Random access techniques for data transmission over packet-switched radio channels*

by LEONARD KLEINROCK and FOUAD TOBAGI
University of California
Los Angeles, California

INTRODUCTION

Terminal access to computer systems has long been and continues to be a problem of major significance. We foresee an increasing demand for access to data processing and storage facilities from interactive terminals, point-of-sales terminals, real-time monitoring terminals, hand-held personal terminals, etc. What is it that distinguishes this problem from other data communication problems? It is simply that these terminals tend to generate demands at a very low duty cycle and are basically *bursty* sources of data; in addition, these terminals are often geographically distributed. In the computer-to-computer data transmission case, one often sees high utilization of the communication channels; this is just not the case with terminal traffic. Consequently, the *cost* of providing a dedicated channel to each terminal is often prohibitive. Instead, one seeks ways to merge the traffic from many terminal sources in a way which allows them to share the capacity of one or a few channels, thereby reducing the total cost. This cost savings comes about for two reasons: first, because of the economies of scale present in the communications tariff structure; and secondly, because of the averaging effect of large populations which permit one to provide a channel whose capacity is approximately equal to the sum of the *average* demands of the population, rather than equal to the sum of the *peak* demands (i.e., the law of large numbers). This merging of traffic and sharing of capacity has been accomplished in various ways such as: polling techniques, contention systems, multiplexing, concentrating, etc. Many of these are only weak solutions to the problem of gathering low data rate traffic from sources which are geographically dispersed.

In this set of papers,[1–6] we suggest another solution to the terminal access problem, namely that of packet switching over radio channels. In such a system, data terminals package their data into constant length segments known as packets to which is added additional control information such as source and destination address, error control bits, etc. All terminals are assumed to share a common (wide-band) radio channel and to be within range and in line-of-sight of a receiver station. When any terminal generates a packet, that terminal follows some transmission protocol which determines when transmission may take place at which time the packet is transmitted using the full channel bandwidth. Depending upon the protocol, more than one terminal might (unfortunately) transmit in overlapping time intervals, in which case these packets may destructively interfere with each other. Whenever the station receives a packet correctly (as determined by the error control sum check), then an acknowledgment is broadcast to the terminal population, identifying which packet was correctly received. If a terminal receives no acknowledgment after some appropriate timeout interval, then it knows that its packet was "destroyed" and must take some action to cause a retransmission attempt. The key point is that all terminals are simultaneously sharing a single channel; this offers a solution which handles the geographical dispersion of terminals and which at the same time takes advantage of the available cost savings mentioned earlier. Moreover, this solution is highly effective when terminals are mobile (police cars, fire trucks, taxis, ambulances, army vehicles and personnel, etc.) and/or when the environment is itself hostile (natural dangers or man-made dangers).

The use of radio packet switching is relatively new* and has been reported upon in the recent literature. The ALOHA system[7] at the University of Hawaii is not unlike the system we have in mind, and the description of experience with this system as it impacts the current study is described in these proceedings.[2] In 1973, a series of papers describing the use of packet switching in satellite radio channels was published in these proceedings;[8–10] the satellite problem is very similar to the terminal radio problem, with the key distinction being the enormous difference in the propagation delay (roughly ¼ second for a stationary satellite as opposed to small fractions of a millisecond for line-of-sight ground radio).

The Advanced Research Projects Agency of the Department of Defense, recently undertook a new effort whose goal is to develop new techniques for packet radio communication

---

* This work was supported by the Advanced Research Projects Agency of the Department of Defense (DAHC-15-73-0368).

* On the other hand, digital (pulse) systems using radio propagation are not new—e.g., telegraphy, radar, etc. Here we restrict our comments to addressed packets. The most well known example of a packet switched wire network is the ARPANET.[11]

among geographically distributed, fixed or mobile, user terminals and to provide improved frequency management strategies to meet the critical shortage of r.f. spectrum. The research presented in this paper is an integral part of the total design effort of this system which encompasses many other research topics. A number of these are considered in this set of papers. In this paper, we are concerned with *one* aspect of design and analysis, namely the consideration of various random access protocols, their behavior, and the difficult problem of controlling a channel which must carry its own control information. Specifically, we do *not* investigate the networking issues when radio relays (repeaters) are required to extend the range of the terminals; such issues (layout, routing, etc.) are dealt with in References 1 and 4. We consider an environment in which all terminals are within radio range and line-of-sight of a common receiver station. One of the first protocols studied in conjunction with ground radio and satellite packet switching was "pure ALOHA" as mentioned above. In this mode, users are permitted to transmit any time they desire. If they receive an acknowledgment within some predetermined time-out period, then they know their transmission was successful. Otherwise they assume a multi-access collision occurred and they must retransmit. To avoid the same collision again (and forever!) any one of many schemes may be used for introducing a random retransmission delay, thereby spreading the conflicting packets over time. It is known that the maximum fraction of successful packet transmissions on the average is simply $\frac{1}{2}e(\approx 18$ percent) for random ALOHA.[7] This is abominably small compared to the maximum of 100 percent successful if transmission were perfectly scheduled to avoid all collisions. A second method for using the radio channel is to modify the completely unsychronized use of the ALOHA channel by "slotting" time into segments whose duration is exactly equal to the transmission time of a single packet (assuming constant length packets). If we require each user to start his packets only at the beginning of a slot, then when two packets conflict, they will overlap completely rather than partially, providing an increase in channel efficiency. This method is referred to as "slotted ALOHA."[9] The optimum performance of this system is twice that of random ALOHA, namely $1/e(\approx 37$ percent); this is still poor. Not only is the *capacity* of the ALOHA channels wanting, but so too is the average delay $\mathfrak{D}$ until successful transmission; we give the throughput-delay characteristic later in Figure 7.

Let us compare slotted ALOHA to Frequency Division Multiple Access (FDMA) which is a common method for partitioning a channel into a given number of separate subchannels which are assigned on a point-to-point basis between user pairs; synchronous Time Division Multiple Access (TDMA) is equivalent to FDMA so far as we are concerned here (we neglect guard bands). The fixed channel assignment in FDMA is effective in preventing collisions but succeeds in this at the expense of possibly poor utilization of each channel since the smoothing effect of a large population is absent. To analyze FDMA, we adopt the following assumptions: (a) an assumed finite (but large) population of M users; (b) each user generates a new fixed length packet (of b_m bits) according to a Poisson process at a rate Λ per second; (c) the total channel has a bandwidth of W hertz modulated at 1 bit/hertz-sec (giving a channel capacity of W bits/sec). Thus, with M users in this FDMA mode, each is assigned a channel of W/M bits/sec. Each such channel behaves as an $M/D/1$ queueing system giving an average time in system $\mathfrak{D}$ (waiting plus transmission) as follows:[12]

$$\mathfrak{D}=\frac{\rho\left(1-\frac{\rho}{2}\right)}{1-\rho} \tag{1}$$

where $\rho=Mb_m/W$.

We are assuming that queueing is permitted at each terminal. However, the analysis for slotted ALOHA assumes an infinite population of users with an aggregate input rate of $M\Lambda$ packets per second and this produces an upper bound on delay. (We note that a finite population model with M users at rate Λ and with queueing permitted will produce fewer collisions than the infinite population would since each terminal will avoid conflicts among its own packets).

