log \frac{P^*(\mathbf{x}_m)}{P(\mathbf{x}_m)}. \tag{16}$$

In Appendix B we study $\hat{\ell}_m(t,n)$ and find that it differs from $\hat{\ell}(\mathbf{x}_m)$ only in an end-effect term. In $\hat{\ell}(\mathbf{x}_m)$ this term depends on x_1 and x_m, whereas in $\hat{\ell}_m(t,n)$ the corresponding term is a function of t and n.

Since

$$P(\mathbf{x}_m) = P(x_1) \prod_{i=2}^{m} P(x_i/x_{i-1}),$$

we may write

$$P(\mathbf{x}_m) = P(x_1)\, P_{10}^{t_{10}}\, P_{11}^{t_{11}}\, P_{01}^{t_{01}}\, P_{00}^{t_{00}}$$

so

$$\hat{\ell}(\mathbf{x}_m) = \log \frac{P*(x_1)}{P(x_1)} + t_{10} \log \frac{P_{10}^*}{P_{10}} + t_{11} \log \left(\frac{1 - P_{10}^*}{1 - P_{10}}\right) + t_{01} \log \frac{P_{01}^*}{P_{01}} + t_{00} \log \left(\frac{1 - P_{01}^*}{1 - P_{01}}\right). \quad (17)$$

Note that a trunk with mean occupancy ρ is busy and idle with probability ρ and $1 - \rho$ respectively, independently of the state it is in (normal or killer). Thus,

$$\log \frac{P*(x_1)}{P(x_1)} = 0 \text{ and eq. (17) can be written}$$

$$\hat{\ell}(\mathbf{x}_m) = [(\alpha - a)t_{10} - at_{11}] + [(\beta - b)t_{01} - bt_{00}], \quad (18)$$

where the parameters b and β are defined by

$$b = \log \left(\frac{1 - P_{01}}{1 - P_{01}^*}\right) \quad (19a)$$

$$\beta = b + \log \left(\frac{P_{01}^*}{P_{01}}\right), \quad (19b)$$

and the parameters a and α are defined as in Section 4.1 (eqs. (9a) and (9b)). $P_{0,1}^*$ and $P_{1,0}^*$ correspond to $P_{0,1}(\rho,r)$ and $P_{1,0}(\rho,r)$ with $r = r_0$.

Before discussing the symmetric structure of the optimum statistic [eq. (18)], we examine the $P_{0,1}$ characteristics for the M/M/1-loss model. Using eqs. (2) and (3), we can obtain $P_{0,1}$ vs mean-trunk occupancy ρ for a normal ($r = 1$) and killer ($r = r_0$) trunk. Figures 6 and 7 are plots for the 200- and 100-second sampling option, respectively, with a mean holding time of 180 seconds. It is clear from Fig. 6 that a $0 \rightarrow 1$ transition is just marginally more likely to occur on a killer trunk than on a normal trunk with a 200-second sampling rate. Although, the difference in the $0 \rightarrow 1$ transition probabilities between a normal and killer trunk increases substantially with the 100-second sampling rate, it is clear that these differences are still quite small—compared to the spread between the $P_{1,0}$ and $P_{1,0}^*$ plots (see Figs. 2 and 3). Note that eqs. (2) and (4) show that

$$\frac{P_{0,1}^*}{P_{0,1}} = \frac{1 - \exp \dfrac{-r_0 s}{1 - \rho}}{1 - \exp \dfrac{-s}{1 - \rho}} > 1 \text{ for } r_0 > 1 \quad (19c)$$

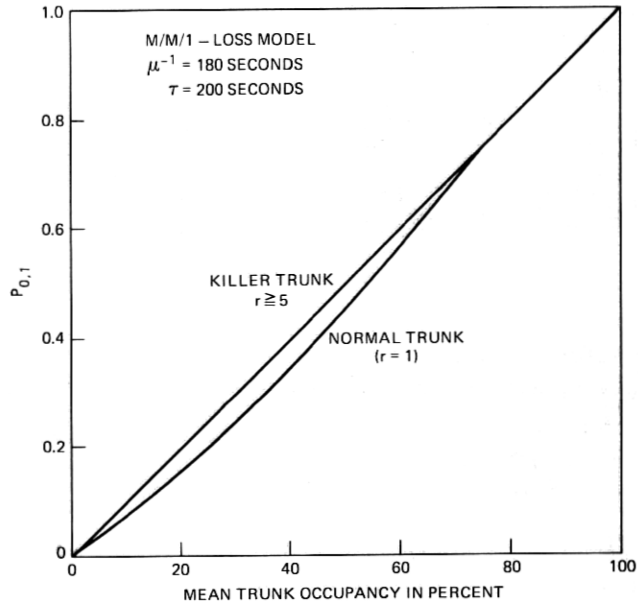


Fig. 6— 0 → 1 transition probability for the 200-second sampling option.

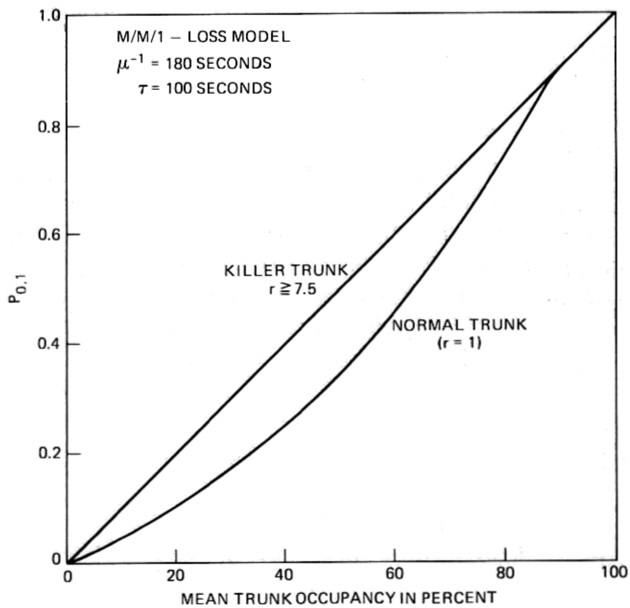


Fig. 7— 0 → 1 transition probability for the 100-second sampling option.

and, hence, we have $\beta > b > 0$. Using eqs. (4) and (19c), we see that $P^*_{0,1}/P_{0,1} = P^*_{1,0}/P_{1,0}$ and, therefore,

$$\beta - b = \alpha - a \tag{19d}$$

Equation (18) shows that the optimum statistic is the sum of two *symmetric* statistics:

(*i*) The statistic $[(\alpha - a)t_{10} - at_{11}]$, which is essentially the *ad hoc* statistic (see eq. (13) and related discussion).

(*ii*) An *additional* statistic $[(\beta - b)t_{01} - bt_{00}]$, which weights $0 \rightarrow 1$ transitions positively (evidence of a killer) and $0 \rightarrow 0$ transitions negatively (evidence of a normal trunk).

Note that by interchanging the role of 0 and 1 in either of these two statistics, we obtain the other—b is obtained from a and β is obtained from α by replacing $P_{1,0}$ with $P_{0,1}$.

By using eq. (11a) and the analogous equation

$$m - n(m) = t_{00}(m) + t_{01}(m) + x^c_m, \; (x^c_m = 1 - x_m) \tag{20}$$

in eq. (18), the optimum statistic can be written

$$\hat{\ell}(\mathbf{x}_m) = [\alpha t_{10}(m) - an(m)] + [\beta t_{01}(m) - b(m - n)] + e_1(x_m), \tag{21}$$

where the end-effect term $e_1(x_m)$ is given by

$$e_1(x_m) = ax_m + bx^c_m.$$

To implement $\hat{\ell}(\mathbf{x}_m)$ with only $t(m)$ and $n(m)$ available, necessitates estimating both $t_{10}(m)$ and $t_{01}(m)$ by $t(m)/2$. That is, using eqs. (12a) and (12b) in eq. (21) yields.

$$\hat{\ell}(\mathbf{x}_m) = \left[\alpha \frac{t(m)}{2} - an(m)\right] + \left\{\beta \frac{t(m)}{2} - b[m - n(m)]\right\} + e(x_1, x_m), \tag{22a}$$

where

$$\begin{aligned} e(x_1, x_m) &= (\alpha - \beta)\left(\frac{x_1 - x_m}{2}\right) + e_1(x_m) \\ &= \left(\frac{a - b}{2}\right)(x_1 + x_m) + b^{\dagger} \end{aligned} \tag{22b}$$

or

$$\hat{\ell}(\mathbf{x}_m) = \left(\frac{\alpha + \beta}{2}\right) t(m) - (a - b)n(m) - bm + e(x_1, x_m). \tag{23}$$

† Recall that $\alpha - \beta = a - b$ [eq. (19d)].

In the development of the ad hoc algorithm, we assumed that the statistics corresponding to successive accumulation periods are *independent*. We conclude this section by examining the independence assumption and with some remarks on implementation.

Thus, turning to multiple accumulation periods, suppose $\mathbf{x}_m^i$, $i = 1,2,\cdots,k$ are the (unobservable) bit streams for k successive (and contiguous) accumulation periods, where $\mathbf{x}_m^i = (x_{(i-1)m+1},\cdots,x_{im})$. Assuming stationary traffic, and noting that $\{x_i\}_{i=1}^{km}$ is Markovian, we can write

$$P(\mathbf{x}_m^1,\cdots,\mathbf{x}_m^k) = \prod_{i=1}^{k} P(\mathbf{x}_m^i) \times \prod_{i=1}^{k-1} \frac{P(x_{im+1}/x_{im})}{P(x_{im+1})} \tag{24}$$

and, therefore,

$$\hat{\ell}(\mathbf{x}_m^1,\cdots,\mathbf{x}_m^k) = \sum_{i=1}^{k} \hat{\ell}(\mathbf{x}_m^i) + \sum_{i=1}^{k-1} \log \left\{\frac{P^*(x_{im+1}/x_{im})/P^*(x_{im+1})}{P(x_{im+1}/x_{im})/P(x_{im+1})}\right\}, \tag{25}$$

where $P(\cdot)$ and $P^*(\cdot)$ denote the distribution under H_0 (trunk normal) and H_1 (trunk killer), respectively. But, as we have seen,

$$\frac{P^*(x_{im+1}/x_{im})}{P^*(x_{im+1})} = 1 - \exp\left\{\frac{-r_0 s}{1-\rho}\right\} \tag{26}$$

and hence x_{im+1} and x_{im} are essentially independent for r_0 sufficiently large. Therefore,

$$P^*(x_{im+1}/x_{im}) = P^*(x_{im+1})$$

and, hence, eq. (25) may be written as

$$\hat{\ell}(\mathbf{x}_m,\cdots,\mathbf{x}_m^k) = \sum_{i=1}^{k} \hat{\ell}(\mathbf{x}_m^i) - \sum_{i=1}^{k-1} I(x_{im};x_{im+1}),$$

where

$$I(x_{im};x_{im+1}) = \log \frac{P(x_{im+1}/x_{im})}{P(x_{im+1})} \tag{27}$$

is recognized as the mutual information random variable, which plays a central role in information theory.[12] It is well known[12] that (under H_0) $E\{I(x_{im};x_{im+1})\}$ is non-negative, and hence to ignore the end-effect term

$$\sum_{i=1}^{k-1} I(x_{im};x_{im+1})$$

by implementing the statistic

$$\sum_{i=1}^{k} \hat{\ell}(\mathbf{x}_m^i)$$

would tend to make a normal trunk look more like a killer trunk on the average. In Appendix C, however, we show that the mean end effect $E\{I(x_{im};x_{im+1})\}$ is negligible compared to the mean statistic update $E\{\hat{\ell}(\mathbf{x}^i_m\}$.

The "optimal" sequential algorithm is implemented in the same manner as the ad hoc algorithm (see eqs. (10a) and (10b)) except that now, corresponding to the sequence of accumulations $(t_i,n_i,m_i)i = 1,2,\cdots$, we define r_i by

$$r_i = \left(\frac{\alpha + \beta}{2}\right) t_i - (a - b)n_i - bm_i.$$

The term t_i denotes the total number of transitions in the ith accumulation interval. As is the case for the ad hoc algorithm (which corresponds to $b = \beta = 0$), the weights are functions of the trunk occupancy estimate $\hat{\rho}_i = n_i/m_i$. In a practical nonstationary environment, no claims of optimality are made or implied. The term "optimal" is applicable only in the context of the equilibrium (e.g., stationary) model with known trunk occupancy.

