A SENSITIVITY STUDY OF TRAFFIC-PARAMETER ESTIMATION PROCEDURES USED FOR ENGINEERING TRUNK GROUPS

M. Berger and S. R. Neal
Bell Telephone Laboratories
Holmdel, New Jersey, USA

ABSTRACT

The Bell System is presently converting its basis for trunk engineering and trunk servicing to Wilkinson's Equivalent Random method. To use the method, the mean and the variance-to-mean ratio (peakedness) of the offered traffic must be estimated.

Using a combination of traffic data, simulation and mathematical analysis, we investigate the accuracy and sensitivity of the Equivalent Random method and various estimates of traffic peakedness to phenomena such as customer reattempts, day-to-day and intra-day variations in measured load and blocking, and trunk outage.

Our results show that significant estimation errors can occur when these phenomena are not properly accounted for. Where necessary, we indicate methods which can be used to remove resulting biases from the estimates.

1. INTRODUCTION

The accuracy of estimated trunk-group requirements depends heavily on the validity of the single-hour load-loss relations as well as the traffic-parameter estimation-procedures used in the engineering process. The mean and variance-to-mean ratio (peakedness) of the offered traffic are the traffic parameters required to implement the trunk engineering and servicing procedures which are based on Wilkinson's Equivalent Random method.

In this paper, we first study the errors which are inherent in the Equivalent Random method. For purposes of trunk estimation, we have found that such errors tend to be canceled out by certain errors caused by customer reattempts. Accordingly, we present a model for retrials and study the interactions through the use of a simulation. The results are contained in Section 2.

In Section 3 we consider four different estimates of traffic peakedness. The sensitivity of the estimates to day-to-day and intra-day variations in measured load and blocking is illustrated and an order of preference for the estimates is obtained.

Section 4 contains a discussion of the effects of trunk outage on estimates of peakedness as well as estimated load-loss relations. Techniques are presented for effectively removing the biases caused by trunk outage.

A summary of our results is given in Section 5. Throughout the paper we use the telephone-network terminology of Reference 1.

2. INTERACTIONS BETWEEN THE EQUIVALENT RANDOM METHOD AND CUSTOMER RETRIALS

The following is a description of the basic approximation made in the Equivalent Random (ER) method:

If a trunk group serves a stream of traffic which is composed of several overflow streams such that the mean and variance of the combined streams is \( \mu \) and \( \sigma^2 \) respectively, then the average call congestion seen by the several streams is the same as that which would be observed by a single overflow stream having mean \( \mu \) and variance \( \sigma^2 \).

In this section we first study the accuracy of the approximation and then show that for trunk-engineering purposes the errors resulting from the ER method are essentially cancelled out by the effects of customer reattempts.

2.1 THE ACCURACY OF THE ER METHOD

To investigate the accuracy of the ER method, we simulated a trunk group consisting of \( c \) trunks, \( 10 \leq c \leq 150 \), and an arrival process composed of \( N \) independent streams of overflow traffic each having specified mean and variance. The service times had a negative-exponential distribution and blocked calls were cleared from the system.
The basic idea for our test was to hold the arrival process fixed but obtain various levels of blocking (and a corresponding set of measurements) by changing only the number of trunks in the final trunk-group.* For each fixed value of \( c \), we simulated 20 hours of system operation. Using the resulting data and the ER method, we obtained 20 estimates of traffic peakedness which were then averaged (to obtain \( z^* \) as described in detail in Section 3). In each case, since we had complete knowledge of the arrival process, we could test the standard estimates of blocking and peakedness which are based on the ER method.

Because the arrival process was held constant, we expected that the estimates of peakedness would be about the same for each value of \( c \). However, the estimates actually changed as \( c \) changed (we varied \( c \) so that the blocking went from one to 50 percent).