Equation (1) for FDMA is compared with the results for delay in slotted ALOHA with an infinite population (see Reference 8 and Figure 7 below) as follows. We consider the (M, Λ) plane in Figure 1, in which we represent constant $\mathfrak{D}$ contours. Comparing the delay performance of the two systems, we note that when we are in presence of bursty users (small Λ), slotted ALOHA can support many more users than FDMA, for the same packet delay. For example, at $\mathfrak{D}=0.1$ sec, slotted ALOHA can support a number of users which is over 3 orders of magnitude greater than the number that FDMA can support when $\Lambda=10^{-3}$ packet/sec; as Λ increases (i.e., as the burstiness decreases), this difference reduces until at $\Lambda\approx 5$ the two systems can support roughly an equal number of users. Beyond this point, FDMA is superior. This crossover point clearly depends upon the value of $\mathfrak{D}$ examined. In fact, slotted ALOHA can support total traffic only in the range $M\Lambda b_m/W<1/e\cong.37$ and beyond

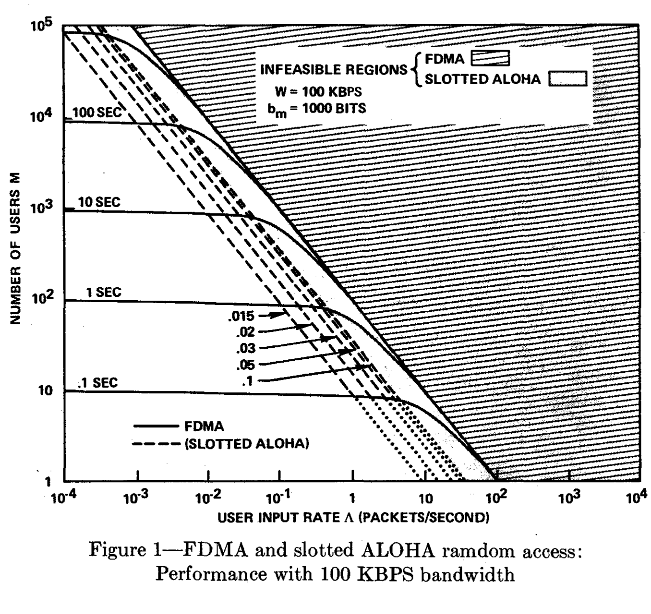

Figure 1—FDMA and slotted ALOHA ramdom access: Performance with 100 KBPS bandwidth

that, FDMA will always be superior until it too saturates at $M\Lambda b_m/W=1$; this tradeoff is clearly evident in the curves of Reference 10.

The above result can be alternatively presented in the following manner. Let M be some large number, say 1000. Figure 2 shows constant $\mathfrak{D}$ contours in the (W, Λ) Plane. Again we note that if we are in presence of bursty users, in order to achieve the same small delay, FDMA requires a bandwidth larger than slotted ALOHA by as much as three orders of magnitude. This factor is exactly equal to M as $\Lambda \rightarrow 0$ since in this region queueing effects are insignificant; in this limit the delay $\mathfrak{D}$ is simply the packet transmission time (observe the flatness of the curves in Figures 1 and 2), which for FDMA is $\mathfrak{D}=Mb_m/W$ and for slotted ALOHA is $\mathfrak{D}=b_m/W$. It is also obvious here, for the same total bandwidth W, that FDMA will give M times the delay as compared to slotted ALOHA. This gain diminishes as Λ increases, until finally as $M\Lambda b_m/W \rightarrow 1/e$ the situation reverses as mentioned above.

Finally, let us fix Λ and consider the delay contours in the (W, M) plane. Figure 3 corresponds to $\Lambda=10^{-1}$ packets per second. Such input rates correspond again to bursty users. We note again that in order to support a large number of users, FDMA requires a larger bandwidth for the same delay performance.

It is all too evident from the above comparison that random access is by far superior to FDMA or TDMA when the environment consists of large populations of bursty users. However, we note that slotted ALOHA itself does not use the channel as efficiently as we might hope and this prompts one to inquire as to other, superior, protocols; such an inquiry is the subject of this paper. Following we consider two random access modes which we refer to as "Carrier Sense Multiple Access" (CSMA) and "Split-channel Reservation Multiple Access" (SRMA).

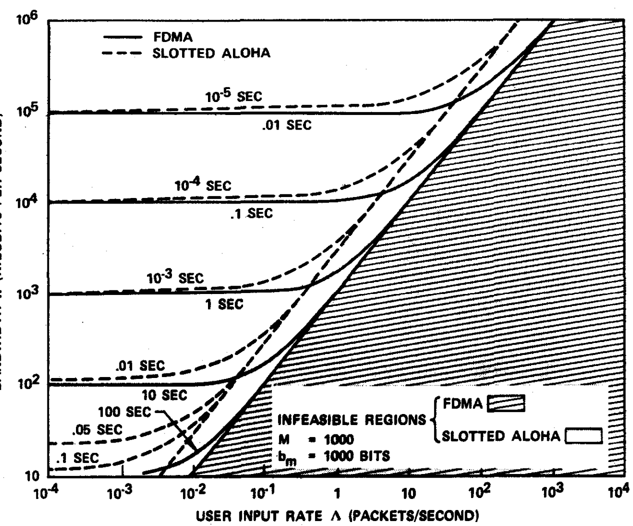

Figure 2—FDMA and slotted ALOHA ramdom access: Bandwidth requirements for 1000 terminals

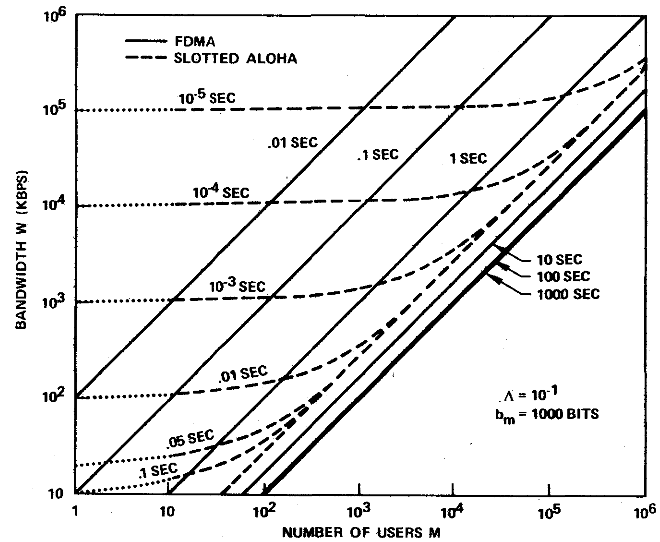

Figure 3—FDMA and slotted ALOHA ramdom access: Performance for $\Lambda=10^{-2}$ packets per second

CARRIER SENSE MULTIPLE ACCESS MODES

The radio channel considered in this paper is characterized as a wideband channel with a propagation delay between any source-destination pair which is very small compared to the packet transmission time.* This suggests a new approach for using the channel; namely, the Carrier-Sense Multiple Access (CSMA) mode. In this scheme one attempts to avoid collisions by listening to (i.e., "sensing"†) the carrier due to another user's transmission. Based on this information about the state of the channel, one may think of various actions the terminal may take. Three protocols will be considered which we call "persistent" CSMA protocols: the 1-Persistent, the Non-Persistent, and the p-Persistent CSMA. Below, we present the protocols, and display the throughput-delay performance for each.

In this paper we omit the proofs for conciseness and clarity of presentation; the details of these proofs are to be found in a series of forthcoming papers.[13–15]

CSMA transmission protocols and system assumptions

The various protocols considered below differ from one another by the action (pertaining to packet transmission)

* Consider, for example, 1000 bit packets transmitted over a channel operating at a speed of 100 Kilobits per second. The transmission time of a packet is then 10 mseconds. If the maximum distance between the source and the destination is 10 miles, then the (speed of light) packet propagation delay is of the order of 54 μseconds. Thus the propagation delay is a very small fraction (a = 0.005) of the transmission time of a packet. On the contrary, when one considers satellite channels [8] the propagation delay is a relatively large multiple of the packet transmission ($a \gg 1$).

† Sensing carrier prior to transmission is a well-known concept in use for (voice) aircraft communication. In the context of packet radio channels it was originally suggested by D. Wax of the University of Hawaii in an internal memorandum dated March 4, 1971.

that a terminal takes after sensing the channel. However, in all cases, when a terminal determines (by the absence of a positive acknowledgment) that its transmission was unsuccessful, then it reschedules the transmission of the packet according to a randomly distributed retransmission delay. At this new point in time, the transmitter senses the channel and repeats the algorithm dictated by the protocol. At any instant a terminal is called a *ready terminal* if it has a packet ready for transmission at this instant (either a new packet just generated or a previously conflicted packet rescheduled for transmission at this instant).

