TRUNKING CONSIDERATIONS—CHARACTERISTICS OF TRAFFIC AFFECTING TRUNK FORECASTING AND SERVICING

GENERAL ENGINEERING CONCEPTS

NETWORK OPERATIONS METHODS

CONTENTS PAGE

Figures

CONTENTS PAGE

1. GENERAL

1.01 Trunk configurations range in complexity from a single nonalternate group to the total North American network. To engineer and administer them effectively requires recognizing and dealing with a variety of traffic characteristics. It is, therefore, essential that the trunk engineer have knowledge of these characteristics and the ability to evaluate the effects and factors (both external and system induced) which affect them. This section in combination with Section 780-400-111, which deals with traffic characteristics in general, will provide the necessary background.

1.02 Whenever this section is reissued, the reason for reissue will be given in this paragraph.

1.03 References in this section to methods, service levels, and equipment quantities are based on the American Telephone and Telegraph Company recommendations.

1.04 For the standard meaning of terms and definitions used in this section, see Section 780-400-305, "Glossary of Trunk Facilities Terms and Definitions."

1.05 As used in this section, the term "traffic" denotes the flow of messages through a communications system. The characteristics of traffic are determined partly by the behavior of telephone subscribers and partly by properties of the telephone system; and various aspects of customer behavior are influenced to a greater or lesser degree by system characteristics. Better appreciation of this interaction between traffic and the telephone network allows the trunk engineer to better plan for satisfactory and economical service.

1.06 All actual service systems have finite capacities, and the telephone network is no exception. Thus the trunk engineer cannot arrange for the system to serve unlimited potential demands. Service objectives based on realistic economies introduce blocking by limiting the number of trunks, just as all other components of the system are also limited in number. Traffic characteristics are influenced by network configuration, grade-of-service objectives, tariffs, equipment design, engineering methods, and other factors which influence customers' behavior. Not all varieties of traffic, as distinguished by their variability, are equally desirable from the viewpoint of efficient and equitable utilization of telephone facilities, and the trunk engineer must keep in mind that traffic is affected by changes in all of the realms just mentioned. In particular, as illustrated in many of the following paragraphs, measures which reduce the ***variability*** of offered loads are nearly always desirable.

2. VARIABILITY OF OFFERED LOAD

A. Poisson Traffic

2.01 During a time-interval when there is no trend in the calling rate, usually the case for the busy hour, the traffic offered by telephone subscribers to a central office, and in many cases to a trunk group, has been found to fit the mathematical concept known as the "Poisson process." A Poisson traffic input process is characterized by three conditions:

(1) The calls arrive one at a time.

(2) The number of calls arriving in any given time interval is independent of the numbers arriving in any collection of time intervals which do not overlap the given interval.

(3) The distribution of the number of calls arriving during any time interval depends only on the length of the interval.

Every arrival process which satisfies these conditions is a Poisson process and vice versa. It follows from these conditions that the average number of calls arriving during any time period is proportional to the length of the period and the probability distribuition of the number of arriving calls in the Poisson distribution.

2.02 Traffic offered as a Poisson process, "Poisson traffic" for short, may also be called "purely random." (In some Bell System and other literature it has been called simply "random"; but this conflicts with the general usage that "random" refers to ***any*** traffic that is not wholly predictable or deterministic.) The "randomness" of a Poisson process is illustrated by the fact that if it is known that a call, which arises in such a process, arrives during some time interval, then it is just as likely to arrive during any part of the interval as during any other part of equal length. In Fig. 1, a typical distribution pattern of Poisson traffic is illustrated by showing the number of simultaneous calls in progress during an hour on a trunk group with no blocking encountered. The quantity of traffic, or "load" as it is more commonly referred to, is defined as the average number of calls per average call holding time; it is also the average number of servers that could be kept fully occupied by the incoming calls if there were no blocking. Therefore in Fig. 1A, the average offered load is equal to the average number of simultaneous calls (which is equivalent to the average number of occupied servers if no blocking is encountered), which happens to be 9.5. There are two units in general use for measuring the traffic intensity or load. One of these, which is most common in engineering practice abroad, is the "erlang." From the above definition, if 190 calls with an average holding time of 180 seconds (0.05 hour) are offered in an hour to a trunk group, the total trunk group offered load would be (0.05) (190) = 9.5 erlangs. In Bell System traffic engineering, loads are usually expressed in "CCS"; one CCS is 100 call-seconds per hour. One erlang is therefore 36 CCS. (For the above example the offered load would be 342 CCS.)