4.3 The ad hoc algorithm reviewed

The assumption that $t_{10}(m)$ conditioned on the switch count $n(m)$ is binomially distributed, is the basic assumption in the development of the ad hoc statistic. Although this assumption is incorrect (as we will soon see), the ad hoc statistic is essentially (except for an end-effect term) one of two symmetric statistics whose sum is the optimum statistic. Our purpose in this section is to examine the binomial assumption and to explain the relationship found between the ad hoc and optimal statistics.

Since the optimal statistic was developed for a trunk modeled as a server in an M/M/1-loss system, it is natural to examine the binomial assumption (eq. (6)):

$$P[t_{10}(m) = t/n(m) = n] = b[t;n,P_{1,0}(\rho,r)], \tag{28}$$

where $P_{1,0}(\rho,r)$ is given by eq. (4) in this context. Consider a killer trunk with r sufficiently large and suppose $\mathbf{x}_m = (x_1,\cdots,x_m)$ is the bit stream for a killer trunk during some accumulation period. Then, for all practical purposes [see eq. (26)], the trunk states x_i $i = 1,2,\cdots,m$ are independent and identically distributed Bernoulli random variables:

$$P(x_i = x) = \begin{cases} \rho & \text{if } x = 1 \\ 1 - \rho & \text{if } x = 0. \end{cases}$$

Thus, it is clear that the switch-count distribution on a killer trunk is the binomial:

$$P(n(m) = n) = b(n;m,\rho). \tag{29}$$

Now suppose $A_m(t,n)$ denotes the number of binary m-tuples having exactly t $1 \rightarrow 0$ transitions and n ones. If $\mathbf{x}_m = (x_1,\cdots,x_m)$ is a sequence of trunk states for a killer trunk with $n(m) = n$, it is clear that each such sequence has probability

$$P(\mathbf{x}_m) = \rho^n(1-\rho)^{m-n}$$

and therefore

$$P(t_{10}(m) = t, n(m) = n) = A_m(t,n)\rho^n(1-\rho)^{m-n} \tag{30}$$

for a killer trunk (with r sufficiently large). Equations (29) and (30) show that for a killer trunk,

$$P(t_{10}(m) = t/n(m) = n) = \frac{\binom{m-n}{t}\binom{n}{t}}{\binom{m}{n}}, \tag{31}$$

where we have used the fact that

$$A_m(t,n) = \binom{m-n}{t}\binom{n}{t}. \tag{32}$$

It is interesting to note that while our assumed distribution for a killer trunk (28) differs from the correct distribution (31)—note that (31) is independent of ρ—there are some interesting similarities. For example, the assumed distribution peaks in the vicinity of $(n+1)(1-\rho)$ and has mean equal to $n(1-\rho)$† whereas the true distribution peaks in the vicinity of $(n+1)[1-(n/m)]$ and has mean equal to $n[1-n/m]$. Note that for "typical" realizations $(x_1,\cdots,x_m)$, we have

$$\frac{n}{m} \doteq \rho$$

and, hence, the two distributions have the same general location and scale. [In fact, expression (31) is a hypergeometric distribution, which converges to (28) as $m \rightarrow \infty$ if $n = \rho m$ (Ref. 13).] Thus, although incorrect, the binomial distribution approximates the true distribution of the killer trunk.

The following result helps to put the relationship between the ad hoc and the optimal statistics in perspective.

Lemma 2: If $\{x_i\}$ is a binary state stationary Markov chain with transition probabilities $P_{0,1}$ and $P_{1,0}$, and if $\mathbf{x}_m = (x_1,\cdots,x_m)$, then we

† $P_{1,0}(\rho,r) \rightarrow 1-\rho$ as $r \rightarrow \infty$.

have

$$P(\mathbf{x}_m) = b(t_{10};n,P_{1,0}) \times b(t_{0,1};m-n,P_{0,1}) \times q, \tag{33a}$$

where

$$q = \frac{P(x_1)(1-P_{1,0})^{-x_m}(1-P_{0,1})^{-x_m^c}}{\binom{n}{t_{10}}\binom{m-n}{t_{01}}}. \tag{33b}$$

Proof:

$$\begin{aligned} P(\mathbf{x}_m) &= P(x_1)\prod_{i=2}^{m} P(x_i/x_{i-1}) \\ &= P(x_1)\, P_{1,0}^{t_{10}}\, P_{1,1}^{t_{11}}\, P_{0,1}^{t_{01}}\, P_{0,0}^{t_{00}} \\ &= P_{1,0}^{t_{10}}\,(1-P_{1,0})^{t_{11}}\, P_{0,1}^{t_{01}}\,(1-P_{0,1})^{t_{00}} \times P(x_1). \end{aligned}$$

Using eqs. (11a) and (18) to express t_{11} and t_{00} in terms of t_{10} and t_{01}, respectively, yields the result.

Thus, given a binary state stationary Markov chain $\{x_i\}$, it is clear from the above lemma that the log likelihood ratio

$$\hat{\ell}(\mathbf{x}_m) = \log \frac{P*(\mathbf{x}_m)}{P(\mathbf{x}_m)}$$

formulated for the two hypotheses

H_0: $\{x_i\}$ Markovian, characterized by $(P_{0,1},P_{1,0})$

H_1: $\{x_i\}$ Markovian, characterized by $(P^*_{0,1},P^*_{1,0})$

is the sum of three terms:

$$\hat{\ell}(\mathbf{x}_m) = \log\frac{b(t_{10;n,P^*_{1,0}})}{b(t_{10};n,P_{1,0})} + \log\frac{b(t_{01};m-n,P^*_{0,1})}{b(t_{01};m-n,P_{0,1})} + \log\left(\frac{q^*}{q}\right).$$

The first term is the ad hoc statistic $(\alpha t_{10} - an)$, the second term is the additional statistic $[\beta t_{01} - b(m-n)]$, and the third term is an end-effect term $(ax_m + bx^c_m)$.

$$\log\left(\frac{q^*}{q}\right) = ax_m + b\,x^c_m + \log\frac{P*(x_1)}{P(x_1)} = ax_m + bx^c_m,$$

since, $P*(x_1) = P(x_1) = \rho$.

The ad hoc algorithm, although based on the approximate binomial distribution, is very attractive for a number of practical reasons:

(*i*) For the 200-second sampling option, it is essentially optimum in a *practical* sense, since the $P_{0,1}$ characteristics for a normal and killer trunk are not far enough apart to exploit (see Fig. 6).

(*ii*) When we exploit the grouping information for the 5XB trunk group in Section V, it will become obvious that the additional part of the optimum statistic is quite sensitive to the trunk occupancy and, hence, to the trunk-selection procedure modeled. (The sensitivity of the "additional" statistic to trunk occupancy stems from the fact that $P_{0,1}$ is "almost" proportional to ρ.) On the other hand, we will see that the ad hoc 5XB algorithm is relatively robust to minor perturbations in the trunk occupancy (and therefore to the trunk-selection procedure) and hence might be expected to perform well in a real 5XB environment.

V. THE 5XB TRUNK-GROUP ALGORITHMS

In addition to utilizing individual trunk switch-count and transition accumulations the 5XB group algorithms exploit the following:

(*i*) The identity of all trunks common to a group.
(*ii*) The trunk-selection procedure.

The resulting 5XB group algorithms typically are faster* than their individual trunk counterparts and are also less sensitive to the groups nominal holding time.

5.1 The 5XB trunk-group model

For the purposes of this paper, we model a 5XB trunk group (with all trunks normal) as an M/M/N-loss model with *random selection of idle trunks.*[2] The same assumptions apply if the group contains one or more killer trunks, but in this case we assume that killer trunks have a mean holding time equal to $1/r$ that of the normal mean holding time. In addition to being convenient theoretically, this idealized model has also been very useful in developing the 5XB group algorithms presently implemented in ICAN.

If all N trunks are normal, the random selection rule implies that all trunks have the same mean occupancy. In Ref. 2, the birth and death equations for the above model with a single killer trunk were solved in closed form, and in Ref. 14 this was generalized to an arbitrary number of killer trunks. These analytic results turn out to be quite useful, and in what follows we will need the following results derived in Ref. 2.

Theorem 1: For the above 5XB trunk group model having a single killer trunk with parameter r and an offered load of a erlangs, the blocking probability $\hat{B}(N,a,r)$ and the mean occupancy $\rho_r^*(N,a,r)$ of the killer trunk are given by:

* For given type 1 and 2 errors, the group algorithms typically have a considerably smaller mean decision time than their individual trunk algorithm counterparts.

$$(i)\ \hat{B}(N,a,r) = \frac{NB(N,a)}{Nr - (r-1)a[1 - B(N,a)]} \tag{34a}$$

$$(ii)\ \rho_r^*(N,a,r) = \frac{1}{1 + \frac{rN}{a} - r[1 - B(N-1,a)]}, \tag{34b}$$

where $B(N,a)$ is the usual erlang B blocking associated with an M/M/N-loss system with all trunks normal and an offered load of a erlangs.

It is easy to see that the occupancy ρ_r of each of the $N - 1$ normal trunks must satisfy the conservation equation:

$$r\rho_r^* + (N-1)\rho_r = a[1 - \hat{B}], \tag{34c}$$

and the trunk-group occupancy ϕ_r^* is defined by:

$$\phi_r^* = \frac{\rho_r^* + (N-1)\rho_r}{N}. \tag{34d}$$

(For a 5XB trunk group having a killer trunk with parameter r: ρ_r^* and ρ_r denote the mean occupancy for a killer and normal trunk and ϕ_r^* denotes the mean group occupancy.)

Although eqs. (34a) through (34d) define an implicit relationship between $\rho_r^*(N,a,r)$ and $\phi_r^*(N,a,r)$, it will be very useful to have a simple explicit relationship. If the blocking term in eq. (34b) is ignored and if we "associate" ϕ_r^* with a/N, an approximation suggested is:

$$\rho_r^* \doteq \frac{\phi_r^*}{r - (r-1)\phi_r^*}. \tag{35a}$$

This approximation, although quite good for large N, is rendered obsolete by the following exact result:

Theorem 2: Consider a 5XB trunk-group model with all trunks normal and mean-group occupancy ϕ. (We will let ϕ denote the (mean) group occupancy for a 5XB trunk group with all trunks normal.) *If one of the trunks is replaced by a killer with parameter r, then*

$$\rho_r^* = \rho(\phi,r),$$

where,

$$\rho(\phi,r) = \frac{\phi}{r - (r-1)\phi}. \tag{35b}$$

Of course

$$\phi = \frac{a[1 - B(N,a)]}{N}$$

where $B(\cdot\,,\cdot)$ is the usual erlang B blocking expression.

This surprising result, which follows easily from eq. (34b), is proved in Appendix D. As a consequence of this theorem, "5XB group occupancy" will be used to denote the occupancy of a 5XB group with all trunks normal.

The mean occupancy of the normal trunks in a 5XB trunk group model having a single killer trunk no longer is given exactly by ϕ. But the following result, derived in Appendix D, shows that ϕ is a good approximation.

Theorem 3: Consider a 5XB trunk group model with N trunks having a single killer trunk with parameter $r \geqq 1$. The mean occupancy (ρ_r) of the $N - 1$ normal trunks satisfy

$$\phi \geqq \rho_r \geqq \left\{ \frac{r - \left(\frac{N}{N-1}\right)(r-1)\phi}{r - (r-1)\phi} \right\} \times \phi, \tag{36}$$

where ϕ is the mean-group occupancy with all trunks normal.

