Thesis Summary:

On-line Call Admission for High-Speed Networks

Anja Feldmann

September 22, 1995

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Daniel D. Sleator, Co-Chair
Bruce Maggs, Co-Chair
Allan Fisher,
F. Thomson Leighton, MIT

Abstract

New networking technologies such as asynchronous transfer mode (ATM) can carry a wide variety of traffic – voice, video, and data – over a single digital network. Since some of these services depend on bandwidth guarantees, the network should employ some mechanism for preventing network links from becoming overly congested. This includes making intelligent decisions about whether to accept a call, and if so, when to schedule it, and how to route it. Such strategies are called call admission algorithms. Challenges in the design of such algorithms include high bandwidth requirements and the on-line nature of the problem: when a call comes in, the algorithm has to make an immediate decision about scheduling the call. A call's duration can only be determined by observing how long it takes to complete.

A necessary first step in designing call admission algorithms is to understand the nature of traffic. In this thesis we develop traffic models for the connection behavior of traffic currently observed on the Internet. We characterice the interarrival times, the bandwidth requirements, and the durations of TCP connections. Heavy-tailed distributions, especially the Weibull distribution, yield statistically better models than traditional models such as exponential distributions. We also find evidence of the self-similar nature of connection interarrival times. This implies that the burstiness of current TCP traffic, and most likely future multi-media traffic, differs substantially from conventional traffic models.

To cope with bursty traffic, high bandwidth demands, and unknown call durations we suggest algorithms that may delay calls. We use competitive analysis and simulations to evaluate our algorithms. Our results include an analysis of the most basic greedy call admission algorithm. We show, among other results, that GREEDY is $\Theta(\log n)$-competitive with respect to network utilization on n-node trees for arbitrary unknown durations and arbitrary bandwidth, and that no algorithm can be better than $\Omega(\log\log n/\log\log\log n)$-competitive.

Our simulations show that call admission algorithms with delay are robust; even when subjected to very bursty traffic and high bandwidth calls, and achieve better network utilization without affecting many more calls than algorithms without delay. We observe that given a certain achieved network utilization the percentages of calls rejected by a greedy algorithm with delay is about the same as that of a greedy algorithm without delay. Therefore delay may offer the right compromise between the network service provider and the users. We show both in theory via competitive analysis and in practice via simulations that greedy call admission algorithms that may delay calls are superior to those that do not.

Thesis Summary

Recent advances in network technology suggest that in the future various classes of traffic – voice, video, and data – will be transmitted on the same circuit-switched high-speed network. A network with this capability is called an integrated service network. Until now different networks based on complementary technologies have been used for different traffic classes, such as the phone network for voice traffic and the Internet for data traffic. Besides the cost of having two separate networks, the lack of an integrated network has hampered the development of multi-media applications, in which streams of different traffic need to be carefully synchronized. Furthermore the new technology allows the promise of flexible bandwidth allocation, bandwidth reservations, and latency[1]. Achieving this will require addressing many challenging new network management questions.

Users of an integrated network might, for example, decide at any time to hold a video conference between Pittsburgh and San Francisco. This involves submitting a request to a service provider for a connection between the two nodes of the network. In this request the user specifies the desired bandwidth and quality of service she or he needs. Since in general the user does not know how long he will need the connection (e.g., for a video conference), he is not necessarily willing to commit himself to a predetermined duration. The user expects to obtain the desired connection within a reasonable time and with a guarantee that he will receive the service specified in the request. Examples of services include voice service (constant bit rate traffic), uncompressed video service (constant bit rate traffic that requires roughly 500 times the bandwidth of voice service), bulk file transfer service (variable bit rate traffic), computer connection service (variable bit rate traffic that requires roughly 15 times the bandwidth of voice service), or some combination thereof.

Call admission

The service provider operates a network with different capacities along different links which can be modeled as a graph in which each edge is labeled by the capacity of the link. For the service provider to guarantee service, certain shared resources of the network, such as bandwidth and internal buffer capacities, have to be reserved along some path between the source and the destination nodes of each connection. Since several requests might require the same network resources a *congestion control strategy* prevents network links from becoming overly congested. As part of its congestion control strategy the service provider uses a *call admission* algorithm to decide when to grant which request along which path.