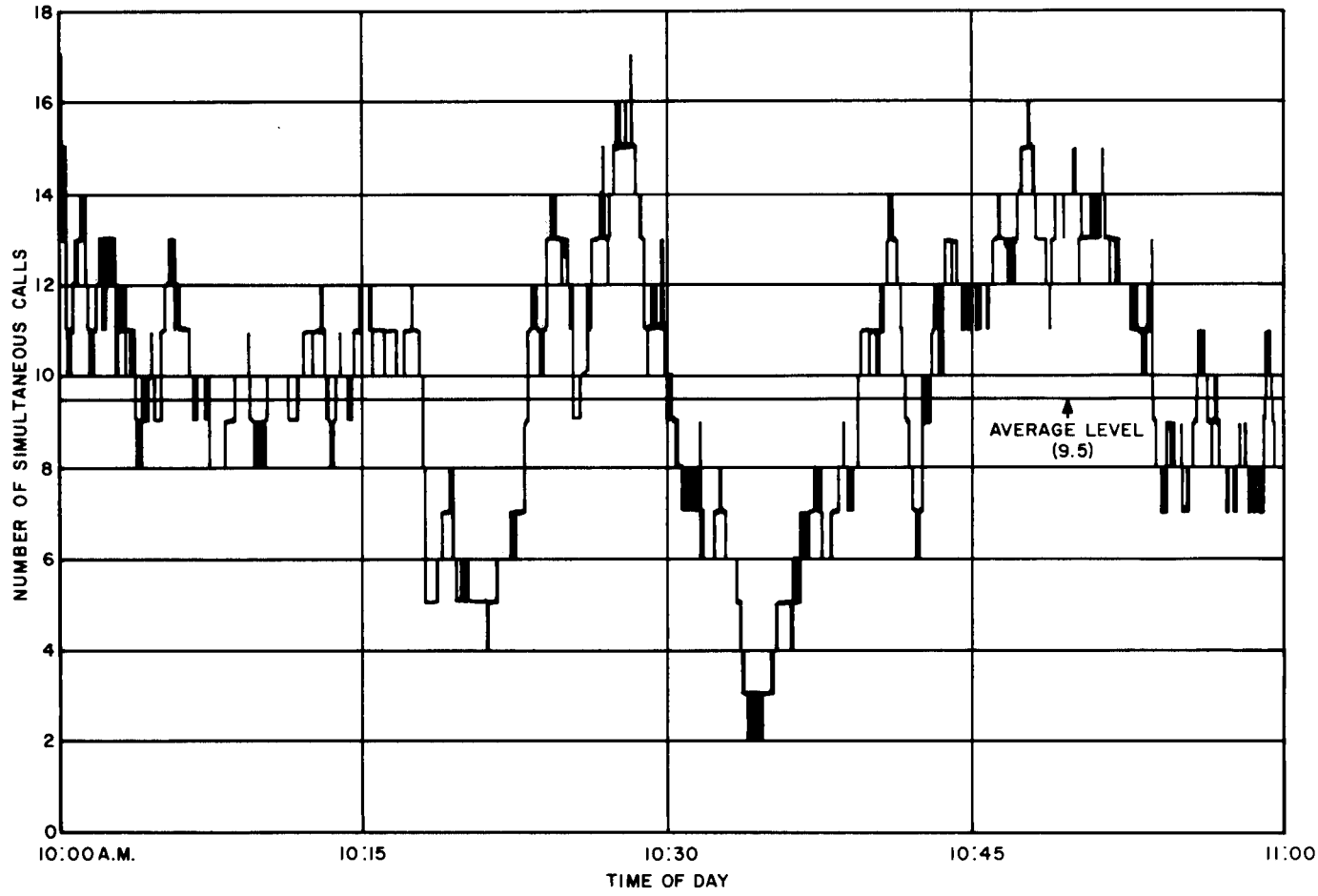

Fig. 1 — Poisson Traffic

B. Non-Poisson Traffic

2.03 In the Bell System, traffic distributions which do not possess Poisson characteristics have been referred to as "nonrandom." This use of the term is not meant to imply that such traffic is scheduled or predictable in detail, but merely that it is non-Poisson or not ***purely random.*** Non-Poisson traffic may occur naturally because of the calling habits of subscribers. Traffic load fluctuates regardless of the time interval in which it is viewed. Load can vary within the hour, the Bell System standard measurement interval, from day-to-day in the same clock hour (day-to-day variation) and from season to season.