We found that the ER method tends to overestimate blocking when the actual blocking is above ten percent and underestimate it when it is below ten percent. If follows that, when the ER method is used to estimate the peakedness \( z \) of the input traffic using measurements of call congestion and usage, the peakedness estimates \( z \) will be too small when the blocking is above ten percent and too large at the other extreme. Using the simulation we found that the results were not too sensitive to \( N \) provided \( N \geq 2 \). However, the relative errors tend to increase as the disparity of peakedness values for the input streams increases.

As an example we computed the biases analytically for \( N = 2 \) input streams. The results are shown in Figure 1 where the relative error in the peakedness estimates vs. blocking is plotted.

For purposes of trunk engineering, one is interested in the impact of these biases on estimated trunk requirements. The resulting trunking errors for \( z \) much larger than \( z \) are not severe since in this case the blocking is low and the system is performing adequately. However, at the higher values of blocking, one might conclude that the underestimates of \( z \) would lead to an underprovision of trunks.

Such a conclusion would be valid except that the problem becomes more complicated at higher blocking because of the increasing occurrence of customer reattempts. In the next section we define a model for customer reattempts and then discuss the interaction between the errors resulting from the reattempts and those inherent in the ER method.

### 2.2 Customer Reattempts in an Overflow Network

When a final trunk-group in an overflow network is underprovided, the traffic offered to the group consists of the actual first-attempt load together with a reattempt load. Since the reattempt load will vanish when the group is properly engineered, only the first-attempt load should be used to estimate trunk requirements. Therefore it may be necessary to include in the trunk-engineering procedures a method for estimating (from the usual traffic measurements) the first-attempt load on a final trunk-group.

To this end, consider the simple traffic-flow model shown in Figure 2.

---

*Our study did not consider the possibility of some significantly under-engineered high-usage groups.
Because it is very difficult to determine first-attempt loads analytically, the following discussion is based on
simple approximations which were used successfully in our study: Let $L_i$ denote the load carried by the $i$th
PHU group. Then the apparent (total) load offered to the
system is $L_i/(1-B_i)$, and since the load due to reattempts is $fr_0$, it follows that the first-attempt load is

$$a^* = L_i/(1-B_i) - fr_0. \quad (1)$$

A proportion $B_i$ of the calls that return to the $i$th PHU
group will again overflow onto the final so that the
reattempt load on the final is approximately

$$B_i fr_0 + B_i fr_0 \approx A \frac{1}{1-B} \text{ where } A = \frac{1}{1-B} (1)$$

Consequently the first-attempt load for the final is approximately

$$a_f = \frac{1}{1-B} R f r_0 + \frac{1}{1-B} R f r_0 \approx \frac{1}{1-B} (1)$$

If one of the traffic streams is first routed (not over­
flow) we merely set the corresponding blocking probability $B_i = 1$ in (2).

These results are easily extended to include any number
of PHU groups as well as intermediate high-usage (IHU)
groups. However such corrections may be inconvenient to
use because of the difficulty in obtaining the
ratios $r_i$.

For such cases, let us assume that the trunk groups in
the subtending network function as a single group, and
use $r$ to denote the proportion of the load on the final
which is first-routed (not overflow). The resulting
estimate of first-attempt load is

$$a_f = \frac{1}{1-B} [R f (1+r)] \frac{1}{1-B} \quad (2)$$

where $\beta$ denotes a composite blocking for the subtending
network. The blocking $\beta$ is a function of the number of
network stages and the economic COS (ECCS) criterion
used for engineering the high-usage groups. The following
were found by simulation to be satisfactory values for $\beta$:

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>2-Stage Network</th>
<th>3-Stage Network</th>
<th>4-Stage Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toll Network</td>
<td>0.3</td>
<td>0.25</td>
<td>0.2</td>
</tr>
<tr>
<td>Local Network</td>
<td>0.15</td>
<td>0.1</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Although the estimate shown in (3) appears simple, it
does require the ratio $r$ which may not be readily
available. In such a case, an average value might be
used. A limiting case is $r = 1$ which is correct for
full-direct groups and results in a third estimate

$$a_f = \frac{1}{1-B} [R f (1+r)] \frac{1}{1-B} \quad (3)$$

Essentially (2) and (3) are generalizations of (4) which
account for the fact that some of the reattempts will be
carried by the subtending network. The estimate (4)
assumes that all reattempts return to the final group.

