THE ROLE OF TRAFFIC ENGINEERING IN THE ENVIRONMENT OF A REGULATED MONOPOLY TELEPHONE SERVICE

R. MINA, J. M. KRAUSHAAR
Federal Communications Commission,* Washington, D. C., U. S. A.

ABSTRACT
In a regulated monopoly telephone environment the traffic and General Telephony concerned with the application of traffic theory as it relates to plant provisioning, plant servicing and utilization efficiency. This paper will deal with some aspects of service criteria and the specific conditions which are present in this environment. It is hoped that some of the possible pitfalls in the application of traffic theory to network design. We shall attempt to bring a greater awareness of the interactive effects of policies and their possible impact. We hope that cognizance of these problems and new approaches to regulation presented herein should ultimately lead to more effective standards and engineering techniques for plant provisioning and surveillance.

INTRODUCTION
In the United States the telephone industry and government have a rather unique relationship as compared to other countries. Because the U.S. telephone system is privately owned and is fundamentally a monopoly, a governmental regulatory structure has been set up to insure that the public interest is served. This structure comprises a set of regulatory bodies at the State level which have jurisdiction over local and intrastate facilities and the Federal Communications Commission which has jurisdiction over interstate facilities. The country is served by a number of single and multistate telephone companies which are subsidiaries of two major companies, American Telephone and Telegraph Company (AT&T), and Western Electric and Electronics Corporation (Bell). There are over 1,000 independent companies. A large number of these independent companies are owned by holding companies such as Continental Telephone Corporation. In addition, the bulk of interstate intercompany message telephone service is handled by AT&T Long Lines, one of AT&T's operating arms. Both AT&T and GTE have research and manufacturing facilities, of which Bell Laboratories and Western Electric are the most widely known.

It has been recognized for some time that effective regulation requires the establishment of criteria and indices which may be analyzed and tracked over time. Up to now monetary criteria, such as rate of return, depreciation, plant in service, etc. have been used almost exclusively in connection with accounting techniques and procedures to establish appropriate regulatory action and policy. In this paper we have developed several ideas in this context which emphasize some of the policy elements and figures of merit based upon traffic engineering and the plant provisioning process. The overall philosophy inherent in our approach is one of maintaining high utilization of plant over time without adverse impact on the customer. Achieving utilization objectives through an engineering process can be an invaluable complement to the monetary regulatory process. To this end, we hope to illustrate a number of facets of regulation which emphasize the engineering processess, in an attempt to focus on several critical areas where the engineer can be used to stimulate economic in the plant provisioning and planning processes, as well as to assure that the welfare of the customer is kept on par with that of the stockholders.

We shall focus on some aspects of service criteria and certain applications of traffic theory which affect the economy of operation. It should be noted that engineering design policies in certain combinations can result in a combined effect which may adversely influence the trunk quantities, switching terminations, and multiplexing (carrier) facilities to be constructed. Three topics, namely service criteria, day-to-day variation, and utilization measurement have thus been chosen to illustrate some of the possible pitfalls in the application of traffic theory to network design. We shall attempt to bring a greater awareness of the interactive effects of policies and their possible impact. We hope that cognizance of these problems and new approaches to regulation presented herein should ultimately lead to more effective standards and engineering techniques for plant provisioning and surveillance.

1. SERVICE CRITERIA
One of the main factors in the engineering process of plant provisioning and servicing in the trunking service criteria. In the limited space of this paper we shall discuss some of its aspects and applications in the toll network in the United States.

The trunking service criteria ultimately determine the proportion of call attempts which cannot find their way to their destination due to limitations in the number of trunking paths or in other elements of the network. These criteria are often defined in terms of blocking or grade of service, i.e. the proportion of unsuccessful call attempts which fall due to economic reasons over a fairly long period of time. Although this matter is a policy decision based upon a number of interacting considerations, it is ultimately desired to provide an overall quality of service which satisfies the customer and is economically feasible. Nonetheless, the interest of a company operating in a regulated environment may be served by liberalizing the service criteria to protect against forecasting errors, to minimize lost profits due to calls which do not reach their destination on the first attempt, and to minimize possible adverse criticism. It is therefore desirable that the quality of service in general and the trunking service in particular be authorized by the regulatory body. The role of the traffic engineer should be evident in such a process.