A terminal may, at any one time, either be transmitting or receiving (but not both simultaneously). However, the delay incurred to switch from one mode to the other is negligible. All packets are of constant length and are transmitted over an assumed noiseless channel (i.e., the errors in packet reception caused by random noise are not considered to be a serious problem and are neglected in comparison with errors caused by overlap interference). The system assumes non-capture (i.e., the overlap of any fraction of two packets results in destructive interference and both packets must be retransmitted). We further simplify the problem by assuming the propagation delay τ (small compared to the packet transmission time) to be identical* for all source-destination pairs.

1-Persistent CSMA

The 1-Persistent CSMA protocol is devised in order to (presumably) achieve acceptable throughput by never letting the channel go idle if some ready terminal is available. More precisely, a ready terminal senses the channel and operates as follows:

- If the channel is sensed idle, it transmits the packet with probability one.
- If the channel is sensed busy, it waits until the channel goes idle (i.e., *persisting* on transmitting) and only then transmits the packet (with probability one—hence, the name 1-Persistent).

A slotted version of the 1-Persistent CSMA can be considered in which the time axis is slotted and the slot size is τ seconds (the propagation delay). All terminals are synchronized and are forced to start transmission only at the beginning of a slot. When a packet's arrival occurs during a slot, the terminal senses the channel at the beginning of the next slot and operates according to the protocol described above.

Non-Persistent CSMA

While the previous protocol was meant to make "full" use of the channel, the idea here is to limit the interference among packets by always rescheduling a packet which finds the channel busy upon its arrival. On the other hand, this scheme may introduce idle periods between two consecutive non-overlapped transmissions. More precisely, a ready terminal senses the channel and operates as follows:

- If the channel is sensed idle, it transmits the packet.
- If the channel is sensed busy, then the terminal schedules the retransmission of the packet to some later time according to the retransmission delay distribution. At this new point in time, it senses the channel and repeats the algorithm described.

A slotted version of this Non-Persistent CSMA can also be considered by slotting the time axis and synchronizing the transmission of packets in much the same way as for the previous protocol.

p-Persistent CSMA

The two previous protocols differ by the probability (one or zero) of not rescheduling a packet which upon arrival finds the channel busy. In the case of a 1-Persistent CSMA, we note that whenever two or more terminals become ready during a transmission period, they wait for the channel to become idle (at the end of that transmission) and then they all transmit with probability one. A conflict will also occur with probability one! The idea of randomizing the starting times of transmission of packets accumulating at the end of a transmission period suggests itself for interference reduction and throughput improvement. The scheme consists of including an additional parameter p, the probability that a ready packet persists ($1-p$ being the probability of delaying transmission by τ seconds). The parameter p will be chosen so as to reduce the level of interference while keeping the idle periods between any two consecutive non-overlapped transmissions as small as possible.

More precisely, the protocol consists of the following: the time axis is slotted where the slot size is τ seconds. For simplicity of analysis, we consider the system to be synchronized such that all packets begin their transmission at the beginning of a slot.

Consider a ready terminal:

- If the channel is sensed idle, then
  - —with probability p, the terminal transmits the packet.
  - —with probability $1-p$, the terminal delays the transmission of the packet by τ seconds (i.e., one slot). If at this new point in time, the channel is still detected idle, the same process above is repeated; otherwise, some packet must have started transmission, and our terminal schedules the retransmission of the packet according to the retransmission delay distribution (i.e., acts as if it had conflicted and learned about the conflict).
- If the ready terminal senses the channel busy, it waits until it becomes idle (at the end of the current transmission) and then operates as above.

* By considering this constant propagation delay equal to the largest possible, one gets lower (i.e., pessimistic) bounds on performance.

Note that 1-Persistent is the special case of p-Persistent with $p=1$.

Throughput equations

We assume that our traffic source consists of a very large number M of users who collectively can be approximated by an independent Poisson source with an aggregate mean packet generation rate of λ packets/second. This implies that each user will generate packets infrequently and each packet can be successfully transmitted in a time interval much less than the average time between successive packets generated by a given user.

In addition, we characterize the traffic as follows. We have assumed that each packet is of constant length requiring T seconds for transmission. Let $S=\lambda T$. S is the average number of new packets generated per transmission time, i.e., the input rate normalized with respect to T. If we were able to perfectly schedule the packets into the available channel space with absolutely no overlap or space between the packets, we would have $S=1$; therefore, we also refer to S as the *channel utilization*, or *throughput*. The maximum achievable throughput for an access mode is called the *capacity* of the channel under that mode.

Each user delays the transmission of a previously collided packet by some random time (introduced to avoid repeated conflicts) whose mean is $\bar{X}$ (chosen, for example, uniformly between 0 and $X_{\max}=2\bar{X}$). Since conflicts can occur, the traffic offered to the channel from our collection of users consists of new packets and previously collided packets. This increases the mean offered traffic rate to G packets per transmission time T, where $G \geq S$.

Our two further assumptions are:

(A1) The average retransmission delay $\bar{X}$ is large compared to T.

(A2) The interarrival times of the point process defined by the start times of all the packets plus retransmissions are independent and exponentially distributed.

We wish to solve for the channel capacity of the system for all of the access protocols described above. This we do by expressing S in terms of G (as well as other system parameters). The channel capacity is obtained by maximizing S with respect to G. Note that S/G is merely the probability of a successful transmission and G/S is the average number of times a packet must be transmitted or scheduled until success.

The basic equations for the throughput S are expressed in terms of a (the ratio of propagation delay to packet transmission time) and G (the offered traffic rate) as follows:*

1-Persistent CSMA

$$S=\frac{G[1+G+aG(1+G+aG/2)]e^{-G(1+2a)}}{G(1+2a)-(1-e^{-aG})+(1+aG)e^{-G(1+a)}} \tag{1}$$

* For proofs, the reader is referred to Reference 13.

Slotted 1-Persistent CSMA

$$S=\frac{Ge^{-G(1+a)}[1+a-e^{-aG}]}{(1+a)(1-e^{-aG})+ae^{-G(1+a)}} \tag{2}$$

Non-Persistent CSMA

$$S=\frac{Ge^{-aG}}{G(1+2a)+e^{-aG}} \tag{3}$$

Slotted Non-Persistent CSMA

$$S=\frac{aGe^{-aG}}{(1+a)(1-e^{-aG})+a} \tag{4}$$

p-Persistent CSMA

$$S(G,p,a)=\frac{(1-e^{-aG})[P_s'\pi_0+P_s(1-\pi_0)]}{(1-e^{-aG})[a\bar{t}'\pi_0+a\bar{t}(1-\pi_0)+1+a]+a\pi_0} \tag{5}$$

where P_s', P_s, $\bar{t}'$, $\bar{t}$ and π_0 are defined in Reference 13. We note that

$$S(G, p\to 0, a=0)\to\frac{G}{G+e^{-G}}$$

In Figure 4 for $a=0.01$, we plot S versus G for the various access modes introduced so far and show the relative performance of each. We also summarize these results in the following table:

PROTOCOL	CAPACITY C
Pure ALOHA	0.184
Slotted ALOHA	0.368
1-Persistent CSMA	0.529
Slotted 1-Persistent CSMA	0.531
0.1-Persistent CSMA	0.791
Non-Persistent CSMA	0.815
0.03-Persistent CSMA	0.827
Slotted Non-Persistent CSMA	0.857
Perfect Scheduling	1.000

While the capacity of ALOHA channels does not depend on the propagation delay, the capacity of a CSMA channel does. An increase in a increases the "vulnerable" period of a packet and reduces its capacity. This also results in "older" channel state information from sensing. In Figure 5 we plot, versus a, the channel capacity for all of the above random access modes. For large a, we note that slotted ALOHA (and even "pure" ALOHA) is superior to any CSMA mode since decisions based on partially obsolete data are deleterious; this effect is due in part to our assumption about the constant propagation delay.