### 2.3 INTERACTIONS

To observe the interactions between the errors inherent in
the ER method and those caused by customer reattempts,
we modified the simulation used earlier so as to include
customer reattempts.

#### 2.3.1 RETRIAL MODEL

There are three basic features in the retrial model used in
our simulation.

1. A customer who is blocked on his $n$th attempt
subsequently reattempts with probability $P_n$ (and leaves the system with probability $1 - P_n$). Based on a study by Wilkinson and Radnik\(1\), we used $P_1 = .5, P_2 = .65, P_3 = .75, P_4 = .8$ and $P_n = .85$ for $n \geq 5$.

2. A fixed proportion of the reattempts (we used ten percent) occurs after a short, fixed
waiting time (we used ten seconds).

3. The remaining reattempts occur after waiting a
random length of time which has a negative­
exponential distribution. The mean waiting­
time was set at one mean holding-time, i.e.,
200 seconds.

### 2.3.2 SIMULATION RESULTS

Using the simulation, we investigated the relative accuracy
of (3) and (4) when the estimates were used with the ER
method to predict trunk requirements. We then compared the
results with those obtained when no adjustment was made for
reattempts (i.e., $\beta = 0$).

Twelve different networks, representative of local and toll
configurations, were studied. The blocking on each final
group was varied from one percent to 55 percent by changing the
number of trunks in the final group. A summary of the
relative trunking-errors which occurred at about 20 percent
blocking on the final trunk groups is shown in Figure 3.

Figure 3: Relative trunking-errors for various load
estimates at 20 percent blocking on the final
group.

The results are of a similar nature for other values of
blocking although the spread in the results increases as the
blocking increases. In general we found that the
maximum relative trunking-errors did not exceed the
corresponding blocking probabilities.

As was expected, neglecting the effect of retrials tends
to cause an overprovision of trunks. However, the resulting
errors are reduced considerably by the bias in the ER
method at high blocking (see Section 2.1). In fact, in
some of the cases, the errors in the ER method dominate and
cause undertrunking even though reattempts were present in
the measurements.

When the estimate (3) was used, the trunking errors were
almost unbiased although both undertrunking and over­
trunking did occur. Finally, use of (4) together with the
ER method almost always resulted in undertrunking.

### 2.4 TRAFFIC DATA

To substantiate our conclusions concerning the practical
accuracy of the ER method, we have compared the blocking
probability predicted by the ER method with that actually
measured on a number of trunk groups in several overflow
networks. Typical results are shown in Figure 7 and
discussed in Section 4.1.2.
2.5 DISCUSSION

Our studies show that there are errors inherent in the ER method. However, for trunk-engineering purposes the errors are essentially cancelled out by the effects of customer reattempts. The accuracy of the ER method can be improved slightly by modifying the estimates of offered load to account for reattempts.

We conclude that when sufficient information is available, (3) provides a good estimate of first-attempt load. However, when it is not practical to use (3), a reasonable second choice is simply to make no adjustment for reattempts (i.e., set $f = 0$) keeping in mind that the trunking errors resulting from unreflected reattempts are partially compensated for by the biases in the ER method.

3. ESTIMATION OF TRAFFIC PEAKEDNESS

We consider four ways to estimate traffic peakedness and illustrate the sensitivities of the estimates to variations in load and blocking.

3.1 THE ESTIMATES

In the Bell System, estimates of traffic parameters are normally based on 20 single-hour measurements made during a time-consistent busy-season busy-hour. One has the choice of first averaging the 20 measurements and then estimating the parameters, or first obtaining 20 single-hour estimates of each parameter and then averaging the values in some fashion. When traffic peakedness is desired, there are two additional possibilities:

The peakedness may be estimated directly from the measured usage and blocking on the trunk group serving the traffic. The estimate is obtained by iteration using the ER method to find that value of peakedness which would give rise to the observed values of carried load and blocking. In this case we call the estimate the measured peakedness.