EVOLUTION OF PRESENT TRUNKING SERVICE CRITERIA
Since the early days of automatic telephony the trunks and switches at each switching stage in local central offices of the Step-By-Step type have been engineered at a low probability of blocking such as P.01. The average service in 20 time-consistent busy hours in the busy season has become the service objective for essentially all types of trunk groups engineered on a probability basis. In the toll network the P.01 blocking standard is applied to the toll access, toll completing and final intratoll trunks.

Since the service criteria apply to the final or full group in each network cluster - they control the

A network cluster is a set of high usage trunk groups whose traffic overflows to a common final trunk group.
overall level of blocking within each network cluster. Remembering that over 80% of the traffic is carried by the high-usage groups which are engineered to economic criteria, the use of P.01 is in effect equivalent to better than P.002 with respect to the primary routed traffic. Apparently one of the main reasons for that practice was to assure that the primary traffic routed directly to the final route would experience the desired (P.01) grade. However, the subsequent practice of providing special high-usage groups for such parcels of traffic, now called "parallel protection high-usage groups" appears to have made the specification of P.01 with respect to the peaked traffic offered to the final route anachronistic. Other techniques for treating this problem can be found in the literature [1,2]. Nonetheless, the P.01 specification has been recommended by the CCITT 2/ not only for provisioning of trunks in final international routes but for "national" trunks as well.

The subscriber views the blocking criteria as significant only insofar as the overall probability of completing the connection to the called party is affected. In a well maintained network it is known that 20 - 30% of call attempts do not result in voice connection for reasons beyond the control of the operating company. This includes 'called line busy', 'no answer', and customer dialing errors. In addition, some 2 - 4% of the attempts fail due to equipment malfunction and blocking. In general, equipment malfunction has a more adverse effect on service than blocking.

Newstead [4] illustrated by convincing evidence that even a ten fold increase in the level of blocking will not be noticeable by the users and will only result in a small decrease in the percent of call completion. Molnar [3] also states, "Congestion due to limitations in trunking [is] insignificant, compared with the number of unsuccessful tries that are due to circumstances unrelated to the structure of the telecommunication plant. This being virtually unnoticeable to most users, congestion in trunking could conceivably substantially be increased without adversely affecting service."

Various attempts have been made to establish a sound economic basis for the selection of service criteria, but this question is still unresolved. Molnar [3] in his paper presented at the Seventh ITU in 1973 states "At the present the prevailing view in the telecommunications field about the grade of service is just about the same as could be said nearly a quarter of a century ago." He concludes by stating, "the grade of service which is one of the most basic, most elusive and until recently the most neglected tasks of teletraffic engineering .... Even so it is becoming more and more under attack as being arbitrary, often inefficient or uneconomical and lacking realistic correlation between performance and service quality as accepted by the users."

PRESENT AND PROPOSED BLOCKING CRITERIA

Examination of the mathematical expressions for the call congestion for the three types of traffic dealt with in practice, namely, smooth, random (Poisson) and rough (peaked) traffic show that the respective criteria do not express the same thing. For example, one percent blocking in a group of servers to which peaked traffic is offered is not equivalent to one percent blocking in a group of equal size to which an equal load of Poisson traffic is offered. It is not rational, it would appear, that blocking criteria should be specified relative to the distribution of the offered traffic or be equated to the equivalent Poisson base developed from the Equivalent Random Theory [6]. For example, one percent blocking on a final group consisting of 84 trunks to which is offered a load of 62 erlangs with peakedness factor 2 is equivalent to about .5% blocking with respect to the equivalent load of 130 erlangs of Poisson traffic offered to an equivalent group consisting of 353 trunks.