A congestion control strategy that employs call admission differs from traditional strategies, e.g., those used in the current Internet, in that it is *preventive* rather than *reactive*. Reactive congestion

[1]Latency is the delay experienced by a bit of information between the time that it is injected into the network by the sender and the time that it is arrives at the receiver.

control strategies alleviate congestion by degrading the service of already established calls, e.g., by preemption or loss of packets or cells. This causes an infringement on service guarantees. Due to the high bandwidth-to-delay ratio of high-speed networks, the time until an action is initiated by a reactive congestion control strategy cannot be ignored and brings into question the use of reactive congestion control schemes for high-speed networks. Preventive congestion control includes call admission, the topic of this thesis, and flow enforcement (policing). Flow policing enforces that each connection does not use more than its allocated share of bandwidth.

The goal of the service provider is to achieve high network utilization while still providing the users with the services they desire. In this thesis we the *on-line* version of the problem where the algorithm cannot forecast future requests nor the lengths of current connections. Instead the algorithm must make its decision in an on-line fashion.

ATM networks

Some aspects of the call admission problem considered in this thesis are motivated by new network technologies such as asynchronous transmission mode (ATM) [For93, All93, All95], a technology that can support a wide range of traffic classes. ATM uses asynchronous packet-switching of cells of small fixed size at its lowest level. At the abstraction level of interest to the call admission problem it is fundamentally connection-oriented. Its link transfer rates range from less than 64 Kbit/s to more than 2.4 Gbit/s. Currently the most common link transfer rates are 100 Mbit/s or 155.52 Mbit/s based on TAXI or SONET multi-mode fiber interfaces. Each link can support several thousand connections.

Most of the applications mentioned above have strong minimal bandwidth and maximal latency requirements. In ATM the small size of the cells overcomes the problem of the uncertain delay experienced by nodes in packet-switched networks with their variable length packets. Since transmission of different types of data can be interleaved in the small cells, the length of time that a cell carrying delay-sensitive traffic must wait is short. The key difference between ATM and other connection-based networks is that ATM uses asynchronous scheduling while connection-based networks like telephone lines usually use synchronous scheduling at the cell level. Therefore ATM can reserve bandwidth over a period of time while other connection-based networks need to specify the exact positions of the data of each connection in the overall stream of data. Thus ATM uses virtual connections, which allow flexible and dynamic bandwidth reservation, while phone lines use static bandwidth reservation schemes.

The aspects of ATM network that are of interest to us are that ATM networks are connection-oriented and provide flexible bandwidth guarantees. Therefore we use connection-oriented abstraction of the network to develop algorithms for the call admission problem. Each request specifies two nodes and the bandwidth of the requested connection. A request does not necessarily specify the duration of the connection. The call admission algorithm may schedule several requests on the same link at the same time if the sum of their bandwidth requirements is less than the link bandwidth, and it is not necessary to specify how the bandwidth is shared.

Bandwidth requirements of applications

It is commonly expected that traffic from applications such as telephone, television, virtual reality, and typical computer applications will be sharing the same physical network. The expected bandwidth requirements of some possible applications are summarized in Table 0.1.

Example	Information Type	Bit Rate
Telephony	Audio	64 Kbit/s
Television	Video	1 – 45 Mbit/s
Virtual reality	3-D Graphic	100 Mbit/s
Packet oriented data	Data	$\approx$ 1 bit/s –100 Mbit/s

Table 0.1: Bandwidth requirements of possible applications

We can see that the bandwidth requirements of these different applications differ dramatically. In addition, it is not valid to assume that the bandwidth requirements are much less than the link bandwidth.

Treatment of call requests

A call admission strategie specifies the treatment of call requests whose requirements cannot be satisfied immediately. Calls can be *rejected (blocked)* or *delayed (queued)*, or some combination of the two. This work analyzes the benefit of allowing requests to be delayed.

Perception of rejection vs. delay of calls

Allowing a call admission algorithm to delay some calls by a small amount should be acceptable in practice because in practice almost all calls will experience some delay independent of the call admission algorithm. Consider initiating a phone call, setting up a telnet connection, or sending an email message. It is in fact likely that any of these actions will encounter some delays, caused by latency, transit time, etc., which are outside the control of a call admission algorithm. These delays are routine and almost always acceptable.