A second estimate is obtained from the subtending network. The mean and variance of each stream of traffic submitted to the trunk group in question is computed by using the ER theory. The mean and variance of the superposition of all the input streams are then estimated by summing the individual means and variances. The ratio of the total variance to total mean is an estimate of peakedness which we call the peakedness from below.

One can first average the daily measurements and then compute the two estimates of peakedness, or first compute 20 values of each estimate and then use a weighted average for each estimate. We use $z$ and $z_B$ to denote respectively the measured peakedness and peakedness from below when the averages of the measurements are used. Similarly, $z^*$ and $z^B$ are respectively the weighted averages of the 20 single-hour estimates. That is, if $(a_i z_i), i = 1, \ldots, 20$, are 20 values of measured load and peakedness, then

$$z = \frac{1}{20} \sum a_i z_i$$

To compare the four estimates, we used data from a number of trunk groups. Typical results are shown in Table 1 for 21 trunk groups originating from a No. 5 Crossbar office. (In this office, maintenance outage was not included in the usage measurements.) The table is structured so that for the local network, all of the IHU groups shown overflow to the final trunk group. However, there were more than one hundred FHU groups in the local network of which only six typical cases are shown. For the toll network, the subtending network for the final group contained only five FHU groups and one IHU group, all of which are shown.

The table contains the average busy-hour blocking, the average offered load, the trunks intended and the average traffic service. Since there was a considerable amount of trunk outage during the study period, the four peakedness estimates were computed twice for each group; once using the trunks intended and once using the average number in service.

<table>
<thead>
<tr>
<th>PEAKEDNESS ESTIMATES</th>
<th>WITH TRUNKS INTENDED</th>
<th>WITH AVERAGE NUMBER IN SERVICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.1 PEAKEDNESS FROM BELOW</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Except for the trunk group IHU-2 in the local network and the two final groups, there are no significant differences between $z_B$ and $z$. Notice that for the group IHU-2, $z_B$ is larger than $z$, but, for the two final groups, $z_B$ is smaller than $z$. Hence, the difference between the two estimates, which reflects the effects of day-to-day variations in the offered loads, can be either positive or negative. However, the estimate $z_B$, which is computationally easier to obtain, has about the same accuracy as $z$. Thus, it appears that the average data can be used to estimate $z_B$ accurately without any correction for day-to-day variations in the offered loads.

3.1.2 MEASURED PEAKEDNESS |
| There are several groups with significant differences between $z$ and $z^*$. For example, for the final group in the local network, the estimates $z$ and $z^*$ are 4.24 and 6.93, respectively. Notice that the difference between $z$ and $z^*$ is positive for some groups and negative for others.

Although the estimate $z$ is computationally easier to obtain, it tends to be biased by day-to-day variations in the offered loads. Hence, when $z$ is used for trunk engineering, it should be adjusted for day-to-day variations. Such a procedure is presently in use in the...
Bell System but the results are not entirely satisfactory.

3.2 THE VARIABILITY OF THE PEAKEDNESS ESTIMATES

3.2.1 DAY-TO-DAY VARIATIONS

During our study we observed that there is considerably more variation in the hourly values of measured peakedness than in the corresponding values of peakedness from below. As an example, in Figure 4 we compare 20 time-consistent busy-hour values of measured peakedness with the corresponding estimates of peakedness from below for an IHU group with an average of 30.7 trunks in service and 32 intended. The trunk outage was negligible in the subtending PHU groups.

Figure 4: Hourly measured peakedness vs. hourly
peakedness from below for an IHU trunk group.

The coefficient of variation of the measured peakedness is 25.2 percent but is only 5.6 percent for the peakedness from below. The large variability in the hourly measured peakedness contributes to the observed differences between $z$ and $z_B$ shown in Table 1.