2.04 Non-Poisson traffic distributions can also occur as a result of the network itself; in particular, these distributions are induced by the alternate route trunking arrangements. If the load offered to a high-usage trunk group is Poisson, the load overflowing this group is peaked, (see paragraph 2.07), since the trunk group carries all but the extreme points in the traffic distribution. Since the overflow load consists only of the extreme points that occur sporadically, it has greater instant-to-instant variability than the Poisson.

2.05 ***Peaked traffic*** is illustrated in Fig. 2 by examining the portion of the random offered load in Fig. 1 that would ***overflow*** (not be carried on) a 12-trunk group. Figure 2 shows the number of simultaneous calls that would be carried on an overflow path. From Fig. 2, it can be seen that there are more occurrences of large numbers of calls, and also longer intervals when few or no calls are present than in the Poisson distribution illustrated in Fig. 1.

2.06 Another type of non-Poisson traffic encountered is "smooth" traffic, which is less variable than Poisson traffic. One example of ***smooth***

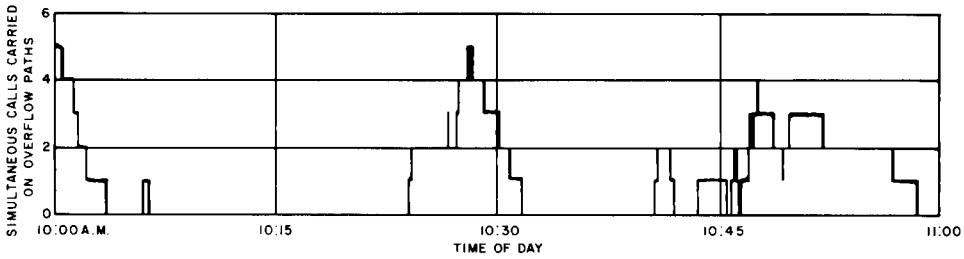

Fig. 2—Peaked Traffic

traffic is illustrated in Fig. 3 by examining the portion of the Poisson offered load in Fig. 1 that would be ***carried*** if that traffic were offered to a 12-trunk group with blocked calls cleared (see the discussion below on retrials).

2.07 A measure of the variability in the offered traffic load is ***peakedness,*** defined as the variance of the number of simultaneous calls divided by the mean number of calls, assuming all calls can be served when offered. Poisson traffic has a peakedness of one. Overflow traffic consisting essentially of the peaks clipped off the top of a random traffic distribution (see Fig. 2) tends to come in separated bunches and has a peakedness greater than one, commonly ranging up to about four. Carried traffic, as illustrated in Fig. 3, results from clipping a Poisson distribution and hence has a peakedness value less than one.

2.08 To illustrate the effect on blocking of peaked versus Poisson traffic, an observed distribution of overflow calls from eight high-usage groups offered to a final group has been plotted in Fig. 4. The Poisson distribution having approximately the same number of average calls is indicated by the dashed line. From a service standpoint, the area of interest is restricted to the extreme right-hand portion of the distribution, which is concerned with the proportion of time during which such a large number of calls are present that blocking or NC ("no circuit") might result. The Poisson assumption (dashed line) suggests that if 14 trunks were provided, virtually no blocking would occur. But given the actual peaked distribution, it is apparent that significant blocking would be encountered since more than 14 simultaneous calls did occur. At times, the number of observed calls is as high as 18.

2.09 The number of trunks required to attain a specified value of blocking increases as the within-the-hour distribution of offered load changes from smooth to Poisson to peaked. Since the load