Theorems 2 and 3 are proved in Appendix D, where an exact expression for ρ_r is also derived. These results are special cases of general results obtained for the random selection model.[14]

5.2 Exploiting the 5XB Grouping Information

To simplify matters, we assume that a trunk in a 5XB group* with mean-group occupancy ϕ has mean occupancy $\rho(\phi,r)$ given by

$$\rho(\phi,r) = \frac{\phi}{r - (r-1)\phi}, \tag{37a}$$

where $r = 1$ corresponds to a normal trunk. Thus, if a group has no killer trunks, all normal trunks satisfy $\rho = \phi$ and eq. (37a) with $r = 1$ yields the correct occupancy. If, however, the group has a killer trunk, then all normal trunks satisfy inequality (36) so eq. (37a) with $r = 1$ is an approximation that increases in accuracy with the size of the group. Of course eq. (37a) is exact for a (single) killer trunk in a 5XB trunk group.

It is clear from eq. (37a) that

$$\rho(\phi,r) = \frac{\rho(\phi,1)}{r - (r-1)\rho(\phi,1)} \tag{37b}$$

* We use "5XB group" and our idealized model of a 5XB trunk group interchangeably.

and hence $\rho(\phi,r)$ is typically much smaller than $\rho(\phi,1)$. (For $\phi = 0.50$ and $r = 10$, $\rho(\phi,r) = 2/11\ \rho(\phi,1)$. Thus, it would appear that considering the $1 \rightarrow 0$ transition probability as a function of the 5XB group occupancy would effectively "spread" the $P_{1,0}$ characteristics in Figs. 2 and 3 further apart. That is, for a given ϕ, we propose comparing

$$P_{1,0}[\rho(\phi,1)] \text{ and } P^*_{1,0}[\rho(\phi,r)]$$

[rather than $P_{1,0}(\rho)$ and $P^*_{1,0}(\rho)$, as in Section IV].

Denoting the composition $P_{1,0}[\rho(\phi,r)]$ by $P_{1,0}(\phi,r)$, we have

$$P_{1,0}(\phi,r) = (1 - \rho(\phi,r))\left(1 - \exp\left\{\frac{-rS}{1-\rho(\phi,r)}\right\}\right), \tag{38}$$

which is plotted in Figs. 8 and 9 for the 200- and 100-second sampling options, respectively. The normal holding time used in these figures is 180 seconds, and the killer-trunk characteristics are drawn for $r = 5$, 10, and 15.

The increased "spread" between normal and killer $P_{1,0}$ characteristics obtained in this way is simply a consequence of exploiting the distinctly different occupancies of a normal and killer trunk in a 5XB trunk group. Figure 10 is a three-dimensional sketch of the composition of $P_{1,0}$ and ρ. Because all normal trunks in a 5XB group have the same mean occupancy, we see that a *single* $P_{1,0}$ vs ϕ characteristic suffices to describe

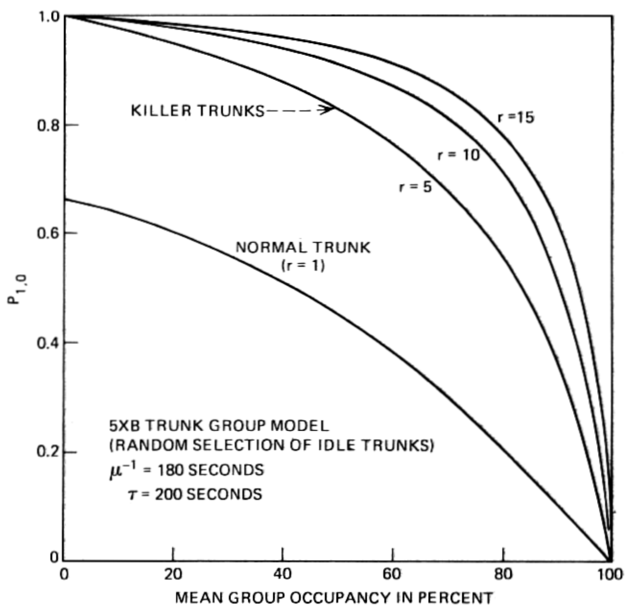


Fig. 8— $1 \rightarrow 0$ transition probability for the 200-second sampling option.

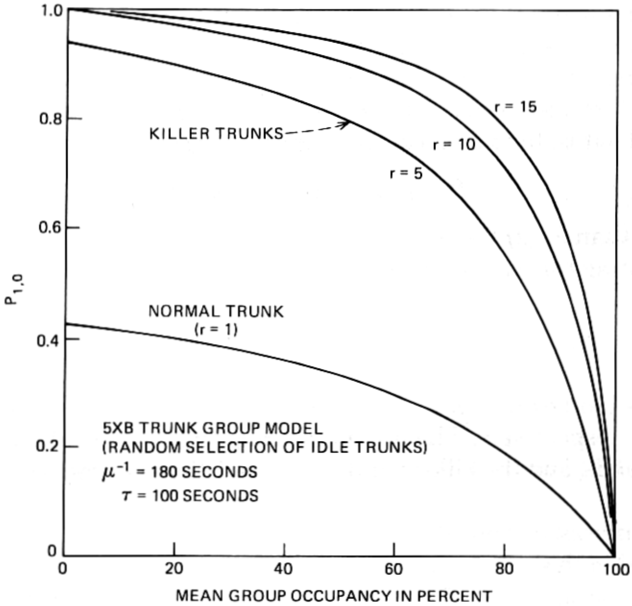


Fig. 9— 1 → 0 transition probability for the 100-second sampling option.

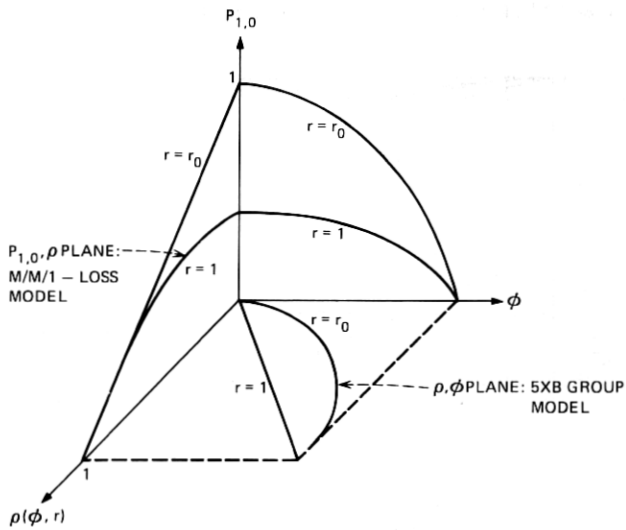


Fig. 10—Sketch of the composition of $P_{1,0}(\rho)$ with $\rho(\phi,r)$.

all normal trunks. This fact allows us to translate the individual trunk algorithm's development to this 5XB context with essentially only notational changes.

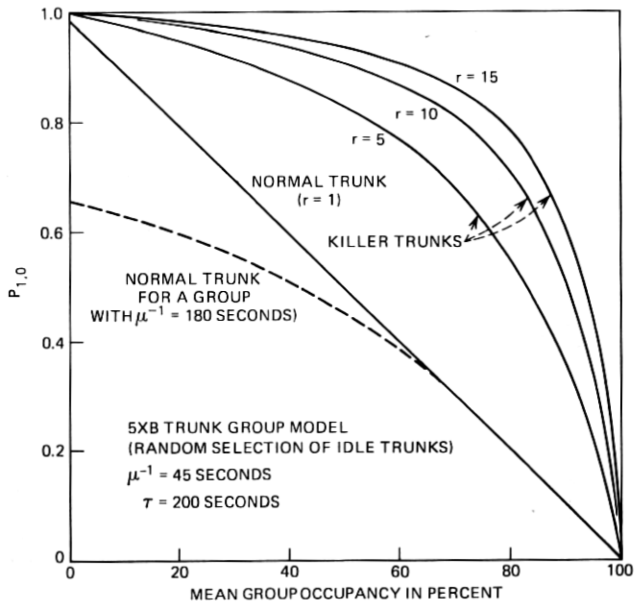


Fig. 11— 1 → 0 transition probability for the 200-second sampling option (mean group holding time = 45 seconds).

Figures 11 and 12 are plots of eq. (38) drawn for a normal group mean holding time of 45 seconds.* The normal trunk characteristic corresponding to 180 seconds is shown in dashed lines. We see that with 5XB grouping information factored into the picture, considerable discrimination exists between both normal trunk characteristics as well as between the normal trunk having a holding time of 45 seconds and the killer trunks. The discrimination that exists between the normal trunks permits us to make the 5XB group algorithm adaptive to the group mean holding time. [Although we will not pursue this topic, the basic idea is that $\Sigma_j t_{10}(j)/\Sigma_j n(j)$ (sums are over all trunks in the group) is an estimate of $p_{1,0}$ and can be used to decide which (of several) normal $p_{1,0}$ characteristics constitutes H_0.]

5.3 The ad hoc and the optimal 5XB group algorithms

We assume that the mean group holding time is known and consider formulating a hypothesis-testing problem similar to that in Section 4.1. Thus, we denote $P(t_{10}(m) = t/n(m) = n)$ by $P(t/n)$ and consider the two hypotheses:

* For normal holding times in the vicinity of 45 seconds, a killer parameter r in the range 3 to 5 probably is typical. An r of 10 or 15 in this context is unrealistic.

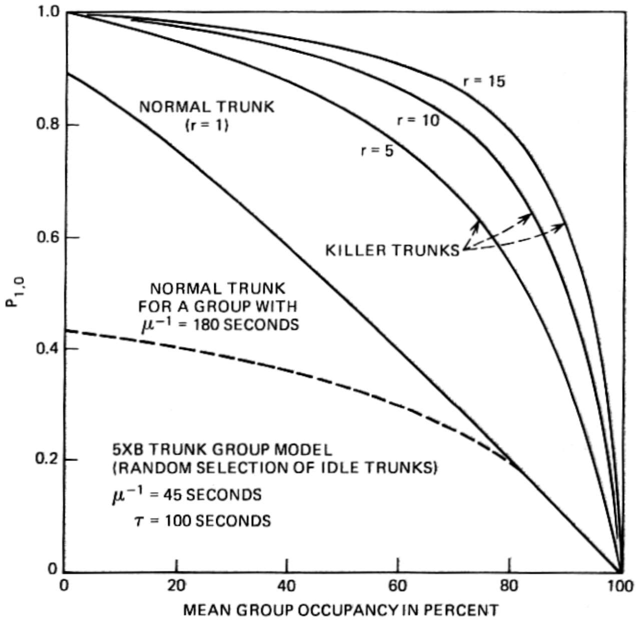


Fig. 12— $1 \rightarrow 0$ transition probability for the 100-second sampling option (mean grouping holding time = 45 seconds).

$$H_0\colon P(t/n) = b[t;n,P_{1,0}(\phi,1)]$$
$$H_1\colon P(t/n) = b[t;n,P_{1,0}(\phi,r)], \quad r \geqq r_0. \tag{39}$$

There are two differences between this formulation and the one in Section 4.1:

(*i*) The trunk occupancy ρ in Section 4.1 is replaced by the group occupancy ϕ.

(*ii*) The alternate hypothesis H_1 is composite since $P_{1,0}(\phi,r)$ for $r \geqq r_0$ are distinct.

The approach taken in dealing with (*ii*) is a natural one often adopted;[11] since $P_{1,0}(\phi,r)$ is monotone increasing in r {this follows from eq. (38) upon noting that $r/[1 - \rho(\phi,r)] = r + \phi/(1 - \phi)$}, then testing between H_0 and the *simple* alternate hypothesis

$$H_1\colon P(t/n) = b[t;n,P_{1,0}(\phi,r_0)],$$

say with type 1 and 2 errors α and β, respectively, implies that if the true state of nature is H_1 with $r = r_0' > r_0$ the resulting type 2 error will not exceed β. With this approach, we can simply translate the ad hoc algorithm results developed in Section 4.1 to this 5XB group context by making the appropriate changes in notation.