3.2.2 HOUR-TO-HOUR VARIATION

In the preceding section, we saw that the day-to-day variation in the peakedness from below is considerably smaller than the corresponding variation in measured peakedness. We have also found the peakedness from below to be more stable from hour-to-hour within a single day.

Typical results are displayed in Figure 5 where we have compared the 20-day averages $z^*$ vs. $z_B$ for each of four busy-hours within the same day.

Figure 5: Average of the hourly measured peakedness vs.
the average of the hourly peakedness from below.

Notice that $z_B$ has a coefficient of variation of only 2.1 percent while the variation in $z^*$ is 23.7 percent. Thus, our data imply that the peakedness from below is less sensitive to intra-day as well as day-to-day variations in the measured load and blocking.

3.3 DISCUSSION

Of the four estimates of peakedness we considered, our results indicate that $z_B$ and $z^*$ are the most reliable. In contrast with $z$ and $z^*$, they are relatively insensitive to day-to-day variations in the measured loads and blocking. We also found $z_B$ to be just as accurate as $z^*$.

In order to use $z_B$ rather than $z$, accurate data for the entire subtending network are required. When it is not feasible to obtain such information, the measured peakedness $z$ or $z^*$ can be used. If $z$ is used, a correction is required to remove the effect of day-to-days in the offered loads. A completely satisfactory correction is not presently available.

4. TRUNK OUTAGE

Whenever traffic measurements are made on a trunk group, it is highly probable that some of the trunks will be out of service. When a trunk is out of service, it is normally "made-busy" and thereby adds one erlang of maintenance outage (sometimes called maintenance usage) reflecting trunk outage in the estimates.

On the other hand, some telephone companies remove the maintenance outage from usage measurements. Also, in many No. 5 Crossbar offices, standard wiring arrangements are not provided to include maintenance outage in usage measurements.

In this section we first illustrate some of the effects of trunk outage on estimates of load-loss relations and traffic peakedness. We then consider methods for reflecting trunk outage in the estimates.
4.1 EFFECTS OF TRUNK OUTAGE ON LOAD-LOSS RELATIONS

4.1.1 PRIMARY HIGH USAGE GROUPS

The scatter plot in Figure 6 represents the observed relation between the single-hour measured load and blocking for a primary high-usage trunk group (the input is normally assumed to be Poisson) having 14 trunks intended and an average of 10.3 trunks in service. The figure also shows four different theoretical load-loss curves which one could associate with the data.

The solid curve (c = 14, z = 1) is the Erlang B relation for 14 trunks and is negatively biased because of the trunk outage. If one assumes that all 14 trunks are working and estimates the peakedness of the input traffic for each data point, then the weighted-average peakedness \( Z^* = 3.21 \). The curve corresponding to \( c = 14 \) and \( z^* = 3.21 \) fits the data somewhat better than the Erlang B relation although it has a small positive bias at the lower end. The point here is that the increased blocking due to trunk outage could be mistakenly attributed to a peakedness greater than one for the incoming traffic \( z = 1 \) would be correct for Poisson input.

The two best fits to the data were achieved by using the average number of trunks in service. When that average was used in the estimation of peakedness, a value close to one \( (Z^* = 1.13) \) resulted. An estimate even closer to one would probably have been obtained had it been possible to use the hourly number of trunks in service.

4.1.2 INTERMEDIATE HIGH USAGE GROUPS

We found similar results on trunk groups which serve overflow traffic. That is, biases in load-loss relations could be effectively removed by using only the average number of trunks in service. As an example, consider Figure 7 where measured blocking versus theoretical blocking is plotted from data taken on an intermediate high-usage group with 32 trunks intended and an average of 30.7 trunks in service.

The theoretical blocking was obtained by using the EN method with \( c = 30.7 \) and data from the subending network (where trunk outage was negligible during the study period).

As one would expect, when the trunk outage was not reflected in the computations, the theoretical blocking contained a negative bias (not shown). Again we note that the bias was essentially removed by using the average number of trunks in service.