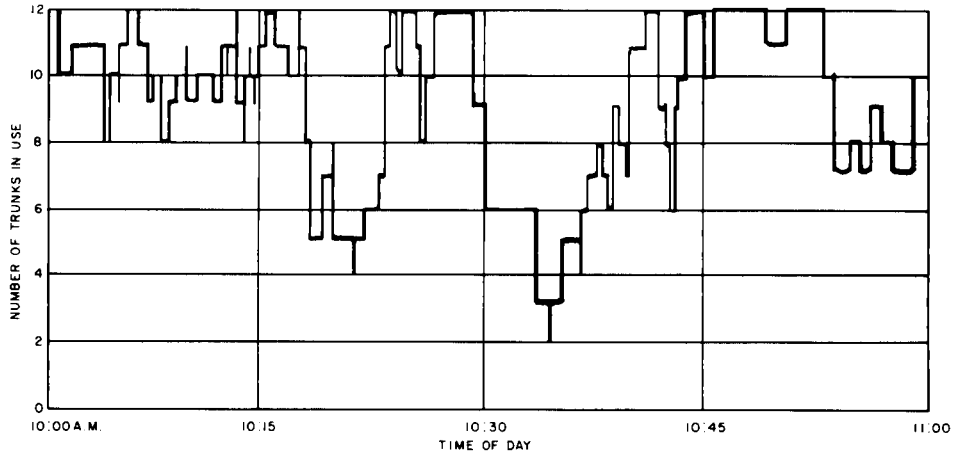

Fig. 3—Smooth Traffic

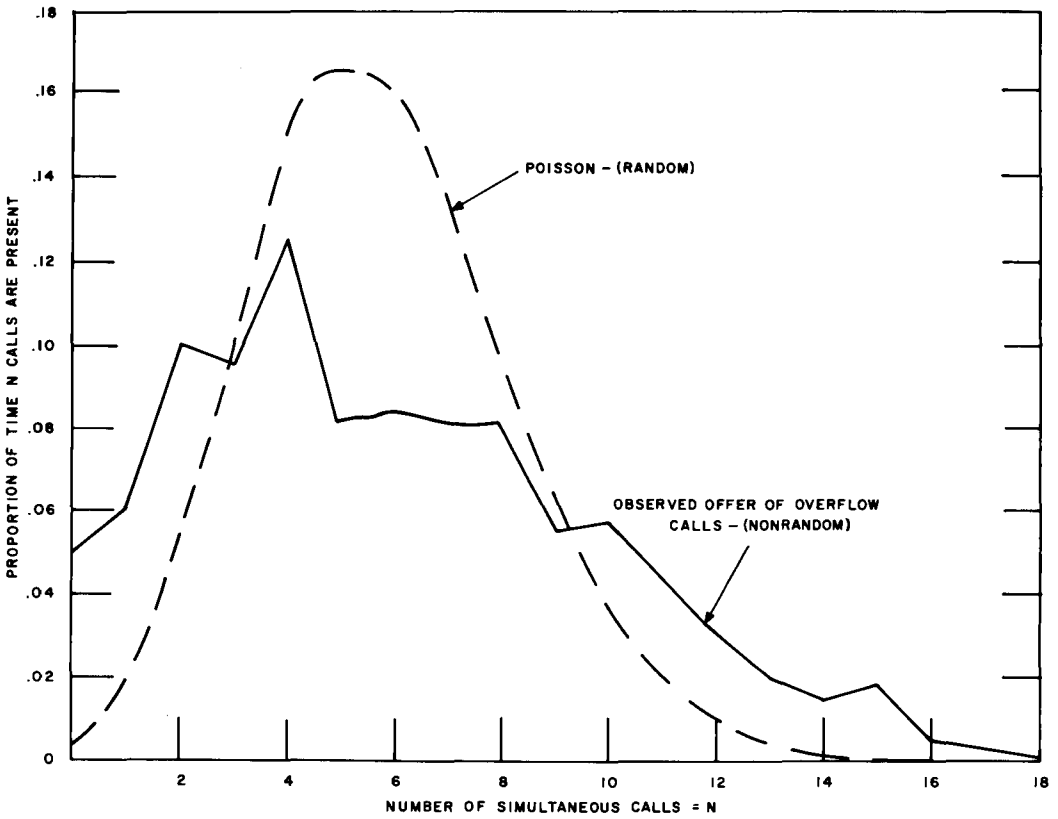

Fig. 4—Comparative Poisson Versus Non-Poisson Distribution of Calls—Given Same Average Offered Load

carried on each trunk decreases, the number of high-usage trunks that can be economically justified, based on the load carried on the last trunk, decreases as the offered traffic becomes more peaked. (This results because, for a given number of trunks, the load carried by the last trunk in the group decreases as the offered traffic becomes more peaked.) The methods recommended for recognizing and coping with peaked traffic are covered in Section 780-400-340.

3. EXTERNAL FACTORS

A. Within-the-Hour Variations

3.01 Underlying most trunking formulas are two key assumptions: that traffic is "steady" (or in more mathematical terms, statistical equilibrium exists) and that engineering is based on hourly average values of blocking. Consider again the Poisson traffic stream depicted in Fig. 1. The phenomenon observed during this one-hour interval is the simultaneous number of calls in progress. If we concentrate on the general trend of the flow of traffic in this hour, it is apparent that the short-term fluctuations rise and fall with some constancy about the average level of 9.5 erlangs. It is also apparent that this average is not changing during this interval. For this condition to exist, on the average the calls arriving and departing from the trunk group throughout the hour must be equal. Such a condition constitutes ***statistical equilibrium.*** The only variations which occur in statistical equilibrium are the fluctuations, such as those shown in Fig. 1, which are an inherent characteristic of the random process representing offered load. These random fluctuations are based on the assumption that subscribers originate calls independently of other subscribers.