With regard to rejection consider the following scenario. After an earthquake in California you want to call your friend in San Diego. Since many other people have the same idea, the network becomes overloaded and you get a busy signal; i.e., your call has been rejected. Most likely you will redial the number (or use automatic redial), and in this way re-queue your request, or you might switch to a different network service provider. An alternative more user-friendly design would allow you to just dial once, and, if the network is overloaded, your phone would ring automatically at some later time after the connection has been established. This can also be more efficient since it avoids additional signaling (e.g. computer redialing).

Delaying calls instead of rejecting calls may provide additional information to the call admission algorithm at no extra cost. For example a delayed call may imply that it would have to be rejected by a pure rejection algorithm, or a large number of delayed calls might imply serious network problems and that rerouting attempts should be put off until further notice.

Users who's calls have been rejected will often try to resubmit the call at a later time. How often and how fast this retry happens in summarized by the retry behavior of the users. Traditional conclusions of the analysis of retry behavior that the periods between retrys are rather long and that the retry rates are rather short, e.g., [MW73], must be changed due to the involvement of computers. Computers are more patient, allow automatic redial at a faster speed, and enable applications that do not even involve a human. It is our impression that we will see faster retries and higher retry rates as previously assumed, e.g., see [Sys60].

History of delay vs. rejection of calls

Giving the call admission algorithm the freedom to delay calls is not a new idea but has been used extensively in the past. One example of its application domain is telephony. In the early years of telephony [Fag75, Sch82, Rey88, Jol67], rejection was used for local calls while delay was applied to long-distance calls.

Call admission for local calls functioned much the same way as today; if there are insufficient networking resources available, the call is rejected. Usually, local service was made available at a fixed cost per line per month or quarter. (E.g., the first commercial exchange was established in New Haven, Connecticut, in 1878, for the cost of $1.50 a month, [Fag75, page 477]). Since a typical local network consisted of a full crossbar network, each line serviced only one customer. Providing instantaneous service, i.e., service without delay, has the disadvantage that lines are used very inefficiently. Yet it was possible to tolerate this for local service since the equipment was simple and inexpensive enough.

But since early equipment to bridge long distances was expensive, long-distance calls required a different scheme from local calls. The high costs resulted in charging on the basis of distance and duration of calls instead of on a subscription basis. The high cost and long distances resulted in low traffic density, which in turn justified only one or at most a few lines (trunks) between any points. Therefore it was necessary to find a more efficient way to use these lines: delay of calls. By providing delayed service, lines could be kept busy over long periods. Calls would be queued up in the order of receipt and completed over the lines as they became free.

For a typical call in the early days of telephony a customer would call an operator, provide him with the information about the called and calling party, hang up, and then wait for a call back. The operator would contact the operator at the called site, who would then call the remote party and initiate a call back. Later, after the automation of most of the call setup and routing process, it was possible for the caller to stay on the line for most calls. The operator was responsible for handling all the tasks involved in delaying calls.

This basic operating method did not change until direct customer access to toll trunks (long-distance lines) was introduced. One of the main reasons to introduce direct distance dialing was to eliminate operator involvement in all long distance calls. But since the computer systems, switch hardware, and signaling mechanisms were not capable of handling delayed calls, the phone service now had to be provided on a demand rather than on a delay basis. The switch from a delay based system to a demand based one made it necessary to add many more trunks. During the year of 1955 this was made possible by the following two facts: the price of the equipment had come down and an increased demand justified more trunks. Due to the relatively high cost of undersea cables, it was not until 1970 that international direct distance dialing was introduced.

In the 1960's, intractable technical problems, such as insufficient switch hardware, computing power, and software, caused the elimination of delay from long-distance telephony. Yet with the evolving technology some features have been reintroduced. For example, AT&T introduced enhanced private switched communication service (EPSCS) in 1978 which offers call queuing when a network, a foreign exchange, or wide-area telecommunications service facility is busy [Rey88]. Another example is that phone companies have now reintroduced services such as automatic redial and call return. Other applications that rely heavily on delay are airline or car rental reservation systems.