Thus, the ad hoc 5XB group algorithm can be described as follows: corresponding to the sequence of data accumulations (t_i^j, n_i^j) $i = 1,2,\cdots$ for the jth trunk in a group of N trunks, where t_i^j and n_i^j are the $1 \to 0$ transition and switch-count accumulations, respectively, during the ith accumulation period in which m_i scans are made, define s_i^j and S_i^j $i = 1,2,\cdots; j = 1,\cdots; N$ by

$$s_i^j = \alpha(\phi_i)t_i^j - a(\phi_i)n_i^j \tag{40a}$$

and

$$S_i^j = S_{i-1}^j + s_i^j \text{ with } S_0^j = 0, \tag{40b}$$

where ϕ_i is the group's occupancy during the ith accumulation period. The sequential test for the jth trunk in the group, $j = 1,\cdots,N$ is defined by

(i) Compute S_i^j, $i = 1,2,\cdots$ and defer making a decision as long as $T_0 < S_i^j < T_1$.

(ii) If $i = k$ corresponds to the first accumulation period for which

$$S_i^j \notin (T_0,T_1),$$

then

$$S_k^j \leqq T_0 \Rightarrow \text{trunk } j \text{ is normal}$$

$$S_k^j \geqq T_1 \Rightarrow \text{trunk } j \text{ is a killer.}$$

The weights $a(\phi)$ and $\alpha(\phi)$ are defined by

$$a(\phi) = \log \frac{1 - P_{1,0}(\phi,1)}{1 - P_{1,0}(\phi,r_0)} \tag{41a}$$

and

$$\alpha(\phi) = a(\phi) + \log \frac{P_{1,0}(\phi,r_0)}{P_{1,0}(\phi,1)}, \tag{41b}$$

where $P_{1,0}(\phi,r)$ is defined by eq. (38).

Just as in the individual trunk algorithm, the actual occupancy required to choose the weights a and α is unknown and must be estimated. Thus, the group occupancy ϕ_i during the ith accumulation period is estimated by

$$\hat{\phi}_i = \frac{1}{N} \sum_{j=1}^{N} n_i^j/m_i \tag{42}$$

Note that $\hat{\phi}_i$ is a "better" estimator than $\hat{\rho}_i = n_i/m_i$ (the estimator used in the individual trunk algorithm) in the following sense: given a 5XB group with all trunks normal and mean-group occupancy ϕ (in equilibrium), we have

(*i*) $E(\hat{\rho}) = \rho = \phi$
(*ii*) $E(\hat{\phi}) = \phi$
(*iii*) $\text{var}(\hat{\phi}) < \text{var}(\hat{\rho})$.

In addition to the better occupancy estimate available on a group basis, the fact that the group $P_{1,0}$ characteristics are "flatter and broader" then the individual trunk $P_{1,0}$ characteristic implies that the group algorithm more faithfully tracks the required weights, than does the individual trunk algorithm.

The ad hoc 5XB group algorithm has the same pleasant intuitive interpretation that the ad hoc individual trunk algorithm had (see eq. (13) and related discussion). It is also easy to show how the optimal individual trunk algorithm development of Section 4.2 carries over to the 5XB group context.

Thus, consider the two states of a trunk to be described by:

$$H_0\text{: } \{x_i\} \text{ Markovian, characterized by } \boldsymbol{\theta} = (P_{0,1}, P_{1,0})$$

$$H_1\text{: } \{x_i\} \text{ Markovian, characterized by } \boldsymbol{\theta}^* = (P^*_{0,1}, P^*_{1,0}),$$

where $P_{1,0} = P_{1,0}(\phi,1)$, $P^*_{1,0} = P_{1,0}(\phi,r_0)$, and [see eq. (2)]

$$P_{0,1}(\phi,r) = \frac{\rho(\phi,r)}{1-\rho(\phi,r)} P_{1,0}(\phi,r) \tag{43}$$

with $\rho(\phi,r)$ defined by eq. (37a). The assumptions that lead to a consideration of these two statistical hypotheses as a model of the normal and killer states of a trunk can be found in Section 4.2.

Proceeding as in Section 4.2 leads us to the optimum statistic $\hat{\ell}(\mathbf{x}_m)$ for distinguishing between the two simple hypotheses under consideration:

$$\hat{\ell}(\mathbf{x}_m) = [(\alpha - a)t_{10} - at_{11}] + [\tilde{b}t_{00} - (\tilde{\beta} - \tilde{b})t_{01}] + \log \frac{P^*(x_1)}{P(x_1)}, \tag{44}$$

where the parameters a and α are defined by eqs. (41a) and (41b) and the parameters $\tilde{b}$ and $\tilde{\beta}$ are defined by

$$\tilde{b} = \log \frac{1 - P_{0,1}(\phi,r_0)}{1 - P_{0,1}(\phi,1)} \tag{45a}$$

and

$$\tilde{\beta} = \tilde{b} + \log \frac{P_{0,1}(\phi,1)}{P_{0,1}(\phi,r_0)}. \tag{45b}$$

As in eq. (39), the alternate hypothesis is really composite since $[P_{0,1}(\phi,r), P_{1,0}(\phi,r)]$ for $r \geqq r_0$ are distinct. Since $P_{0,1}(\phi,r)$ is monotone decreasing in r, the approach discussed earlier of treating H_1 as a simple hypothesis with $r = r_0$ is followed. The parameters $\tilde{b}$ and $\tilde{\beta}$ are defined a bit differently than were the parameters b and β for the optimal individual trunk algorithm (see eq. (20a) and (20b)) in order to obtain nonnegative weights. Thus, in the individual trunk algorithm context, we had $P^*_{0,1} > P_{0,1}$ (eq. (19c)), but in the present 5XB group context we have $P_{0,1} > P^*_{0,1}$. The reason for this "flip-flop" is easy to see: for a *given group occupancy* ϕ, we are now contrasting the $0 \rightarrow 1$ transition probability for two trunks which differ, not only in their hang-up rates but also in their occupancies as well. Thus, the difference in the two occupancies dominates the effect that the hang-up rates alone have. Roughly speaking, the $0 \rightarrow 1$ transition probability of a trunk is approximately equal to its occupancy (conditioning on the last scan has little effect) and hence since $\rho(\phi,r_0) \ll \rho(\phi,1)$ it is clear that we should have $P_{0,1}(\phi,r) < P_{0,1}(\phi,1)$. Figures 13 and 14 are plots of $P_{0,1}(\phi,r)$ for the 200- and 100-second sampling options, respectively. In both figures, the killer trunk characteristics have been plotted for $r = 5$, 10, and 15 and a normal mean holding time of 180 seconds is assumed. These figures are very insensitive to the assumed normal mean holding time, since they essentially reflect eq. (37a), which is independent of the mean group holding time.

The "additional" statistic which appears in eq. (44),

$$\hat{b}t_{00} - (\hat{\beta} - \hat{b})t_{01}$$

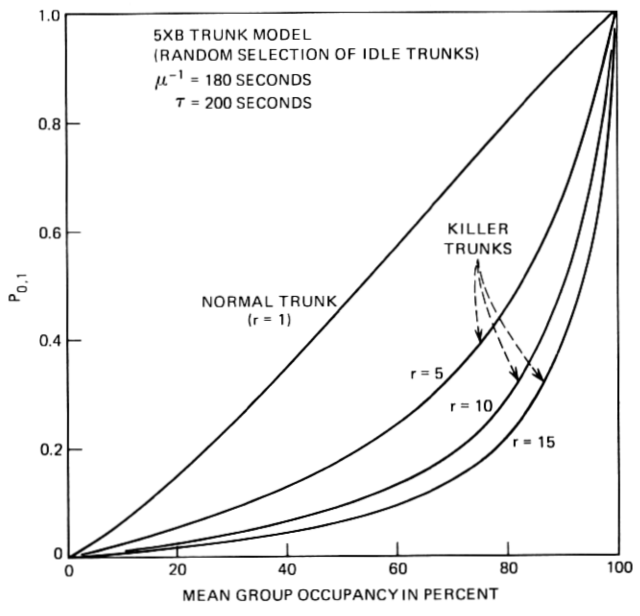


Fig. 13— $0 \rightarrow 1$ transition probability for the 200-second sampling option.

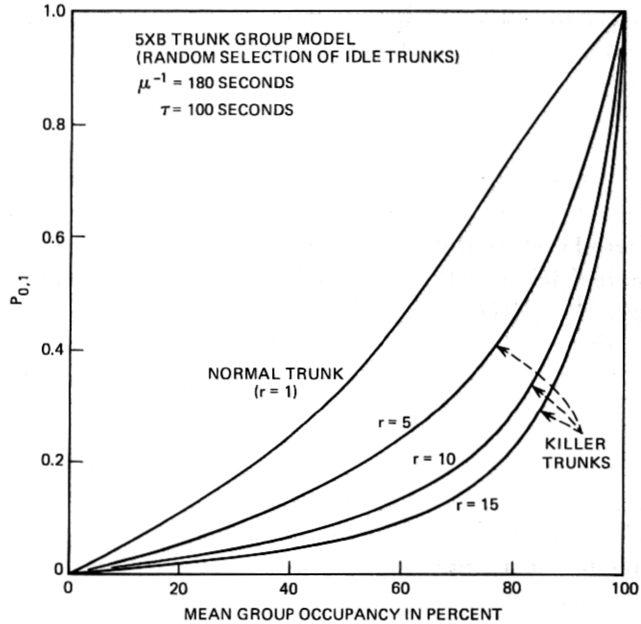


Fig. 14—0 → 1 transition probability for the 100-second sampling option.

shows that 0 → 0 transitions are weighted positively (evidence of a killer trunk*) and 0 → 1 transitions are weighted negatively (evidence of a normal trunk). This additional statistic is strongly influenced by the occupancy of a trunk, and only slightly by its hang-up rate.

Note also that the term $\log P^*(x_1)/P(x_1)$ is nonzero in the 5XB context since

$$\frac{P^*(x_1)}{P(x_1)} = \begin{cases} \dfrac{1}{r-(r-1)\phi} & \text{if} \quad x_1 = 1 \\ \dfrac{r}{r-(r-1)\phi} & \text{if} \quad x_1 = 0. \end{cases}$$

VI. PERFORMANCE OF THE 5XB GROUP ALGORITHMS

In common with all sequential detection algorithms, the time required by the killer-trunk detection algorithms to reach a decision (trunk normal or killer) is a random variable. In this section, we obtain an approximate formula for the *mean time* required by the 5XB group algorithms to reach a decision. This result is used to contrast the performance

* Killer trunks in the 5XB group model have very low occupancy, and hence 0 → 0 transitions are likely.

of the ad hoc and optimal algorithms as well as to point out the considerable effect that the sampling rate has on each. In addition, we also obtain an approximate expression for the false-alarm probability of the 5XB group algorithms. The analysis for the individual trunk algorithms, although differing in several respects from the group algorithms, involves the same sort of considerations and is omitted.

The analysis in this section assumes a server system in equilibrium and, therefore, the mean trunk-group occupancy ϕ is assumed constant. In addition, to simplify the analysis, we assume that ϕ is known; an approximate analysis which does not require this assumption is sketched in Section 6.1. A consequence of this assumption is that the algorithm weights are treated as constants rather than random variables. This assumption is not unreasonable because for multiple-hour accumulation periods var($\hat{\phi}$) is quite small ($\hat{\phi}$ is the switch-count estimate of ϕ). (Var($\hat{\phi}$) has been derived for an M/M/N-loss system.[15])

6.1 Mean statistic update

Corresponding to a sequence of trunk states $x_1, x_2, \cdots, x_m$ in an accumulation period with m scans, define a sequence of transition updates $z_2, \cdots, z_m$ by

$$z_n = \begin{cases} \hat{b} & \text{if} \quad (x_{n-1}, x_n) = (0,0) \\ -(\hat{\beta} - \hat{b}) & \text{if} \quad (x_{n-1}, x_n) = (0,1) \\ (\alpha - a) & \text{if} \quad (x_{n-1}, x_n) = (1,0) \\ -a & \text{if} \quad (x_{n-1}, x_n) = (1,1) \end{cases} \tag{46}$$

The optimum 5XB statistic (eq. (44)) may therefore be written:

$$\hat{\ell}(\mathbf{x}_m) = \sum_{i=2}^{m} z_i + \log \frac{P^*(x_m)}{P(x_m)} \tag{47}$$

In practice the end-effect term cannot be implemented and all the transitions in eq. (44) must be estimated in terms of $t(m)$, $n(m)$, and m. Thus, if we denote the implementable version of eq. (44) by $S_m(\phi, r_0)$,† use eqs. (11), (12), and (20) in eq. (44), and drop all end-effect terms we obtain

$$S_m(\phi, r_0) = \left[\alpha(\phi) \frac{t(m)}{2} - a(\phi) n(m) \right] + \left[\tilde{b}(\phi)[m - n(m)] - \tilde{\beta}(\phi) \frac{t(m)}{2} \right]. \tag{48}$$

† r_0 is the value of the killer parameter used in defining the alternate hypothesis and hence the algorithm weights (see eqs. (41) and (45)).