3.02 If the rate at which the traffic load is offered to a trunk group is not constant but instead changes during an hour, the load is said to be "time-varying" or "nonsteady". In the Bell System, nonsteady within-the-hour traffic distributions are often referred to as being "skewed". Trunking formulas, which form the basis for standard methods of trunk engineering, will call for too few (that is, will "underprovide") trunks, if used in the presence of traffic which has significant within-the-hour

variation. This is due to the shape of the curve showing blocking as a function of offered load in the usual range of blocking. As shown in Fig. 5, the average value of blocking levels associated with a set of differing loads is higher than the blocking corresponding to the average of these loads. The blocking associated with the average, a, of two hourly loads a_1 and a_2 is smaller than the average of the blocking ratios in each hour, $B = (B_1 + B_2)/2$.

3.03 There are circumstances under which traffic is nonsteady (skewed) because of significant reductions in toll tariffs at a particular clock hour. In this situation, the provision of more trunks permits time-variation to appear in the carried load. The number of customers wishing to place calls in the first few minutes of an hour may be extremely high. If trunks are limited, the customers are forced to spread their calls out over a longer time span. As trunks are added, however, the number of calls handled on the initial attempt increases, but because of the suppressed demand the busy-hour service may still remain below standard.

3.04 Systematic or predictable within-the-hour variations have the same general effect that systematic day-to-day variations in load have on load-service relationships (which are discussed below). In the case of systematic day-to-day variations, however, the data collection and processing methods lend themselves to identifying the variance of a collection of daily single-hour loads. To cope with systematic within-the-hour variations in a similar manner would require collecting and processing data at intervals of less than one hour.

3.05 As mentioned previously, existing trunk engineering methods are based on grade-of-service criteria stated in terms of hourly average values of blocking. Such methods do not directly account for within-the-hour variations or skewness. However, when the peakedness of the busy-hour traffic offered to a particular trunk group is calculated from the number of trunks and the observed busy-hour load service relationship, the presence of within-the-hour variation inflates the peakedness factor. The phenomenon can be dealt with by utilizing empirically derived peakedness factors in engineering the trunk group.

3.06 When an observed-load service relationship on a trunk group cannot be reasonably explained by the best fitting trunking theory, taking into consideration the effects of day-to-day variation, peakedness, and customer retrials (see Part 4) as appropriate, the trunk engineer may suspect skewness. To verify that skewness is present, half-hourly readings are normally taken to determine whether there is a significant difference between the two half-hourly readings within an hour. As

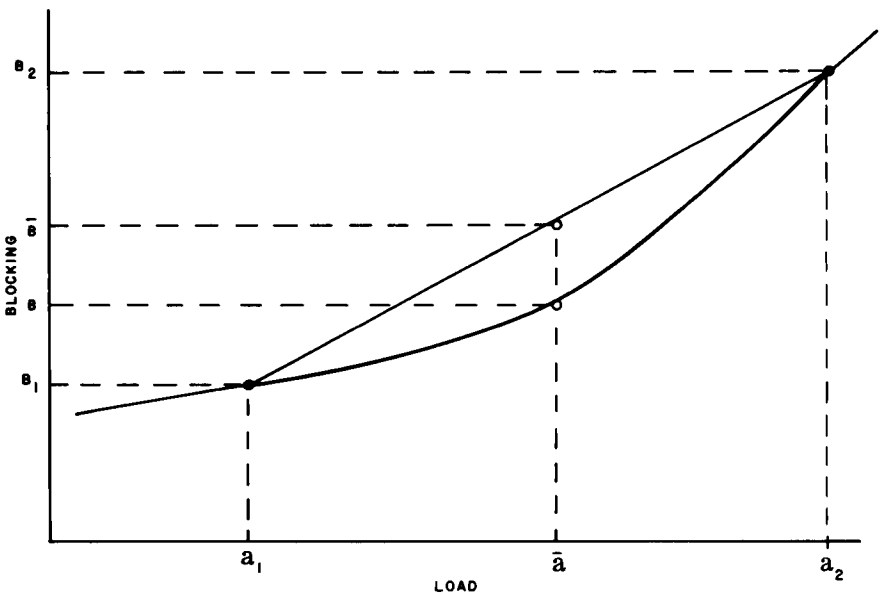

Fig. 5—Load-Service Curve

a rough rule of thumb, within-the-hour variation of the offered load is considered to be present when the hourly blocking is 5 percent or more and the blocking in one half-hour is at least double the blocking in the other half-hour. A more detailed discussion of the service considerations of within-the-hour variations can be found in Section 780-400-325 and a discussion of the effects of these variations on trunk group capacity can be found in Section 780-400-340.