4.2 EFFECTS OF TRUNK OUTAGE ON ESTIMATES OF PEAKEDNESS

4.2.1 MEASURED PEAKEDNESS

As noted above, an increase in measured blocking due to trunk outage can be mistakenly attributed to an increase in traffic peakedness. In this section we indicate how the bias in peakedness estimates behaves as a function of outage and then discuss the following alternatives for accounting for outage in the estimates of peakedness:

(i) Include the maintenance outage in usage measurements and use the intended number of trunks in the computations.

(ii) Exclude the maintenance outage from usage measurements and use the average number of working trunks in all computations.

(iii) Exclude the maintenance outage from usage measurements and use the actual number of working trunks in all computations.

Figure 8 indicates how the measured peakedness increases when the number of trunks out of service increases, the usage measurements do not include maintenance outage, and the number of trunks intended is used in the computations. For instance, if 14 trunks were intended to be in service and the offered load is 32 erlangs, then the measured peakedness increases from 1.8 to 7.2 when ten trunks are out of service (but 14 is used in the computations) and the usage measurements do not include maintenance outage.
However, if maintenance outage is included in usage measurements, the measured peakedness does not change significantly as a function of trunk outage. Thus, alternative (i) is an effective way to remove the bias in measured peakedness caused by trunk outage.

The results in Section 4.1.1 indicate that alternative (ii) will furnish adequate estimates of peakedness provided the variance of the number of trunks out of service is not too large. Hence, if system constraints preclude the use of (i), alternative (ii) is a reasonable second choice. Of course, (iii) is the most accurate of the three but is expensive to use since hourly checks must be made on the number of trunks in service. Moreover, (iii) furnishes only a slight increase in accuracy over (i).

4.2.2 PEAKEDNESS FROM BELOW

Throughout our data studies, we found the peakedness from below to be rather insensitive to trunk outage in the subtending networks. As an example, observe the values of $z_B$ and $z_{\tilde{B}}$ in Table 1 for the final trunk-groups. Although there was substantial outage in the subtending network, the peakedness from below based on average trunks in service agrees closely with that based on the intended number of trunks.

4.3 DISCUSSION

The results of this section imply that using the average number of trunks in service in the ER computations will provide adequate estimates of load-loss relations when maintenance outage is not included in usage measurements.

Trunk outage can cause significant biases in measured peakedness. However, the biases can be almost completely removed by including maintenance outage in usage measurements. An alternative approach is to use the average number of trunks in service in the computations.

In contrast to measured peakedness, the peakedness from below is relatively insensitive to the trunk outages which we found to occur in high-usage trunk groups.

5. SUMMARY AND CONCLUSIONS

5.1 ER METHOD

We considered several aspects of the ER method. Although there are biases inherent in the method, in practice they are essentially cancelled out by the effects of customer reattempts. We presented a method which accounts for customer reattempts and thereby increases the accuracy of trunk-engineering based on the ER method.

5.2 TRAFFIC PEAKEDNESS

We presented four possible estimates ($z, z^*, z_B, z_{\tilde{B}}$) of traffic peakedness and studied their sensitivities to trunk outage and to day-to-day and intra-day variations in offered loads.

The measured peakedness $z$ is much more sensitive than the peakedness from below $z_B$ to trunk outage and variations in offered loads. In fact, when there is a significant relative difference, say 100 percent, between $z$ and $z_B$, the trunk engineer can reasonably conclude that trunk outages are not being treated in a consistent manner or that significant measurement errors are present. We have provided guidelines for properly accounting for trunk outages.

Our preferred estimate of peakedness is $z_{\tilde{B}}$. It is just as accurate as $z_B$ but easier to compute. However, data from the subtending network are required to compute $z_B$.

If such data are not available and the subtending network is not being changed, either $z$ or $z^*$ for measured peakedness can be used to estimate peakedness. The estimate $z$ is easier to compute but day-to-day variations in the offered load cause the estimate to be biased. A completely satisfactory estimate of the bias is not presently available.

References