Now using eqs. (11), (12), and (20) once again, we see that $S_m(\phi,r_0)$ may be written

$$S_m(\phi,r_0) = \sum_{i=2}^{m} z_i + (\tilde{b}x_m^c - ax_m) + (\alpha + \tilde{\beta})\left(\frac{x_1 - x_m}{2}\right). \quad (49)$$

Since the 5XB group is in equilibrium, $z_2,\cdots,z_m$ are identically distributed and it is easily verified that their common mean is given by

$$E(z) = (\alpha P_{1,0} - a)\rho + (\tilde{b} - P_{0,1}\tilde{\beta})(1 - \rho). \quad (50)$$

We recall that ρ (eq. (37a)) as well as the transition probabilities (eqs. (38) and (43)) are functions of the group occupancy ϕ and the state r of the trunk. The weights a, α, $\tilde{b}$ and $\tilde{\beta}$ are only functions of ϕ for a specified choice of the parameter (r_0), which characterizes the alternate hypothesis. Thus, the mean statistic update for the optimum 5XB statistic and its "implementable version" is given by:

$$E\{\hat{\ell}(\mathbf{x}_m)\} = (m-1)E(z) + E\left\{\log\frac{P^*(x_1)}{P(x_1)}\right\} \quad (51a)$$

and

$$E\{S_m(\phi,r_0)\} = (m-1)E(z) + \tilde{b}(1-\rho) - a\rho. \quad (51b)$$

Note that it is easily shown that

$$E\left\{\log\frac{P^*(x_1)}{P(x_1)}\right\} = -\log\left[r_0 - (r_0 - 1)\phi\right] + \frac{r(1-\phi)}{r - (r-1)\phi}\log r_0$$

which is negative for $r = 1$ and positive for $r \geqq r_0$.

Although the increments (transition updates) defined in eq. (46) are identically distributed, they are not independent. In fact, since the state sequence $\{x_i\}$ has been modeled as a Markov chain (see Section 4.2), it is easy to see that the sequence $\{z_i\}$ defined by (46) is also a Markov chain. Relabeling the four natural states,

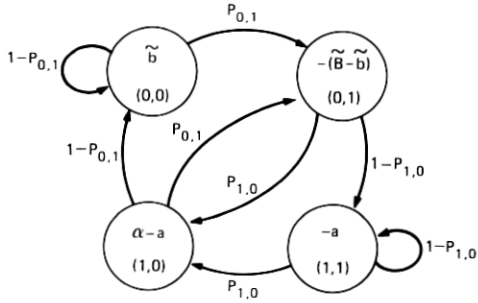
$$\tilde{b},\ -(\tilde{\beta} - \tilde{b}),\ (\alpha - a),\ -a,$$

of the chain $\{z_i\}$ by

$$0, 1, 2, 3,$$

respectively, it is easily seen that the one-step transition matrix π for this chain is given by

$$\pi = \begin{pmatrix} 1 - P_{0,1} & P_{0,1} & 0 & 0 \\ 0 & 0 & P_{1,0} & 1 - P_{1,0} \\ 1 - P_{0,1} & P_{0,1} & 0 & 0 \\ 0 & 0 & P_{1,0} & 1 - P_{1,0} \end{pmatrix},$$



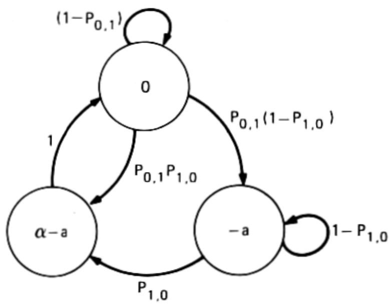


Fig. 15—State diagram for the ad hoc and optimal 5XB group statistics. (a) Optimal statistic. (b) Ad hoc statistic.

where $\pi_{i,j} = P(z_n = j/z_{n-1} = i)$. Of course, the stationary distribution $\boldsymbol{P}$ satisfying $\boldsymbol{P}\pi = \boldsymbol{P}$ is

$$\boldsymbol{P} = [(1-\rho)(1-P_{0,1}),\ (1-\rho)P_{0,1},\ \rho P_{1,0},\ \rho(1-P_{1,0})].$$

The state diagram corresponding to the Markov chain $\{z_i\}$ is shown in Fig. 15a. If $\tilde{b} = \tilde{\beta} = 0$, we obtain the ad hoc algorithm, for which the sequence $\{z_i\}$ is a three-state Markov chain, with natural states 0, $\alpha - a$ and $-a$. The state diagram for this chain is shown in Fig. 15b.

Figures 16 and 17 are plots of the mean statistic update vs group occupancy for the implementable version of the 5XB group algorithms. (Log base 10 is used in this paper.) These figures are drawn for a killer trunk with parameter $r = 10$. Also shown is the corresponding plot for a normal trunk ($r = 1$).

We close this section by indicating an approximate analysis of $E(z)$ which doesn't assume that ϕ is known. Thus, in practice, ϕ is unknown

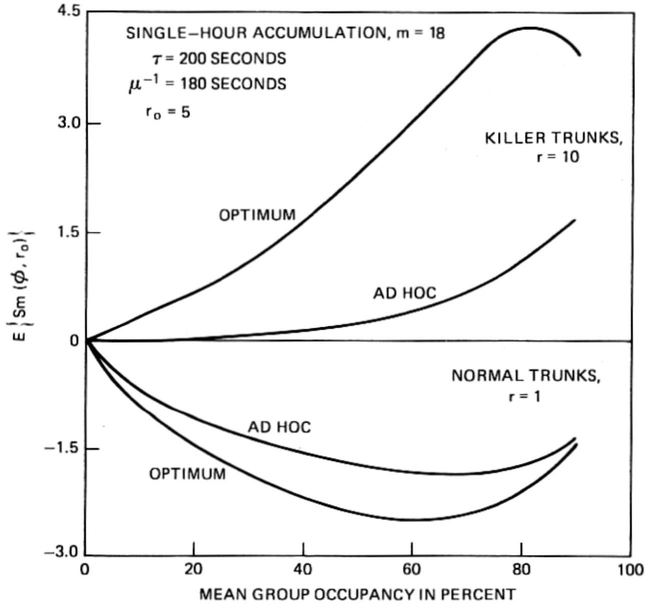


Fig. 16—Mean statistic update for the 5XB group algorithms—200-second sampling option.

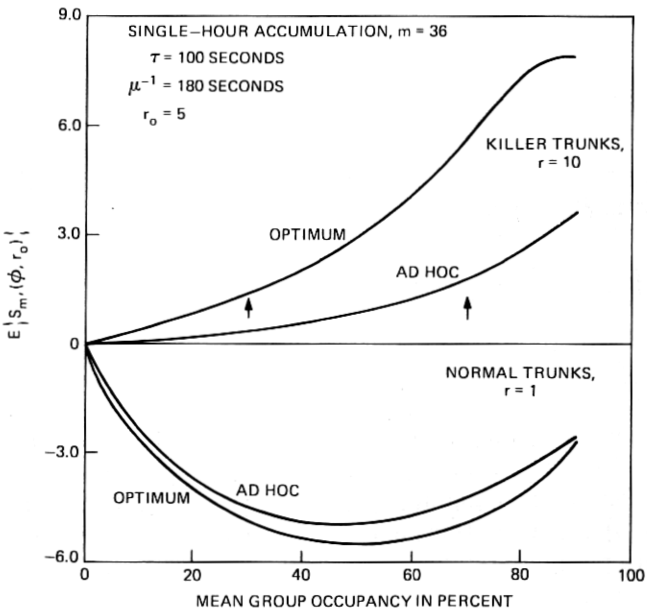


Fig. 17—Mean statistic update for the 5XB group algorithms—100-second sampling option.

and is estimated by $\hat{\phi}$. Hence eq. (46) should read:

$$z_n = \begin{cases} \tilde{b}(\phi_0) & \text{if} \quad (x_{n-1},x_n) = (0,0) \quad \text{and} \quad \hat{\phi} = \phi_0. \\ \vdots \end{cases}$$

Conditioning on $\hat{\phi}$, we have

$$E(z) = E\{E(z/\hat{\phi})\},$$

where

$$E(z/\hat{\phi}) = \tilde{b}(\hat{\phi})P(x_{n-1} = 0,x_n = 0/\hat{\phi}) + \cdots.$$

Now by assuming that (x_{n-1},x_n) and $\hat{\phi}$ are independent,* we get eq. (50) with the constant weights $\alpha(\phi),\cdots$ replaced by the mean values $E\{\alpha(\hat{\phi})\},\cdots$. The mean values can be approximated in either one of two ways:

(*i*) $E\{\alpha(\hat{\phi})\} \doteq \alpha[E(\hat{\phi})] = \alpha(\phi)$, which obviously amounts to assuming ϕ is known.

(*ii*) $E\{\alpha(\hat{\phi})\} \doteq \alpha(\phi) + \dfrac{\text{var}(\hat{\phi})}{2!} \dfrac{d^2\alpha^{\dagger}}{d\hat{\phi}^2}\bigg|_{\hat{\phi}=\phi}$,

which factors the available variance of $\hat{\phi}$ into the picture.

6.2 Mean time to detection

The basic structure of all the detection algorithms in this paper are the same: a statistic s_i is evaluated at the end of the ith accumulation period, $i = 1,2,\cdots$ and a decision is made the first time that the sum $s_1 + s_2 + \cdots$ falls outside an interval (T_0,T_1). Presumably, the random walk type statistic S_i has a negative drift under H_0 (trunk normal) and a positive drift under H_1 (trunk killer). Wald's SPRT always has the appropriate drift: if H_0 and H_1 correspond to the probability distributions $P_0(w)$ and $P_1(w)$, respectively, and if $\hat{\ell}(w)$ is defined by

$$\hat{\ell}(w) = \log \frac{P_1(w)}{P_0(w)},$$

then $E\{\hat{\ell}(w)\} < 0$ under H_0 and $E\{\hat{\ell}(w)\} > 0$ under H_1. The proof is immediate by using the inequality[12]

$$-\sum_i p_i \log p_i < -\sum_i p_i \log q_i,$$

* For reasonable-size trunk groups, we expect very little dependence between the sampled state process of an individual trunk $(x_1,\cdots)$ and the group process $\hat{\phi}$.

† This is a Taylor expansion to second order.

where $\{p_i\}$ and $\{q_i\}$ are distinct probability distributions. In Appendix E, we show that both the ad hoc ($\tilde{b} = \tilde{\beta} = 0$) and the additional ($a = \alpha = 0$) parts of $E(z)$ (eq. 50) have the appropriate drift. This in turn shows that both the ad hoc and additional parts of the 5XB group statistic (eq. (51a)) have the appropriate drift.

Suppose $Y_1, Y_2, \cdots$ are i.i.d. random variables with common mean μ and consider a random walk

$$S_n = \sum_{i=1}^{n} Y_i \quad n = 1,2,\cdots$$

with absorbing barriers at T_0 and T_1. If μ is small compared to T_0 and T_1, then the mean stopping time (mean number of steps to absorption) $E(n)$ is approximately given by

$$\frac{P_0T_0 + P_1T_1}{\mu},$$

where P_0 and P_1 are the probabilities of absorption at T_0 and T_1, respectively. This follows from Wald's identity[11]

$$E(S_n) = \mu E(n)$$

if we approximate the mean value of the random walk at absorption by $P_0T_0 + P_1T_1$.