B. Day-to-Day Variation

3.07 The load offered to a trunk group in any given hour is a sample which fluctuates around the "source" load; that is, the average load that would be offered in the long run if the traffic input process continued indefinitely under statistically identical conditions. However, the ***average*** load-service relationship observed on a trunk group for a series of hours will differ from that predicted by a "constant" load theory.

3.08 The source load on a trunk group will normally vary by hour of the day, the day of the week, and the month of the year. Thus, one frequently finds that the busy-season, busy-hour load on Mondays is significantly different from the busy-season, busy-hour load on Tuesdays, Wednesdays, etc. Figure 6 illustrates day-to-day variation by examining the distribution of the busy hour offered loads over a 20-day period for a primary high-usage group with nine trunks.

3.09 If there is appreciable day-to-day variation in the loads offered to a trunk group, the use of trunking theories with a constant load assumption to interpret the average of a series of busy-hour loads can result in a misstatement of offered loads, overflow loads, blocking, and trunk requirements.

3.10 To estimate the blocking which will occur in a single busy hour, two traffic parameters are required: offered traffic load and a measure of the within-the-hour variation (peakedness). However, to predict the busy season average blocking, an additional parameter is required: the variation in the time-consistent busy-hour load from day-to-day.

3.11 For many years, the effects of day-to-day variation (and customer retrials) were taken into account by using the Poisson blocking formula (or blocked-calls-held model). This formula is also based on Poisson arrivals, but predicts a level of blocking that is higher than predicted in a blocked-calls-cleared system or the Erlang B formula. The Poisson blocking model was used since empirical studies indicated that the blocking levels predicted by this formula more closely matched observed load-service relationships than other existing formulas. The Poisson formula merely accounted roughly for effects which were not represented in the Erlang B model.

3.12 Modern methods of collecting, storing, and processing traffic data have permitted us to find not only the average busy-hour load but also the variance of a collection of daily single-hour loads. New trunk capacity tables, the Neal-Wilkinson tables, which use the three parameters, offered load, peakedness, and day-to-day variation, have been developed at Bell Laboratories and are now in use for engineering grade-of-service trunk groups. A description of the assumptions underlying these tables and of their use is given in Section 780-402-210. A more detailed discussion of traffic theories is presented in Section 780-400-140.

3.13 In addition to the new trunk capacity tables which include day-to-day variation parameters, special methods for adjusting offered loads on high-usage groups in order to account for this variation are available. These procedures are discussed in detail in Section 780-400-335.

C. Holding Times

3.14 The holding times of customers' calls ordinarily tend to form an exponential distribution. The probability of blocking on a trunk group depends only on the offered load and is not affected by the distribution of holding times. However, when certain measurements are not taken or data are missing, as in the case when loads must be estimated from counts of offered or carried calls, the average holding time must be known. (The distribution of the holding time also determines the delay distribution in service systems in which blocked calls are delayed, instead of being cleared. However, trunk forecasters and servicers are not generally concerned with traffic engineering of delay service systems.)