Figure 1 shows the service criteria used for the provisioning of the main links in the toll network in the United States. End-to-end blocking is not an objective in itself because at present there is no way to calculate it. Obviously, it is not the sum of the blocking in the various links in the chain of connection because of non-coincident busy hours, exhaustion dates and other reasons enumerated by Longley [5]. Examination of this figure shows that the blocking is not apportioned relative to the cost of the links. For example, P.002 (the equivalent of P.01 for the final trunks) is used for the provisioning of the expensive intertoll trunks and P.02 is used for the "incoming matching loss (IML)" in the terminating office. The latter can be materially improved by giving toll completing trunks more than one appearance in the terminating link frames at insignificant cost.

To correct for present problems, we propose as an interim measure that the trunks in final routes be engineered so as to allow 1% of the total originating traffic of each network cluster to be blocked in the final route. In addition, it would be desirable to enhance the probability of completion of calls which traverse the bulk of the toll network and reach the terminating toll center. This can be achieved by reducing the blocking on the toll completing facility as illustrated in fig. 1.

FIGURE 1.- SHOWING PRESENT AND PROPOSED BLOCKING CRITERIA IN THE U. S. TOLL NETWORK

2/ Recommendation E.541 CCITT Sixth Plenary Assembly Volume II-2, p. 236 [12].

ITC-9
MINA / KRAUSHAAR-2
2. Validity of the Theory of Day-to-Day Variations

Another element introduced lately in the process of trunk engineering is the theory of day-to-day variations first presented in 1958 by Wilkinson [7,11] and modified in 1976 by S. Neal [8]. Wilkinson sought to find the distribution of trunk traffic offered to a final route taking into consideration the day-to-day load variations. He states "if we can predict the day-to-day joint probability of the average and variance of this load for an alternate route system, we shall have at hand the means for estimating the proportion of days in the busy season during which the network service will drop below any given level" [7].

In this model the blocking is determined on the basis of the average daily blocking:

\[ B = \frac{1}{2} \sum_{i=1}^{n} \]

The average blocking, \( B \), is calculated by formula 10 in reference [8]. The day-to-day variance is assumed to be related to the average busy hour load by the empirical formula \( \frac{W}{2} \) in the same reference. It will be noted that Neal's modifiication sought to separate the random and systematic components inherent in Wilkinson's model. He reduced the daily variance by a small amount, \( \frac{W}{2} \).

If our understanding is correct, then it seems that the objectives of the trunk provisioning process contradicts the specification of the blocking criteria which had been computed on the basis of the average daily load rather than the average daily blocking.

The Neal-Wilkinson Trunk Capacity Tables based on this model have become the standard tables in the Bell System for provisioning and servicing all types of trunks engineered to blocking criteria. These tables have also been recommended by the CCITT 5/ for engineering international circuits. We note that the basic concepts and assumptions of the day-to-day model continue to be accepted and implemented without careful scrutiny and that, to our knowledge, their validity and implications have not been examined at this conference. Because of the far reaching economic impact of this empirical model, we believe that its examination now is timely. In the limited space of this paper we will examine some of the assumptions upon which these tables are based.

Basis of the Neal-Wilkinson Tables

The tables are primarily based on the assumptions that the traffic from day-to-day is time-variant or non-stationary and that the variations within each daily busy hour are stationary (Poisson distributed). These two assumptions, combined, are not in accordance with the ergodic property of traffic used in classical traffic theory which requires that the properties of a sample taken over a fairly long period of time exhibit the same properties as the ensemble:

\[ \{ B(m) f(m) dm \} \]

where \( B(m) \) is the single hour expected blocking and \( f(m) \) is the distribution density of the day-to-day load \( m \), assuming a Pearson Type III distribution. See also 5/.

\[ \text{Var}(m) = 0.13 \left( \frac{m^2 - 2 m - \mu}{m^2} \right) \]

where \( m \) is the average busy hour load, \( \mu \) is its peakness.

\[ \mu = m \] (5/h)

\( \text{Var}(m) \) is the variance of the daily average load, \( z \) is its peakness.

5/ Recommendation E.521, CCITT Sixth Plenary Assembly Volume II.2, p 213 [12].