In our detection theory context, P_0 and P_1 correspond to β and $1 - \beta$, respectively, if the trunk is a killer (β = probability of miss) and $1 - \alpha$ and α, respectively, if the trunk is normal (α = probability of false alarm). If we denote the mean number of accumulation periods needed to reach a decision under H_0 and H_1 by $E(T_N)$ and $E(T_k)$, respectively, and assume (*i*) successive statistic updates are independent and (*ii*) the mean statistic update is small compared to T_0 and T_1, we obtain

$$E(T_N) \doteq \frac{(1-\alpha)T_0 + \alpha T_1}{E\{S_m(\phi,r_0)\}} \tag{52a}$$

and

$$E(T_k) \doteq \frac{\beta T_0 + (1-\beta)T_1}{E\{S_m(\phi,r_0)\}}, \tag{52b}$$

where the mean statistic update $E\{S_m(\phi,r_0)\}$ is evaluated for $r = 1$ in (52a) and for $r \geqq r_0$ in (52b).

The assumption that successive statistic updates are independent is not strictly true if the successive statistic updates are contiguous. However, one expects that the slight (end-effect) dependence will not give rise to very much error.

Let T_{k_0} be the time required to decide (incorrectly) that a killer trunk is normal. Similarly, let T_{k_1} be the time required to decide (correctly) that a killer trunk is a killer. Then we have

$$E(T_k) = \beta E(T_{k_0}) + (1 - \beta)E(T_{k_1}),$$

which suggests that

$$E(T_{k_1}) \doteq E(T_k) \quad \text{if} \quad \beta \ll 1. \tag{53}$$

The moments of the conditional stopping times T_{k_0} and T_{k_1} can be obtained by using a well known technique of Wald's,* and one finds that approximation (53) is reasonable if T_1 and $(-T_0)$ are sufficiently large.

In our faulty-trunk detection context, a type 1 error ("false alarm") may result in the misuse of craft resources (e.g., testing a perfectly good trunk). A type 2 error ("miss") on the other hand, will result in an increased time to detection. Assuming type 1 and 2 errors of 10^{-6} and 10^{-2}, respectively (realistic implementation values), implies approximate thresholds $T_0 = -2$ and $T_1 = 6$ (formulae 4.5c and d, log base 10). With these parameter values, the mean detection time $E(T_{k_1})$ can be approximated by

$$E(T_{k_1}) \doteq \frac{T_1}{E\{S_m(\phi,r_0)\}}. \tag{54}$$

For a normal trunk, the mean statistic update is comparable to T_0 and hence expression (52a) isn't applicable, nor is it needed since the 5XB group algorithm reaches a decision on a normal trunk after one or two updates.

Figures 18 and 19 are plots of the mean-detection time vs group occupancy for the implementable versions of the 5XB group algorithms. [The dashed line portion of Figs. 18 and 19 indicates where approximation (54) involves considerable error (e.g., the region in which $E\{S_m(\phi,r_0)\}$ is a significant fraction of T_1).] It is apparent from these figures that

(*i*) The mean detection time for both the ad hoc and optimal algorithms is enhanced by using the 100-second sampling option. This enhancement is far more pronounced for the ad hoc algorithm.

(*ii*) The optimal algorithm is "faster" than the ad hoc algorithm. This contrast is greater for the 200-second sampling option.

6.3 False alarm probability of the 5XB group algorithms

If $B(\alpha,\beta)$ and $A(\alpha,\beta)$ are the test thresholds that result in type 1 and 2 errors α and β, respectively, for a SPRT, Wald showed[11] that using

* See eq. 158 and 159 in Appendix A.5.2 of Ref. 11. The method ignores (as usual) the "excess over the boundaries" and hence yields approximate results.

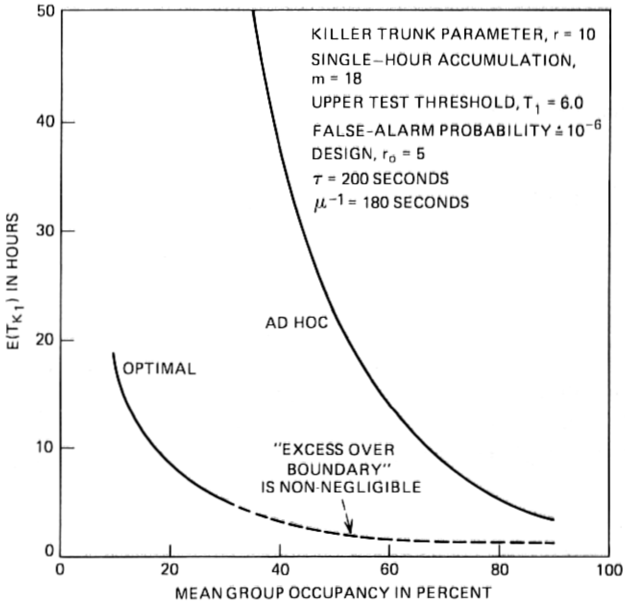


Fig. 18—Mean time to detection for the 5XB group algorithms—200-second sampling option.

thresholds T_0 and T_1 defined by*

$$T_0 = \log\left(\frac{\beta}{1-\alpha}\right) \quad \text{and} \quad T_1 = \log\left(\frac{1-\beta}{\alpha}\right) \text{ (assuming } \alpha + \beta < 1\text{)}$$

yield type 1 and 2 errors α' and β' which satisfy

$$\beta' \leqq \frac{\beta}{1-\alpha} \quad \text{and} \quad \alpha' \leqq \frac{\alpha}{1-\beta}.$$

The proof of this result is trivial and depends only on the assumption that the SPRT terminates with probability 1. This assumption is satisfied by a wide class of SPRTs,[16] including the case of interest to us, where the underlying distribution is "Markovian". Thus, the probability of false alarm (α') for a SPRT (an example of which is the "optimal" 5XB group algorithm) satisfies $\alpha' \leqq 10^{-T_1}$.

Because the ad hoc algorithm is not a SPRT (the true underlying distribution is not binomial, see Section 4.3), we may wish to study the consequences of using the above thresholds appropriate for a SPRT. To do this, we can proceed in (at least) two distinct ways:

(*i*) consider the statistic $S_m(\phi, r_0)$ to be the basic update in the algorithm.

* This assumes that the log likelihood ratio is used in defining the SPRT.

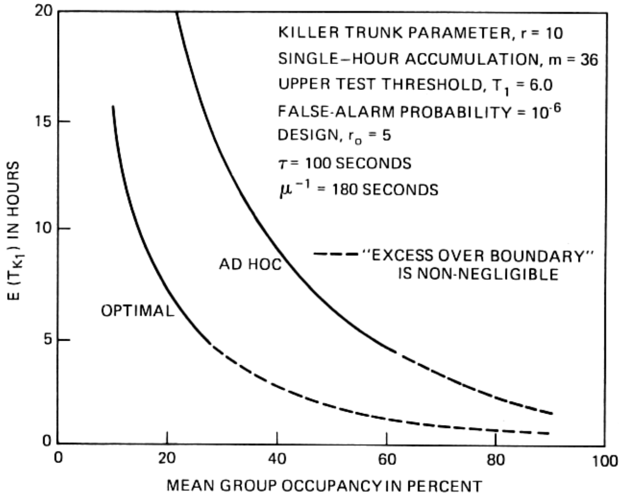


Fig. 19—Mean time to detection for the 5XB group algorithms—100-second sampling option.

(*ii*) Consider the transition update $Z_i(x_{i-1},x_i)$ to be the basic update in the algorithm.

In either case, we study a process of the form $S_j = \Sigma_{i=1}^{j} x_i$, $j = 1,2,\cdots$ with absorbing barriers $T_0 < 0 < T_1$. The advantage of proceeding as in (*i*) above is that the increments x_i may be assumed to be i.i.d. We briefly sketch this approach.

Consider a random walk S_j, $j = 1,2,\cdots$ with i.i.d. increments $\{x_i\}$. If n is a stopping time associated with S_j, $j = 1,2,\cdots$, Wald's fundamental identity[11,17] is given by

$$E\{e^{S_n t}\chi(t)^{-n}\} = 1 \text{ (all } t \text{ satisfying } |\chi(t)| \geqq 1), \tag{55}$$

where

$$\chi(t) = E\{e^{xt}\}$$

is the moment generating function corresponding to the common distribution of the increments. If P_{T_0} and P_{T_1} denote the probabilities of absorption at T_0 and T_1, respectively, then rewriting (55) in a standard way yields

$$P_{T_0}E\{e^{S_n t}\chi(t)^{-n}/S_n \leqq T_0\} + P_{T_1}E\{e^{S_n t}\chi(t)^{-n}/S_n \geqq T_1\} = 1. \tag{56}$$

Now if the x_i take on both positive as well as negative values with nonzero probability and have a nonzero mean, then the equation

$$\chi(t) = 1$$

has only one nonzero root h_0 with the property that h_0 and $E(x)$ have opposite sign.* That is, $E(x) > 0 \Rightarrow h_0 < 0$ and $E(x) < 0 \Rightarrow h_0 > 0$. Assuming that the excess of S_n over the boundaries is small, eq. (56) yields the standard approximation[11,17]

$$P_{T_0} \doteq \frac{e^{h_0 T_1} - 1}{e^{h_0 T_1} - e^{h_0 T_0}} \tag{57a}$$

and

$$P_{T_1} \doteq \frac{1 - e^{h_0 T_0}}{e^{h_0 T_1} - e^{h_0 T_0}} \tag{57b}$$

Note that the probability of false alarm (type 1 error) corresponds to P_{T_1} if the random walk increment is that of a *normal* trunk $[S_m(\phi,r_0),\ r = 1]$. Similarly, the probability of miss (type 2 error) corresponds to P_{T_0} if the random walk increment is that of a killer trunk $[S_m(\phi,r_0),\ r \geqq r_0]$.

To use these approximations, we must compute the moment generating function

$$\chi(\mu) = E\{e^{[\alpha(\phi)t/2 - a(\phi)n]\mu}\} \tag{58}$$

by using the joint distribution $p(t,n)$ derived in Appendix B. [Note that our discussion applies equally well to the additional and optimal statistics. In general, we need $E\{\exp\,[S_m(\phi,r_0)\mu]\}$, where $S_m(\phi,r_0)$ is given by eq. (48).] Choosing the test thresholds T_0 and T_1 for the ad hoc algorithm according to the Wald SPRT formulae [eqs. (10c) and (10d)] typically results in $P_{T_0} < \beta$ and $P_{T_1} < \alpha$.

VII. SUMMARY

A class of killer-trunk detection algorithms has been developed that use the individual trunk usage and transition accumulations available in EADAS/ICUR. Because this data is essentially a sufficient statistic for the Markov chain used to model the (unobservable) sampled data, one of the algorithms developed is Wald's celebrated SPRT.

The detection algorithms developed can be partitioned in two natural ways:

(*i*) By sampling rate (100 or 200 seconds).
(*ii*) According to whether grouping information is used.

The algorithms which do not use grouping information are applicable to all trunks (one way or two way) independent of the type of switching machine used. A version of one of these individual trunk algorithms is currently in use in ICAN, testing trunks on 1XB, XBT and step-by-step

* See Appendix A.2.1 of Wald's original treatise (Ref. 11).

switching machines. The algorithms that exploit grouping information detect killers more quickly but are tailored to a specific switching machine. A "group" algorithm of this type is currently being used to test trunks associated with 5XB switching machines.

In the course of this study several problems of independent interest were studied. These include:

(*i*) The server covariance in a M/G/1-loss and GI/M/1-loss system.

(*ii*) The structure of the likelihood statistic that arises in testing simple hypotheses characterized by a binary valued Markov chain.

(*iii*) The occupancy of nonidentical trunks in a random-selection (Markovian) loss system.

The major conclusion in this study is, of course, that accumulated switch-count and state-transition data on *individual* trunks (based on sampling intervals on the order of a normal holding time) can be used to *reliably* detect abnormally short holding time trunks. Moreover the (near) optimal sequential detection algorithms using this accumulated data are easily exhibited, simple in structure, and intuitively appealing.

VIII. ACKNOWLEDGMENTS

Various stages of this study benefited from discussions with A. E. Eckberg, R. L. Franks, S. Horing, E. J. Messerli, and T. E. Rutt.