3.15 Just as the effect of a tariff may be to produce a time-varying load because of a

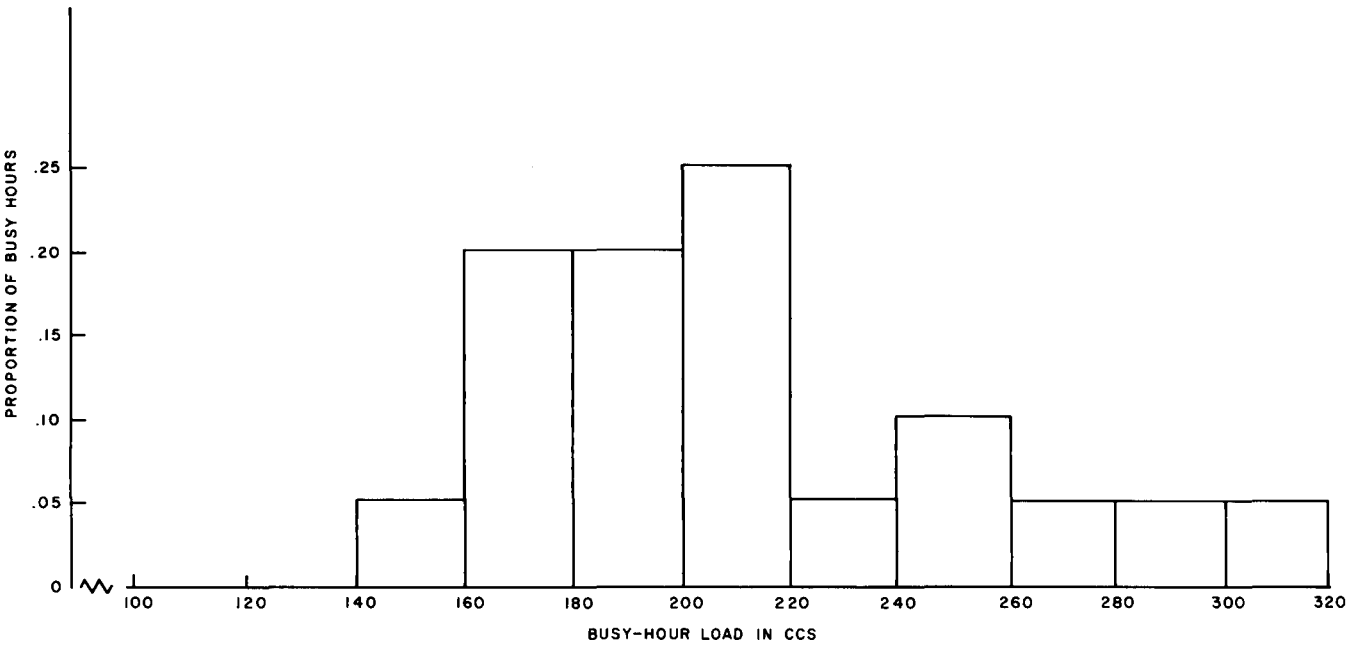

Fig. 6—Example of Day-to-Day Variation (20-Day Period)

rate change at some time of day, so may the initial period provisions of a tariff affect the distribution of holding times. Thus, a rate change may affect load by changing both the calling rate and the average holding time. Similarly, the advent of computer communications and other specialized uses of the trunk network may alter the average holding times that are employed in interpreting various types of load indicators such as overflow registrations.

D. Number of Sources

3.16 The number of sources that can generate calls also affects the traffic input process. The effect of having a finite number of sources is to increase the capacity of a trunk group of given size engineered for a specified blocking objective over what it would be for infinite sources. This increase in capacity results from the fact that with finite sources, as the number of calls present builds up, there are correspondingly fewer idle sources left to originate new calls.

3.17 Figure 7 illustrates the effect upon capacity of varying the number of sources. The capacities in this table were calculated from the finite-source, blocked-calls-cleared model, which is more commonly referred to as the Engset model. In examining Fig. 7, it is evident that the number of sources must be extremely limited to exert any appreciable effect upon the capacity of a trunk group. From a practical standpoint, the number of sources may be considered infinite when the ratio of sources to trunks is greater than 8 to 1.

3.18 In trunk engineering, a finite-source situation exists when a trunk group is dedicated to a particular class of service utilized by a limited number of customers. One example, which a trunk engineer should be aware of, is in a central office where there may be only 10 or 20 coin lines or message-rate customers offering traffic to a specific coin or message trunk group.

3.19 Finite source trunking situations often involve small trunk-groups (2 to 10 trunks) and small loads. In general, the precision with which such low loads can be estimated is limited. When these arguments are coupled with the fact that frequently the trunk group involved is the one and only route for traffic being generated by the customers in question, service protection considerations suggest that use of the Neal-Wilkinson capacity tables is reasonable provided that the number of trunks does not exceed the number of sources.

4. SYSTEM INDUCED FACTORS

A. Blocking and Retrials

4.01 In the mathematical models of service systems, two assumptions are commonly made about

NUMBER OF SOURCES	ENGSET CAPACITY OF 10 TRUNKS IN CCS
11	253
12	243
15	216
20	191
50	173
75	166
100	163
Infinite	163*

* NEAL— WILKINSON capacity for 10 trunks at B01 with no day-to-day variation.

Fig. 7—Effect of Finite Sources on Capacity at B.01

the disposition of calls that are blocked because of NC (no idle trunk available). In trunk engineering, blocked calls are assumed to be "cleared" or abandoned, in which case they vanish as if they had never existed. The Neal-Wilkinson capacity tables are based on this assumption. In some service systems in the telephone system (such as with Automatic Call Distributors), each blocked call is "delayed" until a server becomes available and then occupies the server for a normal holding-time.