We note that today's phone networks have low rejection rates for the following reasons. Over the years phone companies have collected extensive records of the use of the resource of their networks. Such records are refered to as network traces. These traces have for the most part confirmed that human behavior with respect to phone calls is very predictable and can mostly be explained by exponential distributions. One example of the predictability is that the days with the highest bandwidth demands always include Labor Day and Christmas. This predictability, combined with the existance of many trunks, leads to nice statistical models about the load on the phone network and provides us with good bounds on the overall bandwidth demands with small predictable variations. For example, if the time between consecutive arrivals of connection requests follows an exponential distribution, it is very unlikely that the bandwidth demand randomly increases by a factor of two. All in all phone traffic allows and justifies a design for peak usage, i.e., the demands on days like Christmas.

Summary

From this brief summary of the history of telephony we can see that delaying calls has been beneficial in situations where the ratio of the bandwidth demand to the bandwidth supply is small and low network utilization is intolerable. Delay was eliminated from the phone networks in the 1950's since at that time the complexity of automatically delaying calls was beyond the technically tractable.

The high bandwidth requirements of tomorrow's applications and the high cost of high-speed networks will require call admission algorithms that achieve high network utilization. History has shown that delay can provide superior network utilization and we feel that delaying calls is again feasible since today's computers provide us with the technical capability. In this thesis we show that future networks will benefit from delaying calls.

Desired features of call admission algorithms

Since the network cannot forecast future requests by users, it should make its decisions on-line. In addition, a user is hardly ever willing to commit himself to a predetermined duration when initiating a connection request. Therefore the call admission algorithm should (1) allow *unknown durations*, i.e., the duration of a request can only be determined by admitting the call into the network and observing how long it takes to complete. Also (2), as long as the demand on the network is light the user expects good and prompt service. Should the network be overloaded, degradation of service is acceptable; yet (3) if the overload is gone the service should return to normal as quickly as possible.

It is the task of the call admission algorithm to support guaranteed service for all classes of applications – data, voice, and video. (4) xoThe algorithm should be able to cope with the wide range of bandwidth requirements and durations. Since preemption, discontinuation of an already admitted call, constitutes a breach of guarantee by the network provider, it is unacceptable.

Because high speed networks are an emerging technology and are not yet readily available, characterizations of tomorrow's applications are still unknown. We can only speculate about them based on other related data. Yet, because of the high bandwidth requirements, we can expect that their characteristics will change from month to month. This is likely to include changing traffic patterns and increased burstiness [2] of the call arrivals. The call admission algorithm should respond quickly and with little performance degradation to these changing characteristics.

In this thesis we mainly consider the most basic call admission algorithm, the *GREEDY algorithm*, which schedules each call as soon as it can satisfy its requirements. Whenever GREEDY does not already automatically satisfy a feature mentioned above, we will give evidence throughout this thesis that features (1) – (4) above are indeed satisfied.

Another desirable feature of call admission algorithms is to handle advance bandwidth reservations. Yet for the following two reasons advance bandwidth reservations collide with unknown durations. Bandwidth resource committed to currently scheduled calls cannot be used for future reservations because, since their durations are unknown, it is unclear if the calls will complete in time. Bandwidth resources committed to future reservations cannot be used to schedule current calls because they might again not complete before the start time of the reservation. In this work we concentrate on studying unknown duration and refer the resolution of the above conflict as well as all issues connected to economic questions to future work.

Service design space

There are several models in which a service provider may take advantage of delaying calls. For example it may allow unlimited delay of calls. From the service provider's perspective this provides the most flexibility and therefore should give the best results with respect to network utilization. Yet from a user's perspective this is only attractive if in general the delays are fairly short.

One way to limit the length of a possible delay is to force the call admission algorithm to reject all calls that cannot be admitted into the network within a fixed amount of time after they are submitted to the network. Although this restricts the possibilities of the algorithm, the added flexibility of limited delay may still be sufficient to increase the network utilization.

Of course a user might be willing to wait longer for some calls than for others. One possible criterion is the duration of the call. Since presumably the patience of the user is correlated to the duration of the call, a service provider might offer a service where the delay is limited to be proportional to the duration of the call.

Another option is that the network provider responds to a submitted call by supplying an estimated delay. This provides the network provider with flexibility and the user with information about what delay to expect. The major problem with this option is that it requires knowledge about the durations of the calls; we do not study this model.

Approach and outline

A necessary first step in designing call admission algorithms for high-speed integrated service networks is to understand the nature of traffic in such networks. Only once a statistical model to forecast future requests has been established and verified can traditional approaches of analyzing congestion control schemes, such as queueing theory and simulations can be applied. Yet, accurate traffic modeling for high speed networks is currently impossible because access to high-speed networks is not yet readily available and most applications are either not yet available or still under development.