APPENDIX A

Sensitivity of the Transition Probabilities to Modeling Assumptions

To get an idea about the sensitivity of the algorithms to some of the modeling assumptions, the transition probability $P_{1,0} = P(x_{t+\tau} = 0/x_t = 1)$ was studied for the following two cases:

(*i*) M/G/1-loss, where the service distribution function $F(\cdot)$ is the mixed exponential given by

$$F(t) = 1 - d_1 e^{-x_1 t} - d_2 e^{-x_2 t} \quad t \geqq 0$$

(*ii*) GI/M/1-loss, where the arrival process is the switched-Poisson process[8] commonly used to model overflow traffic.

Because the methods used to obtain $P_{1,0}$ for these two models differ, we discuss these models separately.

A.1 The M/G/1-loss model

An observer viewing the server in an M/G/1-loss system sees an alternating sequence of busy and idle intervals. The busy intervals are distributed according to some distribution $F(\cdot)$ and are independent.

The idle intervals are exponentially distributed with mean λ^{-1} (λ = mean arrival rate) and are independent. Thus, the sequence of alternating busy and idle intervals constitutes an alternating renewal process. In this context, the conditional probability

$$P_{1,0} = P(x_{t+\tau} = 0/x_t = 1), \tag{59}$$

where

$$x_t = \begin{cases} 1 & \text{if server is } \textit{busy} \text{ at epoch } t \\ 0 & \text{if server is } \textit{idle} \text{ at epoch } t \end{cases}$$

has already been studied,[18] and we have the following result:

Theorem 4: Consider an M/G/1-loss system in equilibrium with a service time distribution $F(t)$ having Laplace transform $f^*(s)$. If $P^*_{1,0}(s)$ denotes the Laplace transform of $P_{1,0}(\tau)$, then we have

$$P^*_{1,0}(s) = \frac{\mu[1 - f^*(s)]}{s\{s + \lambda[1 - f^*(s)]\}} \tag{60}$$

where λ and μ are the mean arrival and service rates, respectively.

Proof: See Section 7.4 of Ref. 18 [$f_2^*(s) = \lambda/s + \lambda$ and $f_1^*(s) = f^*(s)$].

For the mixed exponential service time distribution mentioned above let

$$x_1 = 2\mu d$$

and

$$x_2 = 2\mu(1 - d), \quad 0 \leqq d \leqq 1.$$

If T is distributed according to this two-parameter (μ,d) family of distributions, then

(*i*) $E(T) = \mu^{-1}$.

(*ii*) $\text{var}(T) = \mu^{-2}\left(\frac{2-\delta}{\delta}\right)$, where $\delta = 4\,d(1 - d)$.

(*iii*) $c(T) = \frac{\sigma(T)}{E(T)} = \left(\frac{2-\delta}{\delta}\right)^{1/2}$.

Thus, the mean is fixed at μ^{-1} and the coefficient of variation satisfies $c \geqq 1$, with equality occurring when $\delta = 1$, which corresponds to the M/M/1-loss system.

Equation (60) is easily inverted for this family of mixed exponential distributions, obtaining the following result:

$$P_{1,0}(\rho,\tau) = (1 - \rho) + \frac{r_1 + \delta}{r_1(r_1 - r_2)} e^{r_1\mu\tau} + \frac{r_2 + \delta}{r_2(r_2 - r_1)} e^{r_2\mu\tau}, \tag{61}$$

where

$$r_1 = -\left(\frac{a+2}{2}\right) + \left[(1-\delta)(1+a) + \left(\frac{a}{2}\right)^2\right]^{1/2}$$

$$r_2 = -\left(\frac{a+2}{2}\right) - \left[(1-\delta)(1+a) + \left(\frac{a}{2}\right)^2\right]^{1/2}$$

$$\delta = 4d(1-d), \quad a = \frac{\rho}{1-\rho}$$

and $\rho = \lambda/(\lambda + \mu)$ is the mean occupancy of the server.

In Fig. 5, $P_{1,0}(\rho,\tau)$ [eq. (61)] is plotted vs ρ for several different values of c (c = 1, 1.5, and 2) assuming

(*i*) μ^{-1} = 180 seconds
(*ii*) τ = 200 seconds.

Also shown is a plot of $P_{1,0}$ vs ρ for a killer trunk with $r = 10$ [μ in (61) is replaced by $r\mu$]. The normal trunk $P_{1,0}$ characteristic with $c = 1$ corresponds to the M/M/1-loss system.

A.2 The GI/M/1-loss model

The covariance function $R(\cdot)$ for the GI/M/1-loss model with a switched Poisson arrival process has the form:

$$R(\tau) = c_1\, e^{w_1\tau} + c_1 e^{w_2\tau} + c_3 e^{-(\omega+\gamma)\tau}, \tag{62}$$

where the coefficients c_i and the exponents w_i are messy expressions involving the three switch parameters ω, γ, and λ (Ref. 8) and the mean service rate μ. The derivation of this covariance function is straightforward but tedious and is therefore omitted. (For the switched Poisson arrival prccess, the Markovian state equations can be solved for $P(x_t,x_{t+\tau})$, where x_t = state of server at epoch t.) Our purpose here is to explain how eq. (62) was used in generating Fig. 4.

If ρ is the mean occupancy of the server ($\rho = E(x_t)$) and a is the offered load in a GI/M/1-loss system, then it is easy to show that the peakedness[6] z satisfies

$$z = a\left(\frac{1-\rho}{\rho}\right). \tag{63}$$

For GI/M/1-loss system the call congestion is $\phi(\mu)$ and is related to the time congestion (ρ) by $a(1-\phi(\mu)) = \rho$. So using $z(\mu) = 1/[1-\phi(\mu)] - a$ yields the result.[2] ($\phi(\cdot)$ is the L.S. transform of the interarrival time distribution.) Therefore, specifying ρ and z uniquely determines a. Hence, with a and z known, we obtain the equivalent random parameters and use the three-moment match to obtain the switch parameters (see

Ref. 8). Using this procedure, we obtain $R(\tau)$ vs ρ parameterized by z. Equation (1) then yields $P_{1,0}$.

APPENDIX B

The Likelihood Statistic Based on the Observable Data

For ease of derivation, the likelihood statistic derived in Section 4.2 was based on the raw data $\mathbf{x}_m = (x_1,\cdots,x_m)$ rather than on the observable data $[t(m),n(m)]$; $t(m)$ and $n(m)$ are defined in Section 2.1. We will now study $\hat{\ell}_m(t,n)$ and verify that the two statistics differ only in their end-effect terms. We will also examine the end-effect term based on $t(m)$ and $n(m)$ and show how it "tracks" the end-effect term based on x_1 and x_m.

We begin by expressing the probability of $\mathbf{x}_m = (x_1,\cdots,x_m)$ in terms of $t(m)$, $n(m)$, x_1, and x_m:

Lemma 3: If $\{x_i\}$ is a binary state stationary Markov chain with transition probabilities $P_{0,1}$ and $P_{1,0}$ and if $\mathbf{x}_m = (x_1,\cdots,x_m)$, then

$$P(\mathbf{x}_m) = P(x_1) \times Q(x_1,x_m) \times P_{1,0}^{t/2}(1 - P_{1,0})^{n-(t/2)} \times P_{0,1}^{t/2}(1 - P_{0,1})^{m-n-(t/2)}, \quad (64)$$

where

$$Q(x_1,x_m) = P_{1,0}^{(x_1-x_m)/2}\, P_{1,1}^{-(x_1+x_m)/2}\, P_{0,1}^{-(x_1-x_m)/2}\, P_{0,0}^{(x_1+x_m)/2+1} \quad (65)$$

and

$$P(x_1) = \begin{cases} \rho & \text{if} \quad x_1 = 1 \\ 1-\rho & \text{if} \quad x_1 = 0 \end{cases}.$$

Proof: This result is obtained from lemma 2 using eqs. (12a) and (12b).

For convenience, we introduce the following notation

(*i*) $f\left(\frac{t}{2};n,P_{1,0}\right) = P_{1,0}^{t/2}(1 - P_{1,0})^{n-(t/2)}$. If t is even then of course

$$f\left(\frac{t}{2};n,P_{1,0}\right) = \frac{b\left(\frac{t}{2};n,P_{1,0}\right)}{\binom{n}{t/2}}.$$

(*ii*) $g\left(\frac{t}{m};m-n,P_{0,1}\right) = P_{0,1}^{t/2}(1 - P_{0,1})^{m-n-t/2}$.

(*iii*) $P_{x,y}(t,n) = P(t(m) = t, n(m) = n, x_1 = x, x_m = y)$.

(*iv*) $S_{x,y}(t,n)$ = number of binary m-tuples satisfying $t(m) = t$, $n(m) = n$, $x_1 = x$, and $x_m = y$.

Thus, we can write

$$P(t(m) = t, n(m) = n) = \sum_{x=0}^{1} \sum_{y=0}^{1} P_{x,y}(t,n) \tag{66}$$

and

$$P_{x,y}(t,n) = f\left(\frac{t}{2}; n, P_{1,0}\right) g\left(\frac{t}{2}; m - n, P_{0,1}\right) P(x) S_{x,y}(t,n) Q(x,y). \tag{67}$$

Equations (66) and (67) imply that

$$P(t, n(m) = t, n(m) = n)$$

$$= f\left(\frac{t}{2}; n, P_{1,0}\right) g\left(\frac{t}{2}; m - n, P_{0,1}\right) \left\{ \sum_{x=0}^{1} \sum_{y=0}^{1} P(x) S_{x,y}(t,n) Q(x,y) \right\} \tag{68}$$

and therefore it is easily seen that:

$$\hat{\ell}_m(t,n) = \log \frac{P^*(t(m) = t, n(m) = n)}{P(t(m) = t, n(m) = n)},$$

$$= \left[\alpha \frac{t}{2} - an \right] + \left[\beta \frac{t}{2} - b(n - m) \right] + E(t,n), \tag{69}$$

where

$$E(t,n) = \log \left\{ \frac{\sum_{x=0}^{1} \sum_{y=0}^{1} P^*(x) S_{x,y}(t,n) Q^*(x,y)}{\sum_{x=0}^{1} \sum_{y=0}^{1} P(x) S_{x,y}(t,n) Q(x,Y)} \right\}. \tag{70}$$

Comparing eqs. (70) and (22), we see that $\hat{\ell}_m(t,n)$ and $\hat{\ell}(\mathbf{x}_m)$ differ only in their respective end-effect terms. The following result is perhaps a bit surprising:

Lemma 4: $\hat{\ell}(t,n) = \hat{\ell}(\mathbf{x}_m)$ *if* t *is odd.*

Proof: t odd $\Rightarrow x_1 \neq x_m \Rightarrow e(0,1) = e(1,0) = a + b/2$ (see eq. (22b)). $\boldsymbol{t}$ odd $\Rightarrow S_{0,0}(t,n) = S_{1,1}(t,n) = 0$ so $E(t,n)$ for t odd may be written:

$$E(t,n) = \log \frac{(1 - \rho) S_{0,1} Q^*(0,1) + \rho S_{1,0} Q^*(1,0)}{(1 - \rho) S_{0,1} Q(0,1) + \rho S_{1,0} Q(1,0)}.$$

Now using (4) $E(t,n)$ for t odd can be manipulated into the following form

$$E(t,n) = \log \frac{(1 - \rho) S_{0,1} Q^*(0,1) + \rho S_{1,0} Q^*(1,0)}{(1 - \rho) S_{0,1} Q(0,1) + \rho S_{1,0} Q(1,0)}.$$

but $P_{0,1}/P_{1,0} = P^*_{0,1}/P^*_{1,0} = \rho/1 - \rho$ which completes the proof.