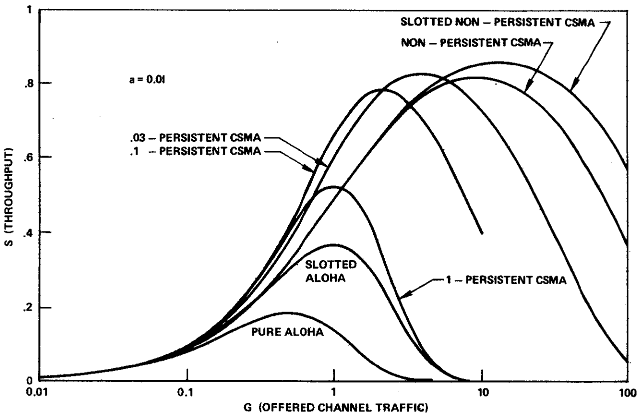

Figure 4—Throughput for the various random access modes ($a=0.01$)

Delay performance

We introduce at this point the expected packet delay $\mathfrak{D}$ defined as the average time from when a packet is generated until it is successfully received. Our principal concern in this section is to investigate the tradeoff between the average delay and the throughput S.

For the present study, it is assumed that the acknowledgment packets are always correctly received with probability one. The simplest way to accomplish this is to create a separate channel to handle acknowledgment traffic. If sufficient bandwidth is provided, overlaps between acknowledgment packets are avoided, since a positive acknowledgment packet is created only when a packet is correctly received, and there will be at most one such packet at any given time. Thus, if T_a denotes the transmission time of the acknowledgment packet on the separate channel, then the time-out for receiving a positive acknowledgment is $T+\tau+T_a+\tau$, provided that the processing time needed to perform the sumcheck and to generate the acknowledgment packet is assumed negligible.

The Delay $\mathfrak{D}$ is a function of S and $\bar{X}$. Thus, for each S,

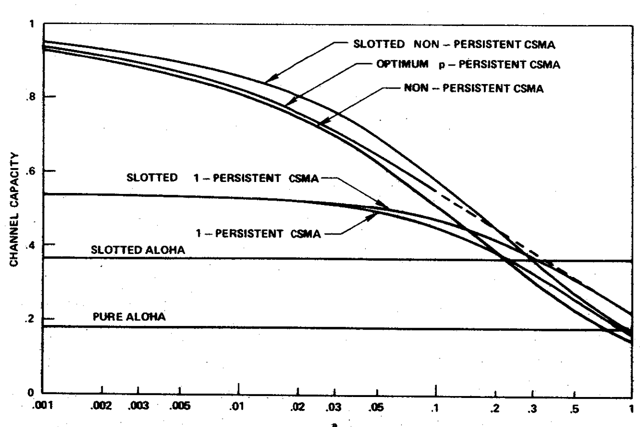

Figure 5—CSMA and ALOHA:
Effect of propagation delay on channel capacity

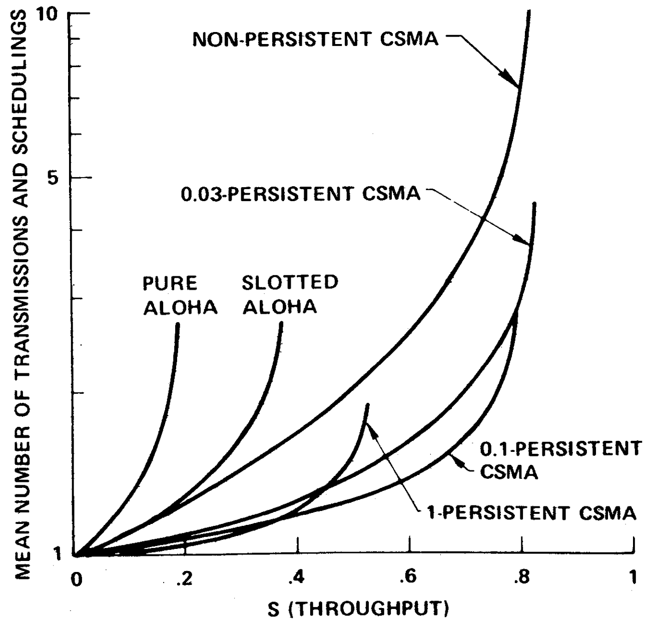

Figure 6—CSMA and ALOHA:
G/S versus throughput ($a=0.01$)

a minimum delay can be achieved by choosing an optimal $\bar{X}$. Such an optimization problem is difficult to solve analytically, and simulation techniques have been employed.

Before we proceed with the discussion of the simulation results, we compare the various access modes in terms of the average number of transmissions (or average number of schedulings) G/S. For this purpose, we plot G/S versus S in Figure 6 for the ALOHA and CSMA modes, when $a=0.01$. Note that CSMA modes provide lower values for G/S than the ALOHA modes. Furthermore, for each value of the throughput, there exists, a value of p such that p-Persistent is optimal. For small values of S, $p=1$ (i.e., 1-Persistent) is optimal. As S increases, the optimal p decreases.

Simulation results

The simulation model is based on all system assumptions presented above. However, we relax the assumptions concerning the retransmission delay and the independence of arrivals for the offered channel traffic.

In general, our simulation results indicate the following:

(1) For each value of the input rate S, there is a minimum value $\bar{X}$ for the average retransmission delay variable, such that below that value, it is impossible to achieve a throughput equal to the input rate. The higher S is, the larger $\bar{X}$ must be to prevent a constantly increasing backlog, i.e., to prevent the channel from saturating. In other words, the maximum achievable throughput (under stable conditions) is a function of $\bar{X}$, and the larger $\bar{X}$ is, the higher is the maximum throughput.

(2) Recall that the throughput equations were based on

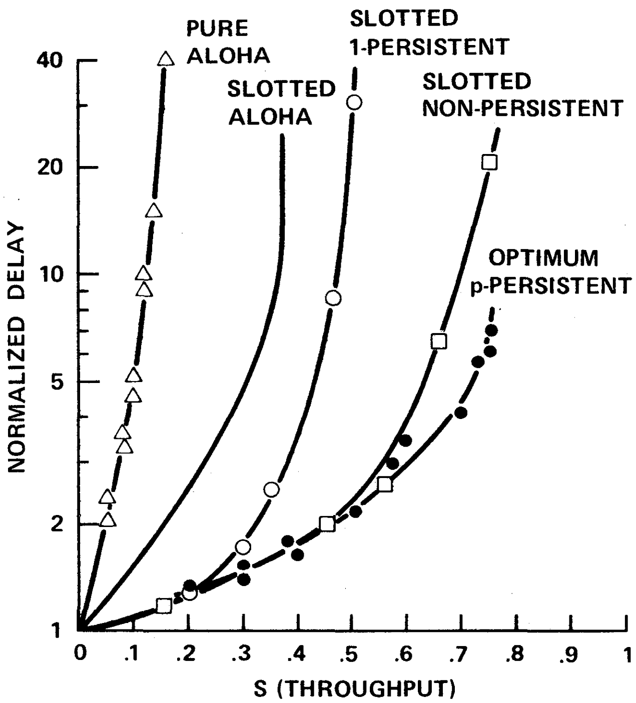

Figure 7—CSMA and ALOHA: Throughput-delay trade-offs from simulation ($a=0.01$)

the assumption that $\bar{X}$ is infinitely large compared to T. Simulation shows that for finite values of $\bar{X}$, larger than some value $\bar{X}_0$ but not too large compared to T, the system already "reaches" the asymptotic results ($\bar{X}\rightarrow\infty$), i.e., for some finite values of $\bar{X}$, assumption (A2) is satisfied and delays are acceptable. Simulation experiments were conducted to find the optimal delay, that is, the value of $\bar{X}(S)$ which allows one to achieve the indicated throughput with the minimum delay.

Finally, in Figure 7, we give the throughput-minimum delay trade-off for the three Carrier Sense Multiple Access modes and $a=0.01$. This is the basic performance curve.

THE EFFECT OF HIDDEN TERMINALS ON CHANNEL CAPACITY FOR CARRIER SENSE MULTIPLE ACCESS

The performance obtained in the previous section (in terms of channel capacity and throughput-delay trade-offs) was based on the (strong) assumption that all terminals were in line-of-sight and within range of each other. There are many instances where this is not the case, forcing us to relax that assumption. Two terminals can be within range of the station but out-of-range of each other, or, they can be separated by some physical obstacle opaque to UHF radio signals. Two such terminals are then said to be "hidden" from each other. It is evident that the existence of hidden elements in an environment affects (degrades) the performance of CSMA. In this section we discuss this effect. (For simplicity, we restrict our study to 1-Persistent and Non-Persistent CSMA protocols only.)