[2] A burst of call is a set of calls arriving within a short time-period. A sequence of calls is bursty if the number of burst is high.

Traffic study

One major traffic contributor and probably the closest approximation to the kind of traffic we expect to see on future high-speed networks is the traffic we currently observe on the Internet, or more specifically the traffic using the transmission control protocol (TCP) traffic. We present the results of a *connection level study of TCP traffic* on an local-area network based on Ethernet technology. The study identifies TCP connections characteristics, such as connection interarrival time, connection duration, and connection bandwidth requirements.

We show that the traditional approach of modeling connection interarrival times, durations and bandwidth requirements by exponential distributions [JK75] is inappropriate. For all call parameters we produce statistical evidence that heavy-tailed distributions such as Weibull, Pareto, and lognormal distributions yield better models. We observe that even within one class of applications there is a wide range of different values for all traffic parameters including durations and bandwidth requirements. While we are unable to provide statistically good models for the distribution of durations and bandwidth requirements we show that the Weibull distribution yields good fits for all connection interarrival times, for almost all application-specific interarrival times, and for time-dependent interarrival times. That heavy-tailed distributions yield better models than the exponential distribution implies that the traffic is much burstier and much less predictable than commonly assumed. The additional observation that the interarrival times of several applications exhibit self-similar behavior confirms that data traffic and probably multi-media traffic differs substantially from conventional telephone traffic and traffic models considered in the literature. Indeed, contrary to common assumptions, it is likely that the burstiness of the request sequence will increase as the traffic load is increased.

Based on these results we believe that call admission algorithms should be evaluated not just based on exponentially distributed parameters, but also for other burstier interarrival times, and even within a worst-case scenario. Competitive analysis provides us with the possibility to analysis the worst case behavior of an on-line algorithm and is therefore a natural choice in the absence of good statistical models for connection characteristics.

Competitive analysis

Competitive analysis [ST85], i.e., comparing the solution that an algorithm produces on a sequence of requests to the best solution that any algorithm could produce on that sequence of requests, does not rely on forecasting future events and as such is an alternative method to analyze call admission algorithms. We say that an algorithm is k-*competitive* if its performance is within a factor of k of the optimal performance on any sequence of requests. In this thesis we use competitive analysis to gauge the benefit of allowing requests to be queued for an unlimited as well as for a limited amount of time.

We use four measures to quantify performance: network utilization (via an inverse makespan[3]), maximum response time, call-admission ratio, and data-admission ratio. Our results include proofs that greedy algorithms are $\Theta(\log n)$-competitive with respect to makespan on n-node trees and $O(\log^2 n)$-competitive on n^2-node meshes for requests with arbitrary durations and bandwidth requirements, a proof that on an n-node tree no algorithm can be better than $\Omega(\log\log n/\log\log\log n)$-competitive with respect to makespan, and a proof that no randomized algorithm can be better than $\Omega(\log n)$-competitive with respect to call-admission and data-admission ratio on an n-node linear array, if each request can be delayed for at most some constant times its (known) duration.

[3]The makespan of a series of calls is the total time required to complete all all of them

In contrast to previous work on competitive analysis, we show that, if a call admission algorithm is allowed to delay calls, a very natural algorithm, GREEDY with a new routing strategy, has good worst-case behavior and has the features we identified as desirable for call admission algorithms. Since all results derived from competitive analysis are independent of the traffic pattern and the specific request sequence, we expect that the performance of GREEDY with delay is robust with respect to the burstiness of the request sequence and traffic pattern even in the presence of arbitrary bandwidth requirements and unknown durations. This intuition motivated us to explore further the benefit of allowing call algorithms to delay calls via simulations to judge their typical performance.

Simulations

Simulations give us the opportunity to explore many different scenarios and compare the performance of algorithms that reject calls to those that delay calls for the same request sequences. We describe simulations for GREEDY with delay, GREEDY with limited delay, and GREEDY with rejection in different scenarios. The networks we use are motivated by local area as well as wide area networks while the characteristics of the request sequences are motivated by the results of the traffic study. Some of the routing algorithms are motivated by the results of the competitive analysis.