If t is even then $x_1 = x_m$,

$$e(0,0) = b$$

and

$$e(1,1) = a.$$

For even t, eq. (70) can be written as

$$E(t,n) = \log \left\{ \frac{S_{0,0} + S_{1,1}}{\left(\frac{1-\rho}{1-P_{01}}\right) S_{0,0} + \left(\frac{\rho}{1-P_{10}}\right) S_{1,1}} \right\}.$$

Thus, for even t, $E(t,n)$ is a complicated function* of t and n. Note however that $E(t,n) \leqq 0$ and

$$E(t,n) = \begin{cases} a & \text{if} \quad S_{0,0}(t,n) = 0 \\ b & \text{if} \quad S_{1,1}(t,n) = 0 \end{cases}$$

It can be shown that

$$\frac{S_{1,1}(t,n)}{S_{0,0}(t,n)} = \frac{n - \frac{t}{2}}{m - n - \frac{t}{2}}$$

if (t,n) is such that $S_{1,1}$ and $S_{0,0}$ are nonzero.

APPENDIX C

The End Effect $E\{I(x_{im};x_{im+1})\}$

The end effect

$$E\{I(x_{im};x_{im+1})\} = H(x_{im+1}) - H(x_{im+1}/x_{im}),$$

where

$$H(x_{im+1}/x_{im}) = (1-\rho)H(x_{im+1}/0) + \rho H(x_{im+1}/1)$$

and

$$H(x_{im+1}),\ H(x_{im+1}/0) \text{ and } H(x_{im+1}/1)$$

are the binary entropy function $\mathcal{H}(x)$ evaluated at ρ, $P_{0,1}$ and $P_{1,0}$, respectively ($\mathcal{H}(x) = -x \log x - (1-x)\log(1-x)$, $0 \leq x \leq 1$).

Table I exhibits $E\{I(x_{im};x_{im+1})\}$ and $E\{\hat{\ell}(\mathbf{x}_m)\}$ as a function of the trunks occupancy ρ—for a normal trunk, with a single-hour accumula-

Table I — End effect as a function of occupancy

	(τ = 100 seconds)		(τ = 200 seconds)	
ρ	$E\{I(x_{im+1};x_{im})\}$	$E\{\hat{\ell}(\mathbf{x}_m)\}$	$E\{I(x_{im+1};x_{im})\}$	$E\{\hat{\ell}(\mathbf{x}_m)\}$
0.10	0.04	−1.36	0.01	−0.22
0.20	0.05	−1.60	0.01	−0.20
0.30	0.04	−1.50	0.01	−0.15
0.40	0.03	−1.21	0.01	−0.09
0.50	0.02	−0.84	0.00	−0.04
0.60	0.01	−0.47	0.00	−0.01

tion period (m = 36 and 18 for the 100- and 200-second sampling options, respectively). The M/M/1-loss model is used with $1/\mu$ = 180 seconds. It is clear that the mean end effect is negligible compared to the mean statistic update.

APPENDIX D

Occupancy Formulae for a Random-Selection Loss System

Theorem: Consider an N server Markovian loss system with random selection of idle servers and

(*i*) $N - 1$ *servers with mean service rate* μ.
(*ii*) *1 server with mean service rate* $r\mu$ (r > 0).
(*iii*) *Mean arrival rate* λ.

Let ρ_r *and* ρ_r^* *denote the mean occupancy of the servers with mean service rates* μ *and* $r\mu$, *respectively, and let* $\hat{B}$ *denote the blocking (call congestion). Also let* ϕ *and* ϕ^- *denote the carried load per server in an N server and* $N - 1$ *server Markovian loss system, respectively, (assuming all servers have rate* μ*), given an offered lead* $a = \lambda/\mu$. *That is,* $\phi = a[1 - B(N,a)]/N$ *and* $\phi^- = a[1 - B(N-1,a)]/(N-1)$, *where* $B(\cdot\,,\cdot)$ *is the Erlang blocking formula. Then*

$$(i)\ \ \rho_r^* = \frac{\phi}{r-(r-1)\phi}. \tag{71}$$

$$(ii)\ \ \hat{B}(N,a,r) = \frac{B(N,a)}{r-(r-1)\phi}\ [B(\cdot\,,\cdot) \text{ is the Erlang blocking formula}]. \tag{72}$$

$$(iii)\ \ \phi \geqq \rho_r \geqq \left(\frac{r-(r-1)[N/(N-1)]\phi}{r-(r-1)\phi}\right)\phi. \tag{73}$$

$$(iv)\ \ \rho_r = \left(\frac{r-(r-1)\phi^-}{r-(r-1)\phi}\right)\phi. \tag{74}$$

Proof:

(*i*) This part of the theorem follows immediately from eq. (34b)

$$\rho_r^* = \frac{1}{1 + \frac{rN}{a} - r[1 - B(N - 1,a)]}$$

by using the well known Erlang B recursion[6]

$$\frac{B(N,a)}{1 - B(N,a)} = \frac{a}{N} B(N - 1,a)$$

and the expression for the mean group occupancy

$$\frac{a[1 - B(N,a)]}{N} = \phi.$$

(*ii*) This part of the theorem follows immediately from eq. (34a) by using the above expression for mean-group occupancy.

(*iii*) The lower bound part of eq. (73) follows from eq. (34c) by noting that $\hat{B}(N,a,r) \leqq B(N,a)$† and using eq. (71) of this theorem. Thus, we obtain

$$\frac{r\phi}{r - (r - 1)\phi} + (N - 1)\rho_r \geqq a[1 - B(N,a)] = N\phi$$

which can be arranged to yield the lower bound in eq. (73). The upper bound in (73) is an immediate consequence of (74) since $\phi^- \geqq \phi$. Thus, it remains to prove (74).

(*iv*) We prove this part as follows: Using eqs. (71) and (72) in the conservation eq. (34c),

$$r\rho_r^* + (N - 1)\rho_r = a(1 - \hat{B}) \tag{75}$$

yields

$$\rho_r = \frac{a(r - 1)(1 - \phi) + (N - r)\phi}{(N - 1)[r - (r - 1)\phi]}. \tag{76}$$

Therefore, eq. (74) holds if and only if

$$[r - (r - 1)\phi^-] \times \phi = \frac{a(r - 1)(1 - \phi) + (N - r)\phi}{(N - 1)}. \tag{77}$$

The right-hand side of eq. (77) can be rewritten as:

$$\text{rhs} = \frac{a}{N(N - 1)}\{r(N - 1 - a) + a - B[N - (a + 1)r + a]\} \tag{78}$$

† This follows eq. (34a) by noting that $a[1 - B(N,a)] \leqq N$.

using

$$\phi = \frac{a[1 - B(N,a)]}{N}.$$

Similarly, using $\phi^{-} = a[1 - B(N - 1,a)]/(N - 1)$, the left-hand side of eq. (77) can be written as:

$$\text{lhs} = \{r - (r - 1)(a/N - 1)[1 - B(N - 1,a)]\}(a/N)[1 - B(N,a)]. \quad (79)$$

Now using the Erlang B recursion formula in eq. (79) and rearranging terms yields eq. (78). Thus, eq. (77) and consequently eq. (74), of the theorem is proved.

APPENDIX E

The Sign of the Mean Statistic Update

The mean statistic update of the optimal 5XB group algorithm is given by eq. (51a):

$$E\{\hat{\ell}(\mathbf{x}_m)\} = (m - 1)\{(\alpha P_{1,0} - a)\rho + (\tilde{b} - \tilde{\beta} P_{0,1})(1 - \rho)\} + \log \frac{P^*(x_1)}{P(x_1)}.$$

The mean update of the ad hoc and the additional statistics correspond to $\tilde{b} = \tilde{\beta} = 0$ and $a = \alpha = 0$, respectively. The following lemma is needed to study the sign of $E\{\hat{\ell}(\mathbf{x}_m)\}$:

Lemma 5: If $0 < q < p < 1$ *define*

$$a = \log\left(\frac{1 - q}{1 - p}\right) \text{ and } \alpha = a + \log \frac{p}{q},$$

then

$$q < \frac{a}{\alpha} < p.$$

Proof: Consider the log likelihood ratio statistic $\hat{\ell}(m)$ for this simple hypothesis testing context: the observed process is $n(m) = \Sigma_{i=1}^{m} x_i$, where x_i are i.i.d. Bernoulli random variables with $P(x_i = 1)$ given by q and p under H_0 and H_1, respectively. Therefore,

$$\hat{\ell}(m) = \log \frac{b(n;m,p)}{b(n;m,q)} = \alpha n - am$$

and hence the mean of $\hat{\ell}(m)$ is $m(\alpha q - a)$ under H_0 and $m(\alpha p - a)$ under H_1. But as we noted in Section 6.2, $E\{\hat{\ell}\}$ is negative under H_0 and positive under H_1 in a general discrete setting.* Thus, we must have

$$\alpha q - a < 0$$

and

$$\alpha p - a > 0,$$

which completes the proof.

An immediate consequence of this lemma is that

$$(i)\ \ P_{1,0}(\phi,1) < \frac{a(\phi,r_0)}{\alpha(\phi,r_0)} < P_{1,0}(\phi,r_0)$$

and

$$(ii)\ \ P_{0,1}(\phi,r_0) < \frac{\tilde{b}(\phi,r_0)}{\tilde{\beta}(\phi,r_0)} < P_{0,1}(\phi,1),$$

where a and α are defined by eqs. (41a) and (41b) and $\tilde{b}$ and $\tilde{B}$ are defined by eqs. (45a) and (45b). Since $r \geqq r_0$ implies that $P_{1,0}(\phi,r) \geqq P_{1,0}(\phi,r_0)$ and $P_{0,1}(\phi,r) \leqq P_{0,1}(\phi,r_0)$, we see that

$$\text{sgn}(\alpha(\phi,r_0)P_{1,0}(\phi,r) - a(\phi,r_0)) = \text{sgn}(\tilde{b}(\phi,r_0) - P_{0,1}(\phi,r)\tilde{B}(\phi,r_0))$$

$$= \begin{cases} \text{positive} & \text{if } r \geqq r_0 \\ \text{negative} & \text{if } r = 1. \end{cases}$$

REFERENCES

1. A. Klimontowicz, "Grade of Service for Full Available Trunk Groups with Faulty Trunks," Proc. Sixth International Teletraffic Conference, Munich, 1970, pp. 212/1–212/6.
2. L. J. Forys and E. J. Messerli, "Analysis of Trunk Groups Containing Short-Holding Time Trunks," B.S.T.J., Vol. 54, No. 6, July–August, 1975, pp. 1127–1154.
3. E. J. Messerli, "Faulty Trunks: A New Look at an Old Problem," Bell Laboratories Record, June 1975, pp. 278–284.
4. W. S. Gifford and J. Shapiro, "Effect of the Change in Mean Holding Time Associated with an Equipment Irregularity on Network Trouble Detection and Customer Service," IEEE Trans. Commun., *COM-21*, No. 1 (January 1973), pp. 1–3.
5. R. E. Machol, Acquiring Data for Network Planning and Control," Bell Laboratories Record, October 1974, pp. 279–285.
6. R. B. Cooper, *Introduction to Queueing Theory*, New York: MacMillan, 1972.
7. E. Parzen, *Stochastic Processes*, Holden-Day, 1962.
8. A. Kuczura, "The Interrupted Poisson Process," B.S.T.J., *52*, No. 3 (March 1973), pp. 437–448.
9. V. B. Iverson, "Analysis of Traffic Process Based on Data Obtained by the Scanning Method, Proc. Seventh International Teletraffic Conference, Stockholm, 1973, pages 224/1–224/10.
10. E. L. Lehman, *Testing Statistical Hypotheses*. New York: John Wiley, 1959.
11. A. Wald, *Sequential Analysis*, New York: John Wiley, 1947.
12. R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.
13. W. Feller, *An Introduction to Probability Theory and Its Applications*, New York: John Wiley, 1957.
14. J. S. Kaufman, "Distribution of Busy Trunks in a Random Hunt Group Containing Killer Trunks," unpublished work.
15. V. E. Benes, *Mathematical Theory of Connecting Networks and Telephone Traffic*, New York: Academic, 1965.
16. Z. Govindarajulu, *Sequential Statistical Procedures*, New York: Academic, 1975.
17. D. R. Cox and H. D. Miller, *Theory of Stochastic Processes*, Wiley: Methuen, 1965.
18. D. R. Cox, Renewal Theory. Methuen; Wiley, 1962.