Definitions and representation of configurations with hidden elements

In the sequel, terminals are in line-of-sight and within range of the station, but not necessarily with respect to each other. By definition, terminal i "hears" (is connected to) terminal j if i and j are within range and in line-of-sight of each other. In order to represent terminal configurations with hidden elements, it is advantageous to partition the population into several groups (say N) such that all terminals in a group hear exactly the same subset of terminals in the population. (This partitioning is easily formed if we know the hearing matrix of the population. See References 14 and 15). Let $h(i)$ be the set of groups that group i can hear.

We shall further assume that each group i consists of a large number of users who collectively form an independent Poisson source with an aggregate mean packet generation rate λ_i packets per second such that $\sum_{i=1}^{N}\lambda_i=\lambda$. Let $S_i=\lambda_i T$ and $S=\lambda T=\sum_{i=1}^{N} S_i$; S is the total throughput of the channel.

Let $\mathbf{S}=(S_1, S_2, \ldots, S_N)$.

We can write $\mathbf{S}$ as $\mathbf{S}=S\mathbf{U}$ such that

$$u_i\geq 0\,\forall i$$

and

$$\|\mathbf{U}\|\triangleq\sum_{i=1}^{N}u_i=1$$

(The vector $\mathbf{U}$ describes a direction in N-dimensional space.) The *capacity* of the channel along the direction $\mathbf{U}$ is defined as

$$C(\mathbf{U})=\underset{0\leq S\leq 1}{\text{Maximum}}\ S$$

such that the set of inputs determined by the vector $\mathbf{S}(\mathbf{U})$ is achievable. Equivalently, we say that a set $\mathbf{S}(\mathbf{U})$ of input rates is feasible if and only if

$$\mathbf{S}(\mathbf{U})\leq C(\mathbf{U})$$

Let G_i denote the mean offered traffic rate of group i ($G_i\geq S_i$). Let $\mathbf{G}=(G_1, G_2, \ldots, G_N)$ and $G=\sum_{i=1}^{N}G_i$. Finally, we consider $\bar{X}$ to be the same for all groups and the assumptions concerning the retransmission delay and the independence of arrivals for the offered traffic to still hold true.

Throughput equations

We recognize that S_i/G_i is merely the probability of success of an arbitrary packet from group i. This quantity is a function of the traffic vector $\mathbf{G}$. By expressing S_i/G_i for each i in terms of $\mathbf{G}$, we obtain a set of equations relating the components of $\mathbf{S}$ to the components of $\mathbf{G}$.

In the case of *independent* groups (i.e., such that terminals in a group do not hear terminals in other groups) for a given $\mathbf{G}$ and under the system and model assumptions stated above, the probability of success of an arbitrary packet from group i is given as follows*

For 1-Persistent CSMA

$$P_{s_i}=\frac{S_i}{G_i}=\frac{[1+G_i+aG_i(1+G_i+aG_i/2)]}{(1+aG_i)e^{-G_i(1-2a)}}$$

$$\prod_{j=1}^{N}\frac{(1+aG_j)e^{-2G_j}}{G_j(1+2a)-(1-e^{-aG_j})+(1+aG_j)e^{-G_j(1+a)}} \tag{6}$$

For Non-Persistent CSMA

$$P_{s_i}=\frac{S_i}{G_i}=e^{G_i(1-2a)}\prod_{j=1}^{N}\frac{e^{-G_j(1-a)}}{G_j(1+2a)+e^{-aG_j}} \tag{7}$$

This set of equations relates the components of the input vector $\mathbf{S}$ to the components of the traffic vector $\mathbf{G}$. For a given input vector $\mathbf{S}$, we can numerically solve for G_i, $i=1, \ldots, N$. This we do by writing the above equations in the form

$$G_i=S_i/f_i(G_1, \ldots, G_N)$$

where f_i is a function of the vector $\mathbf{G}$, and by solving the set of equations iteratively, starting with the initial values $\mathbf{G}=\mathbf{S}$. If the input vector is a feasible one, then the iterative procedure will result in a (finite) traffic vector $\mathbf{G}$, satisfying the above set of equations. Thus the convergence of the iterative procedure determines the feasibility of the input vector $\mathbf{S}$ and the final values G_i/S_i, $i=1, 2, \ldots, N$ give the average number of transmissions and schedulings a packet from group i undertakes before success. This will be our measure of relative performance of the various groups. Some simple examples are treated in the following section.

In the case of *dependent* groups, similar but approximate relationships can be found for the Non-Persistent CSMA protocol. They are expressed as

$$S_i=G_i\frac{\prod_{j\in h(i)}e^{-aG_j'}\prod_{k\in h(i)}e^{-G_k'(1-a)}}{\prod_{l=1}^{N}[G_l'(1+2a)+e^{-aG_l'}]} \tag{8}$$

$$G_i'=G_i\prod_{\substack{j\in h(i)\\ j\neq i}}\frac{1+aG_j'}{G_j'(1+2a)+e^{-aG_j'}} \tag{9}$$

Examples

Here we consider some typical examples of independent groups to which we apply the analytical results found above. Simulation techniques have been used to check the validity of the assumptions on which the analysis was based. We restrict ourselves to $a=0.01$.

* See References 14 and 15 for proof.

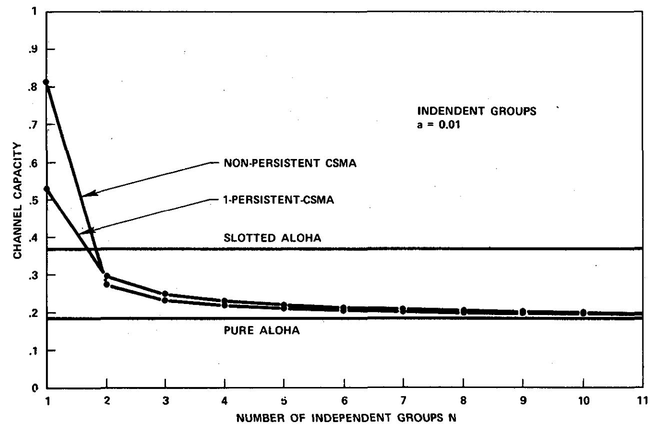

Figure 8—Independent group case: Channel capacity versus the number of groups

Independent Groups Case—A Symmetric Configuration

The population is partitioned into N groups of equal size. For each terminal there exists a fraction β of the population which is hidden, namely $\beta=(N-1)/N(\geq 0.5)$. The channel capacity for various values of N is plotted in Figure 8. Note that the channel capacity experiences a drastic decrease between the two cases: $N=1$ (no hidden terminals, $\beta=0$) and $N=2(\beta=0.5)$. For $N\geq 2$, slotted ALOHA performs better than CSMA. This decrease is more critical for the Non-Persistent CSMA than for the 1-Persistent CSMA as shown in the Figure. For $N>2$, the channel capacity is rather insensitive to N and approaches pure ALOHA for large N.

Independent Groups Case—Complementary Couple Configuration

The previous example did not show the effect of a small fraction of the population being hidden from the rest. In this example the population consists of two independent groups ($N=2$) of unequal sizes such that $\mathbf{U}=(\alpha, 1-\alpha)$ that is

$$S_1=\alpha S$$

$$S_2=(1-\alpha)S$$

Equations (6) and (7) are readily applicable. The channel capacity is plotted versus α for both CSMA protocols in Figure 9. Here again we note that the capacity decreases rapidly as α increases from 0. This decrease is much more critical for the Non-Persistent than for the 1-Persistent. As soon as $\alpha=10^{-2}$, the capacity of Non-Persistent CSMA is only 0.5, as compared to 0.82 when $\alpha=0$. In addition, CSMA performs (capacity-wise) only as good as slotted ALOHA as soon as $\alpha=0.08$ for the Non-Persistent protocol and $\alpha=0.1$ for the 1-Persistent protocol. In both cases, we note that the minimum capacity is obtained for $\alpha=0.5$; this corresponds to the case $N=2$ in the previous example.