Although the traffic intensity varies from day-to-day according to the calling needs of the users, for all practical purposes the assumption has historically been made that a string of 20 busy hours would characterize stationary traffic. Wilkinson's model appears to revolve around the fact that 20 separate time-consistent busy hours are insufficient to represent a stationary capacity. He observed from field measurements on some final routes that the variance of the daily measurements tends to exceed that expected from a stationary process and developed an empirical model to account for this situation.

His studies were not extended over longer periods such as 30 hours as recommended by CCITT to fully quantify day-to-day variations. It is counter to the variance tends to approach a limit with the extension of the period of observation. Although some of these short-comings were recognized by Neal [8], the basic assumptions were not questioned, nor was the effect of limitations of traffic measurement facilities clearly documented.

Use of the second assumption that the traffic within each daily single hour is in statistical equilibrium and exhibits the properties of stationary traffic is also contrary to the basic fundamental principles of traffic theory which require that the traffic parameter(s) used in the Erlang trunking formulas must be the average of many hours. The result of this assumption is the so called "single hour engineering" model in which blocking can be determined from the Erlang formula on the basis of single hourly measurements. It should be noted that the Erlang formula which relates blocking to average offered load is based on a stationary day-to-day period of time is a probabilistic formula, and as such does not define a deterministic relationship for measurements over an arbitrarily short time interval such as one hour.

Kruithoff [9] implies a number of basic criticisms to the "single hour" approach as it is expected that the day-to-day variations and caution against carrying the disparity between real traffic and theory too far. He asserts that the traffic engineer is "faced with the situation that on one hand he has a number of equations at his disposal that are based on the principle of statistical equilibrium, whereas on the other hand he has to deal with traffic data that clearly show that the actual traffic does not satisfy that condition." Despite the fact that the actual traffic may exhibit a non-stationary character, he concludes that "the introduction of a function that describes the daily fluctuations of telephone traffic does not appear to be effective, useful, or even necessary."

The Neal-Wilkinson Trunk Capacity Tables consist of tables computed for high, medium and low day-to-day variations. In addition, there is a table without allowance for day-to-day variation. Partial tables are published in CCITT Volume II.2, p 213 - 223. Examination of the table without allowance for day-to-day variation shows that the traffic capacity of small trunk groups without peakness exceeds that indicated by the Erlang formula. The tables in reference [11] without day-to-day variation and without peakness appear to agree with the Erlang tables. As day-to-day variations are older and new tables show reduced traffic capacity for a given grade of service compared to that given by the tables based on the Erlang theory.

"every sample function of the ergodic process is the same in statistical properties, so that any one can be used to represent the process. If one sample function is inspected over a sufficiently long period of time, then all features of the process will be observed. The record of one observation of a particular realization taken for a long period of time can be cut into pieces of a certain length and the different pieces can be considered as different records of a set of observations of the ensemble."
An example annexed to CCITT Recommendation E.521 (see footnote 5) states that an average load of 11.39 Erlangs with peakedness of 1.3 and high day-to-day variation requires 23 circuits at P.01. A similar calculation without day-to-day variation indicates that 21 circuits are required.

The referenced example suggests that it is both necessary to increase (from 11.15 to 11.39 Erlangs) the composite of individual peak of overflow loads on the basis of the "single hour" principle to arrive at a so called "true overflow load" and to use the (high) day-to-day variation table to determine the required quantity of trunks. The combination of these two processes increases the number of trunks over that determined from Erlang theory from 21 to 23 trunks.

It is possible that the use of these tables may increase the trunk requirements by as much as ten to twenty percent without apparent benefit to the user.

3. SURVEILLANCE OF NETWORK EFFICIENCY

The application of service criteria, engineering techniques and other policies will be ultimately reflected in levels of plant utilization and cost. Measures of plant utilization can be used as a feedback indicator in the process of regulating the level of trunk and plant additions. In the United States there is an increasing interest in measuring and tracking utilization in the telephone network. This is particularly important in times when the cost of raising new capital is high, as is presently the case. Utilization efficiency must then be balanced against desired service standards. The Federal Communications Commission has thus been concerned with plant utilization as it relates to construction cost and the public interest.