4.02 In the telephone system, calls may not be completed for reasons other than trunks not being available, such as don't answers, customer line busy, or called station number change or disconnect. In the trunking theory, these attempts are accounted for by a reduction of the average holding time.

4.03 In the event a call is not completed successfully, the calling party either abandons permanently or decides to try again. Field studies of customer reaction to incompleted calls in a toll environment indicate a relatively consistent pattern of abandonment versus retrial over a wide range of blocking. The retrial probability and time to retry are aspects of customer calling behavior influenced to some extent by the indications for failure to complete.

4.04 Recent Bell Laboratories studies indicate that 19 percent of all unsuccessful attempts in the toll network are retried within 60 seconds, and 34 percent of those retrials are successful. Over 50 percent of these retrials are initiated within two seconds. The retrial rate is relatively higher (42 percent) for those attempts that encounter NC conditions.

4.05 The trunking theories normally employed to interpret trunk group measurements on grade-of-service groups do not effectively account for the effect of customer retrials and abandoned attempts at the higher levels of blocking. Consequently, there is a tendency to overstate first-route offered loads and the number of trunks required to reach a specified blocking objective unless special adjustments are made to these measurements. These adjustments are covered in Section 780-400-345.

B. Full and Limited Access

4.06 Another factor affecting trunk group capacities and load-service relationships is the "access" arrangement, which may be "full" or "limited." Under full-access conditions, every demand for service has access to every available trunk in a group. Limited access, however, means that a

specific demand has access to a predetermined number of trunks that is always less than the number of trunks in the group. Under this condition, it is possible for a call to be blocked when a trunk is idle in another subgroup of the trunk group. Thus, with a limited access arrangement, the probability of blocking is greater than with a full access arrangement.

4.07 Limited access represents a compromise on the part of equipment designers between the cost of connection arrangements and the efficiency of the equipment. A limited-access group is never as efficient as a full-access group. From a trunk engineering standpoint, limited access is a factor which affects the engineering of Step-by-Step, Panel, and No. 1 Crossbar graded groups.

4.08 In order to select the proper trunk capacity tables, it is essential for the trunk engineer to be familiar with switching system trunking features and capabilities as discussed in Section 780-402-170.

C. Measurement Accuracy

4.09 A final factor which influences the traffic load which a trunk engineer must deal with is the trunk group measurement accuracy. Unlike the other factors discussed above, the measurement process does not affect the ***actual*** offered load to a trunk group, but instead only the load which is ***observed.*** It is equally important that the trunk engineer understands all the characteristics of the observed load which serves as the basis for forecasting and servicing, as well as the characteristics of the actual offered load.

4.10 The accuracy of an estimate of offered load is a function of the size of the load, the number of hours observed, the method of measurement employed, and, in some cases, the call holding time. Offered load is often estimated from carried load, hourly values of which are usually measured with a traffic usage recorder. This instrument samples the carried load at discrete intervals. Loads can only be estimated within certain confidence limits. The effects of the size of the observed load and the number of sample values on the estimate of the source load is illustrated in Fig. 8.

4.11 Figure 8 shows the 90 percent confidence interval for the true source load given the observed CCS load. For example, given an observed load of 288 CCS averaged over twenty hours, there is a 10 percent chance that the true load might be lower than 276 CCS or higher than 300 CCS. As evident, the confidence in an estimate of offered load made from trunk group measurements increases as the number of hours of observed data and the size of the load increase.

5. TRUNK ENGINEER'S ROLE

5.01 The trunk engineer, when evaluating load service results, must take into consideration not only the external but also the induced characteristics of traffic. Theoretical capacities or blocking values are precisely that—***theoretical.*** Significant changes in any one of the assumptions

AVERAGE CCS LOAD OBSERVED	NUMBER OF HOURS IN AVERAGE			
	1	5	20	80
36	17-55	27-45	32-40	34-38
288	236-340	265-311	276-300	282-294
2880	2707-3053	2808-2952	2843-2917	2860-2900

Fig. 8—Approximate 90 Percent Confidence Limits—180 Second Holding Time

underlying the traffic capacity tables will distort the load service relationship, so that either another theory must be employed or theory must be supplemented by adjustments derived from experience.

5.02 The trunk engineer should be alert for sustained load service results that are not in conformity with these capacity tables so that steps may be taken to protect both present and future service. The trunk engineer should also communicate regularly with counterparts in other organizations so that traffic affecting changes in the telephone system will be foreseen and planned for in the future.