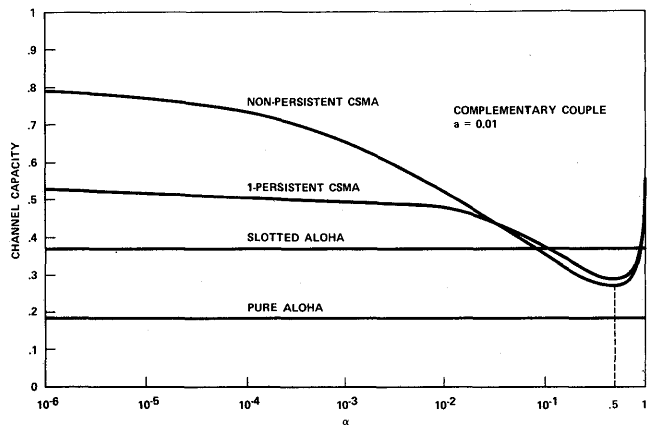

Figure 9—Complementary couple configuration: Channel capacity versus α

In addition, we simulated the 1-Persistent CSMA case for this example and various values of α. The comparison of (S_1, G_1) and (S_2, G_2) relationships obtained from simulation to the results obtained from the analytical model exhibits an excellent match, thus checking the validity of the model.

Examining G_i/S_i for each group, we noted that the large group always performed better than the smaller one. Although we noted for $\alpha \simeq 0.1$ that 1-Persistent CSMA has a capacity only as great as slotted ALOHA, the average number of transmissions $\alpha(G_1/S_1)+[(1-\alpha)(G_2/S_2)]$ was lower (superior) for the 1-Persistent CSMA than for slotted ALOHA.

CARRIER SENSE MULTIPLE ACCESS WITH A BUSY TONE

System operation

In this section we wish to consider a solution to the hidden terminal problem which we call the **B**usy **T**one **M**ultiple **A**ccess mode (BTMA). The operation of BTMA rests on the assumption that the station is, by definition, within range and in line-of-sight of all terminals. The total available bandwidth is to be divided into two channels: a message channel and a busy tone (BT) channel. As long as the station senses a (terminal) carrier on the incoming message channel it transmits a (sine wave) busy tone signal on the busy tone channel. It is by sensing a carrier on the busy tone channel that terminals determine the state of the message channel. The action pertaining to the transmission of the packet that a terminal takes (again) is prescribed by the particular protocol being used. We shall restrict ourselves to the *Non-Persistent protocol* because of its simplicity in analysis and implementation, as well as its relatively high efficiency as shown above. In CSMA, the difficulty of detecting the presence of a signal on the message channel when this message uses the entire bandwidth is minor and therefore is neglected. It is not so when we are concerned with the (statistical) detection of the (sine wave) busy tone signal on a narrow band channel. The detection time, denoted by t_d, is no longer negligible and must be accounted for. The Non-Persistent BTMA protocol is similar to the Non-Persistent CSMA protocol and corresponds to the following. Whenever a terminal has a packet ready for transmission, it senses the busy tone channel for t_d seconds (the detection time) at the end of which it decides whether the BT signal is present or absent. (t_d is a system parameter and its optimal value is discussed below). If the terminal decides that the BT signal is absent then it transmits the packet, otherwise it reschedules the packet for transmission at some later time incurring a random rescheduling delay; at this new point in time, it senses the BT channel and repeats the algorithm. In the event of a conflict, which the terminal learns about by failing to receive an acknowledgment from the station, the terminal again reschedules the transmission of the packet for some later time, and repeats the above process.

Of interest is first, the determination of the channel capacity under a Non-Persistent BTMA protocol and second, the throughput delay characteristics of the latter. The total available bandwidth being the limiting resource, the problem then reduces to selecting the system parameters in order to achieve the best system performance.

Here we make the same assumptions as above. However while the effect of noise is assumed to be negligible on the message channel, we do account for it in the (narrow band) busy tone channel. Each packet is of constant length requiring T_m seconds for transmission on the message channel. Let $S_m=\lambda T_m$. S_m is the average number of new packets generated per transmission time, i.e., this is the input rate normalized with respect to T_m. Under steady state conditions, S_m can also be referred to as the message channel throughput rate and as the *message channel utilization*. Let ψ be the fraction of the bandwidth assigned to the BT channel. Let $S=(1-\psi)S_m$. S is the *overall channel utilization*. The maximum achievable channel utilization is the *capacity* of the channel.

Signal detection

The detection of the busy tone signal is the problem of detecting a signal of known form in the presence of noise. The useful signal is a given function with some unknown parameters, namely, phase and amplitude.† However the observation (detection) time is usually small compared to the "fluctuation time" of these parameters, and the unknown phase and amplitude can be regarded as constant.

The problem of detecting a signal in a background of random noise is a classical statistical problem involving the choice of one hypothesis from two mutually exclusive hypotheses. This has been extensively studied in the litera-

† Because of the mobility of terminals, the signal fluctuates. Thus we assume it to be of unknown amplitude. In the case of fixed terminals, we may idealize the problem to be that of detecting a signal with known amplitude but unknown phase.

ture.[16] The quality of the decision can be characterized by two probabilities:

D Probability of correct detection (in presence of the signal)
F Probability of incorrect detection or false alarm

The detector at the receiver consists of a filter, an integrator and a threshold decision box. Assuming the step response of the busy tone detect filter to be exponential, and considering the same peak power to be used for the busy tone as for the message on the message channel, then the signal-to-noise ratio (SNR) $\mu(t)$ on the busy tone channel at time t is given by

$$\mu(t) = \mu_m \frac{1-\psi}{\psi} (1-e^{-2\psi W t})^2 \tag{10}$$

where

- μ_m is the SNR of the message on the message channel required for suitable operation (typically $\mu_m = 10$)
- ψ is the fraction of bandwidth assigned to the BT channel
- the time constant of the filter exponential rise is taken to be $\frac{1}{2}\psi W$.

Consider now a signal starting at $t=0$ and terminating at $t=T$. Let $D(t)$ be the probability of correct detection at time t after having observed the channel over t_d seconds (t is the time at which the decision is made). $D(t)$ is determined by (See Reference 16).

$$D(t) = F^{(1/1+\mu(u))} \tag{11}$$

where

$$u = \begin{cases} t & \text{if} \quad 0 \le t \le t_d \\ t_d & \text{if} \quad t_d \le t \le T \\ T-t+t_d & \text{if} \quad T \le t \le T+t_d \end{cases}$$

For $t > T+t_d$, the probability of false alarm is F.

Throughput equation

We wish to solve for the channel capacity, given the system parameters F, ψ, W, b_m, τ, t_d. This we do by solving for S in terms of γ (the traffic rate measured in packets per second) and other parameters. The channel capacity is then found by maximizing S with respect to γ.

Contrary to the CSMA modes the fraction of the population which decides to transmit is a function of time. The analytical approach consists of identifying the busy and idle periods and of determining the condition for a successful transmission over the busy period. To keep the analysis simple, some very minor approximations are made yielding a lower bound on throughput as given in the following equation*

$$S \ge S_l = \frac{b_m}{W} \frac{\exp[-\gamma m(0, T_m)]}{\bar{B}+\bar{I}} \tag{12}$$

It can also be shown that in the limit, when $t_d \to 0$, the channel capacity reduces to

$$S = (1-\psi) \frac{1}{2e} \tag{13}$$

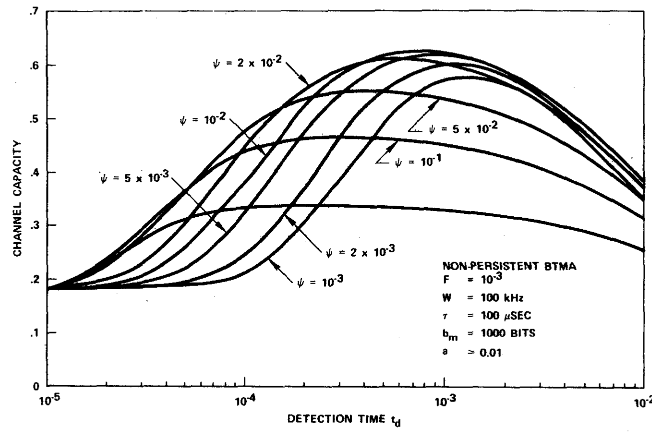

Figure 10—BTMA: Channel capacity versus observation window t_d

Results

The design problem in BTMA consists of maximizing the channel capacity (under the Non-Persistent protocol) by properly selecting the design variables ψ, F and t_d when the number of bits per packet, b_m, and the total available bandwidth W are given. Because of the complicated form of the expressions for S, numerical optimization techniques are used.