An outgrowth of this new interest is the establishment of a group within the Federal Communications Commission charged with monitoring and tracking performance levels and utilization measures for the domestic telephone network. This effort has identified sources of data which have yielded valuable insight into current engineering practices and policies, such as those already discussed. In the area of trunking, utilization measures based upon the Carried CCS/trunk (Erlang/trunk) have been examined along with techniques of comparing the actual number of trunks in service with the theoretical number required, based upon forecast traffic loads. One technique comparing histograms of actual trunk group utilization vs. planned utilization has been employed. Figure 2 illustrates a typical overlay of such histograms.

![Figure 2: Typical Overlay of Actual and Design CCS/Trunk Distributions](image)

Data indicating the carried CCS, number trunks per group and Economic CCS criterion in 1976 on several hundred trunk groups in one company were used to calculate a value for the CCS/trunk group and to develop statistical distributions from these sets of design and actual utilizations. The results showing the design and actual distributions were plotted and superimposed, yielding a good deal of insight into performance. The large observed variance of the resulting actual CCS/trunk histogram compared with that of the design CCS/trunk indicated a need for greater control over trunk utilization and forecasting.

While statistical distributions of trunk utilization coupled with data recording capability may yield deeper insight into problem areas relating to specific trunk groups, the additional computational and data handling effort may not justify such an approach in all cases. It therefore has been useful to develop summary data relating to the mean of trunck utilization for each jurisdiction of interest. The format of a typical computer generated report developed at the POC and based upon Company records is shown in figure 3.

Specific items of interest are as follows:

- **Number of trunks**: In Service
  - With Valid Measurement
  - Theoretical Required
  - Short
  - Per Group

- **Excess Trunks**: Number Percent of total trunks with valid measurement
  - Measured
  - Theoretical

- **Weighted % NC**: By tracking data of this kind, it is possible to separate consistent differences or trends which may require further investigation.

Despite current interest in the topic of utilization, search of past papers reveals that the subject of performance has not heretofore received a great deal of attention. One of the few papers we encountered was entitled "Network Efficiency and Network Planning Considering Telecommunication Traffic Influenced by Time Difference" by Tohru Ota (10) was presented at the Seventh ITC Conference, which discusses time distributions of traffic. It appears that further study is needed to broaden the use of distribution analysis and observation techniques in the areas of plant administration and utilization to include additional variables of interest. We anticipate that further interest will develop in this important area in future ITC Conferences.

![Figure 3: Typical Trunk Utilization Analysis Report](image)
SUMMARY AND CONCLUSION

The primary emphasis in this paper has been placed on trunk provisioning, covering a discussion of several concepts in relation to their impact on trunk utilization from regulatory viewpoints. In a time when the limitations on economic growth are beginning to be felt, we believe that it is becoming increasingly important to be aware of sources of waste and inefficiencies so that the rate payer may receive the most service for his dollar.

We have shown that the present service criteria for plant provisioning and servicing tend to be liberal and are not always clearly defined. This, for example, has become particularly important with the introduction of day-to-day variation and single hour engineering concepts. Although the problems we have outlined are complex, we have proposed a number of changes and clarifications in the service criteria definition which could assist in providing greater economies in the telephone network.

To correct for present problems as is proposed, as an interim measure that trunks in final routes be engineered so as to allow one percent (1%) of the total originating traffic at each network cluster to be blocked in the final route and further that the probability of blocking in the toll completing facility be reduced. This will enhance the probability of completing calls which reach the terminating toll center.

Concerning the theory day-to-day variation we have shown that the Neal-Wilkenson trunk capacity tables show traffic capacity less than that indicated by the Erlang tables. The use of these tables may increase the trunk requirements by as much as ten to twenty percent without apparent benefit to the telephone user.

Finally, we have illustrated sample methods for the surveillance of network trunking utilization efficiency. Consideration of these ideas and interactive effects of separate facets of the service criteria should ultimately lead to a lower cost network required to meet customer demand without noticeable change in service.

ACKNOWLEDGEMENT

The authors are grateful to Mr. C. Cowan and Mr. A. C. Latker of the Common Carrier Bureau for their kind encouragement and support the preparation of this paper.

REFERENCES


(12) CCITT Sixth Plenary Assembly Volume II.2.