The simulations provide evidence that delaying calls may offer the right compromise between achieving good network utilization and good service to the user in the case of high bandwidth calls and bursty traffic. In particular, we show that the ability to delay requests substantially improves the robustness of the call admission algorithm in the sense that it increases the ability of the algorithm to cope with high bandwidth calls and very bursty request sequences while maintaining good network utilization. The amount of delay a call will experience is rather short and for a given network, routing strategy, and request sequence, the number of delayed calls for GREEDY with delay is within a small factor of the number of rejected calls for GREEDY with rejection.

Under the same conditions we observe that given that a network provider wants to achieve a certain network utilization the percentage of calls that are delayed is at most slightly larger than the percentage of calls that are rejected. A GREEDY call admission algorithm that combines delay with a more aggressive routing strategy easily outperforms GREEDY with rejection and fixed path routing call admission algorithm.

Summary

While some traffic (especially real-time video and audio) requires stringent bandwidth guarantees, other traffic (file transfers and electronic mail) offers some flexibility and generally a best effort guarantee is sufficient. Therefore each connection is usually specified to belong to either the group of guaranteed bandwidth traffic or to the group of best-effort traffic. This work on call admission is concerned with requests belonging to the group of guaranteed traffic. If the algorithms discussed in this thesis are used to schedule guaranteed traffic and an appropriate congestion control scheme [KMCL93, KC93, CKB94] is used, best-effort traffic can be used to further increase the utilization of the network.

History shows that, no matter how many resource are provided, humans will find a way to use them. The traffic study has shown us that traffic on tomorrow's networks will be very bursty. Even large bandwidth reserves will only provide limited protection against bursts, and congestion will occur. We show in theory, using competitive analysis, and in practice, using simulations, that call admission algorithms that may delay calls are superior to those that have to reject calls.

Bibliography

[All93] Anthony Alles. ATM in private networking, a tutorial. HUGHES LAN SYSTEMS, 1993.

[All95] Anthony Alles. ATM internetworking, May 1995. An abridged version was presented at Engineering InterOp, Las Vegas, March 1995.

[CKB94] A. Chapman, H.T. Kung, and T. Blackwell. Use of flow control for effective statistical multiplexing and notes on implementation. ATM Forum Contributions 94-0085, January 1994.

[Fag75] M.D. Fagen, editor. *A History of Engineering and Science in the Bell System, The Early Years (1875 – 1925)*. Bell Telephone Laboratories, Incorporated, 1975.

[For93] ATM Forum. ATM User-Network Interface Specification - Version 2.1, May 1993.

[JK75] Norman L. Johnson and Samuel Kotz. *Distributions in Statistics: Continuous Univariate Distributions – 1* . Wiley-Interscience Publication, 1975.

[Jol67] E.H. Jolley. *Introduction to telephony and telegraphy*. Hart Publishing Company, INC., New York City, 1967.

[KC93] H.T. Kung and A. Chapman. The FCVC (flow controlled virtual channels) proposal for ATM networks, 1993. A summary appears in Proc. 1993 International Conf. on Network Protocols, San Francisco, California, October 19-22, 1993, pp. 116-127.

[KMCL93] H.T. Kung, R. Morris, T. Charuhas, and D. Lin. Use of link-by-link flow control in maximizing ATM network performance: Simulation results. In *Proceedings IEEE Hot Interconnects Symposium*, Palo Alto, California, August 1993.

[MW73] Arne Myskja and Odd O. Walmann. An invetigation of telephone-user habits by means of computer techniques. *IEEE Transactions on Communications*, 21(6):663–671, 1973.

[Rey88] R.F. Rey, editor. *Engineering and Operations in the Bell System, Telecommunications in the Bell System in 1982 – 1983*. Bell Telephone Laboratories, Incorporated, 1988.

[Sch82] G.E. Schindler, editor. *A History of Engineering and Science in the Bell System, Switching Technology (1925 – 1975)*. Bell Telephone Laboratories, Incorporated, 1982.

[ST85] Daniel D. Sleator and Robert E. Tarjan. Amortized efficiency of list update and paging rules. *Communications of the ACM*, 28(2):202–208, 1985.

[Sys60] R. Syski. *Introduction to Congestion Theory in Telephone Systmes*. Oliver and Boyd, Edinburgh and London, 1960.