To reduce the dimensionality of the problem, and to provide an easy comparison with the previously analyzed CSMA protocols we restrict ourselves to the following:

- τ (maximum propagation delay) = 100 μsec.*
- $\mu_m = 10$
- $\frac{b_m}{W} = 10^2$ msec.†

We consider two cases for b_m and W:

case I: $b_m = 1000$ bits; $W = 10^5$ Hz
case II: $b_m = 10{,}000$ bits; $W = 10^6$ Hz

* See References 14 and 15 for proof and for the definition of $m(0, T_m)$, $\bar{B}$ and $\bar{I}$.
* The bandwidth is assumed to be modulated a 1 bit/Hz-sec.
† This corresponds to a maximum distance of about 20 miles. The ratio of propagation delay to transmission time of a packet, denoted by a, is, in all cases less than (but very close to) or equal to 0.01.

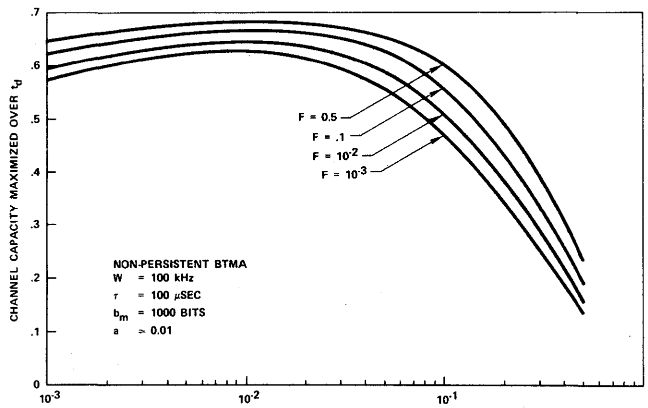

Figure 11—BTMA: Channel capacity (maximized over t_d) versus ψ

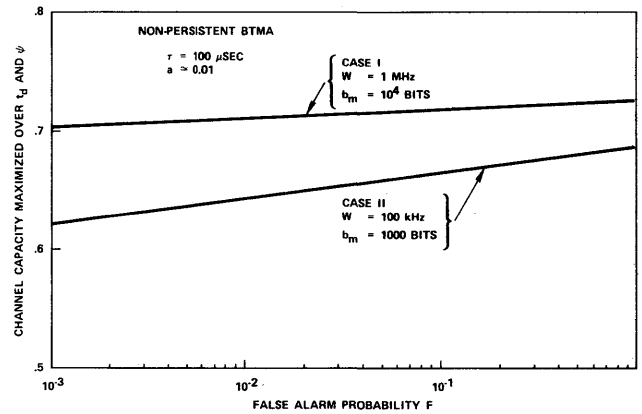

Figure 12—BTMA: Channel capacity (maximized over t_d and ψ) versus F

For $F=10^{-3}$ and various values of ψ we plot in Figure 10 the channel capacity versus the observation window t_d. Similar curves can be plotted for other values of F. For each couple (F, ψ) the channel capacity reaches its maximum at some optimum value of t_d. This optimum is explained by the fact that the larger t_d is, the better is the probability of correct detection $D(t_d)$ when the signal is present during the entire window. However, the larger t_d is, the longer the idle period will be. The effect is reversed as t_d gets smaller.

Note that when the observation window shrinks to 0, the capacity of the channel decreases to $(1-\psi)\frac{1}{2}e$, the capacity provided by the pure ALOHA access mode. Qualitatively speaking $t_d \rightarrow 0$ reduces to very bad detection, and terminals behave in a pure ALOHA mode.

In Figure 11, we plot for various F, the maximum capacity of the channel (maximized over t_d) versus ψ. We note here that the maximum capacity is not very sensitive to *small* variations of ψ. However, there is a certain range of ψ which yields the best performance. For those values of F considered in the graph ($F=10^{-3}$, 10^{-2}, 10^{-1}, 0.5), the optimum ψ is the range $(10^{-2}, 2\times10^{-2})$.

In Figure 12, we plot the capacity (maximized over ψ and t_d) versus F for cases I and II. Note that for both cases the capacity of the channel is a logarithmic function of F. The ultimate performance ($\simeq$0.68 for Case I and $\simeq$0.72 for Case II) is obtained for $F\rightarrow 1$. However, the channel capacity is not very sensitive to variations of F. Case II offers a channel capacity higher than that offered by Case I; we note that this gain does not consider other factors such as increased power requirements.*

To compare the delay performance of BTMA for various values of the system parameters, we first consider the quantity G/S, the average number of transmissions and schedulings that a packet incurs before successful transmission. In Figure 13, we plot, for each value of F, G/S versus S for those values of ψ and t_d yielding the maximum channel capacity. Note that for each value of S there exists a value of F minimizing G/S. However, for relatively small values of S (not too close to the saturation point of the channel) we note that the higher the probability of false alarm F is, the larger is G/S. An explanation can be given by the following fact:

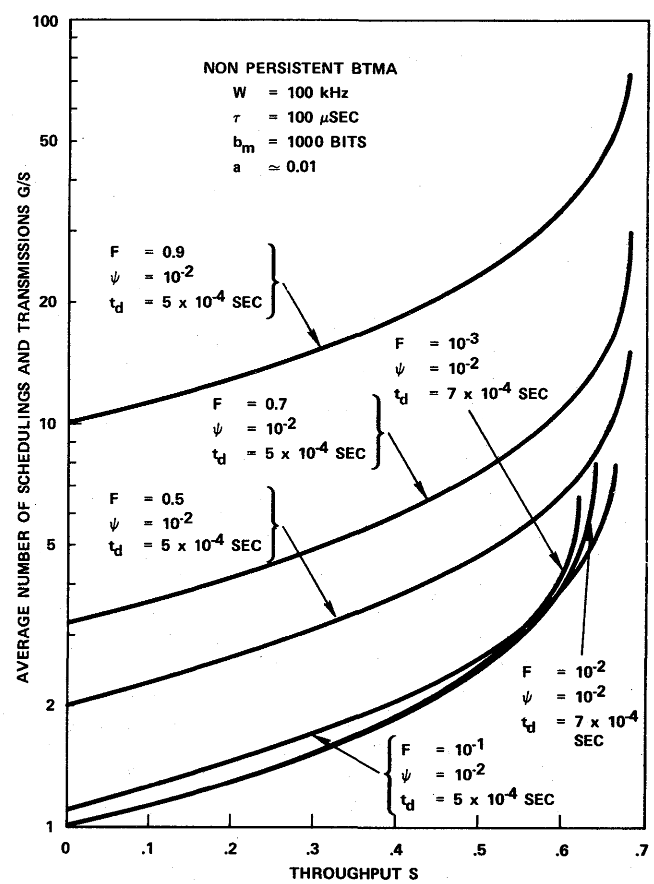

Figure 13—BTMA: Average number of schedulings and transmissions

* The larger the bandwidth is, the better is the correct detection. Thus larger W provides larger channel capacity. However, the channel capacity is always bounded from above by the capacity of CSMA with propagation delay equal to 2τ.[14]

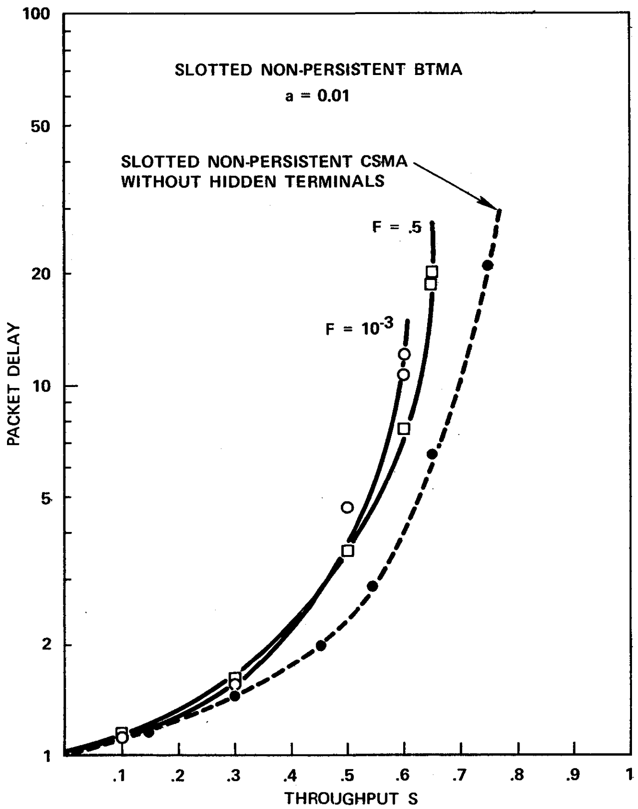

Figure 14—BTMA: Throughput-delay trade-offs ($a=0.01$)

when $G\rightarrow 0$ and $S\rightarrow 0$, the terminal incurs an average number of schedulings and transmissions equal to $1/(1-F)$. This is shown on Figure 13 at $S=0$.

G/S, as a measure of delay, can be of importance since the complexity of the equipment and the implementation of the protocol can be directly related to the number of schedulings and transmissions that a packet incurs. For example, at each scheduling, the terminal has to generate a random number determining the scheduling delay. Of even more importance in evaluating the performance of such a system is the determination of the actual packet delay, defined as the time lapse since the packet is first generated, until the time it is successful. As discussed earlier, the mathematical determination of packet delays is fairly complex, and simulation techniques are employed. For various values of F ($F=10^{-3}$ and $F=0.5$), by selecting the optimum system parameters (ψ, t_d) with respect to channel capacity, we simulated the BTMA mode. In Figure 14 we plot the throughput-minimum-delay* curve for these values of F. It is to be noted that, even though G/S can be significantly affected by F, the minimum delay is relatively insensitive to F. However, for each value of S there exists a value of F which provides the lowest delay. By comparing the lower envelope of these throughput-delay curves to the curve corresponding to the Non-Persistent CSMA without hidden terminals, we note the relatively-good performance of BTMA.

* Delay is minimized with respect to $\bar{X}$. In BTMA, the larger F is, the larger is G/S. The minimum delay is obtained for very small values of $\bar{X}$ since the packet incurs $1/(1-F)$ reschedulings when the channel is idle.

RESERVATION TECHNIQUES

We have shown that, in the presence of a large population of users exhibiting a bursty behavior, FDMA and TDMA produce much higher delays with the same available bandwidth than random multiple access, and in order to achieve the same delay performance, they require a much larger bandwidth; in the latter case, the utilization of the channel is extremely low. In order to increase the channel utilization beyond FDMA and TDMA, statistical multiplexing or Asynchronous Time Division Multiple Access (ATDMA) has been proposed.[17] However, this technique is less attractive in situations where the terminals are geographically spread and/or mobile.

Of more recent interest are "controlled" techniques for transmission from terminals to computer. There are two methods in common usage for wired networks: contention and polling. In a *contention network*, the terminal makes a request to transmit: if the channel is free, transmission goes ahead; if it is not free, the terminal must wait; the station schedules the transmissions either in a prearranged sequence (according to some scheduling scheme) or in the sequence in which the requests were made. In the *polling technique*, the station asks the terminals one by one whether they have anything to transmit. For this, the station may have a polling list giving the order in which terminals are polled. A polling message is sent to the terminal under consideration. If the terminal has some data to transmit, it goes ahead; if not, a negative reply (or absence of reply) is received, and the next terminal is polled.

These controlled techniques are readily applicable to radio networks. They constitute the subject of this section. It has been shown that although polling may allow the system to achieve high utilization of the channel, the delay incurred by a packet is large (mainly for the large M case which is of interest to us) rendering the polling technique less attractive than CSMA and BTMA. The alternative is the use of reservation techniques. In this section, we study the **S**plit-channel **R**eservation **M**ultiple **A**ccess (SRMA) as one implementation of such reservation techniques. The available bandwidth is divided into two parts: one used to transmit control information, the second used for the message itself.

System operation

In the particular scheme considered here, the bandwidth allocated for control is further divided into two channels:

—the request channel
—the answer-to-request channel.

The request channel will be operated in a random access mode (ALOHA or CSMA). Consider a terminal with a message ready for transmission. To initiate the sending of the message, the terminal sends, on the request channel, a request packet containing information about the address of the terminal and, in the case of variable length or multi-packet messages, the length of the message. At the correct reception of the request packet, the scheduling station computes the time at which the message channel will be available and transmits back to the terminal, on the answer-to-request channel, an answer packet containing the address of the terminal and the time at which it can start transmission.

Analysis

The total delay is composed of the two following components:

(i) $\mathfrak{D}_1$, the time for the request packet to be successfully received at the station, and
(ii) $\mathfrak{D}_2$, the time between reception of the request packet at the station and the end of the message transmission.

Let W_m be the bandwidth allocated to the message channel and $\theta = W_m/W$. The answer-to-request channel is an interference-free channel since the station is the only transmitter. That is, answer packets can be queued at the station and transmitted without conflicts. It is possible to give the answer-to-request channel enough bandwidth W_a such that answer packets do not incur any queueing delay at the station. Indeed, if b_r and b_a are the number of bits per request packet and answer-to-request packet respectively, then W_a should satisfy

$$W_a \geq W_r \frac{b_a}{b_r} \qquad (14)$$

where W_r is the bandwidth assigned to the request channel.

Let λ be the average number of messages generated per second. As usual, we shall assume the generation process to be Poisson. The maximum generation rate that the total bandwidth W can ever handle is W/b_m. The channel utilization denoted again by S is then expressed as

$$S = \lambda/(W/b_m) \qquad (15)$$

Since both control packets contain the same type of information, it is reasonable to assume that $b_a = b_r$ and therefore let $\eta = b_r/b_m$. We further let $W_r = W_a$. In this case we have

$$W_r = W_a = \frac{(1-\theta)W}{2} \qquad (16)$$

Consider the request channel operated in a random access mode. The expected delay incurred by a request packet is readily obtained from the simulation results presented earlier.

To estimate the delay $\mathfrak{D}_2$, we assume that the output of the random access request channel defined as the process corresponding to the arrival of successful requests at the station is Poisson with mean λ requests per seconds. We verified the above assumption by examining the distribution of inter-departure times (i.e., time between successive successful packets) of the Non-Persistent CSMA simulator and comparing it to the exponential density function. Except for interarrivals in the range of one or two packet transmission times the match is acceptable and the smaller S_r is, the more valid is the assumption. Under this assumption, the message channel can be modeled as an $M/G/1$ queueing system.[12]

The *maximum bandwidth utilization* is determined by the fact that the throughput on the request channel does not exceed its capacity (under the access mode in use) and the utilization of the message channel does not exceed one.

Numerical results

System Capacity

In Figure 15 we plot system capacity versus η (which represents a relative measure of the overhead due to control information) for the following access modes:

Pure ALOHA SRMA
Slotted ALOHA SRMA
Slotted Non-Persistent Carrier Sense SRMA ($\tau W/b_m = 0.01, 0.05$)

We note that the system capacity in SRMA reaches 1 for very small η. A case of interest considered throughout the paper corresponds to $b_m = 1000$ bits and b_r anywhere from 10 to 100 bits (b_r is directly related to the number of terminals in the population, since addressing information increases with increasing M). Thus, the interesting range for η is 0.01 to 0.1. For $\eta > 0.01$, the effect on the system capacity of the random access used to operate the request channel is important: a large improvement is gained when the request channel is operated in slotted Non-Persistent CSMA as compared to ALOHA. On the other hand, in comparing the capacity of SRMA to the capacity of random access modes, we note that SRMA can be superior only for relatively small values of η.

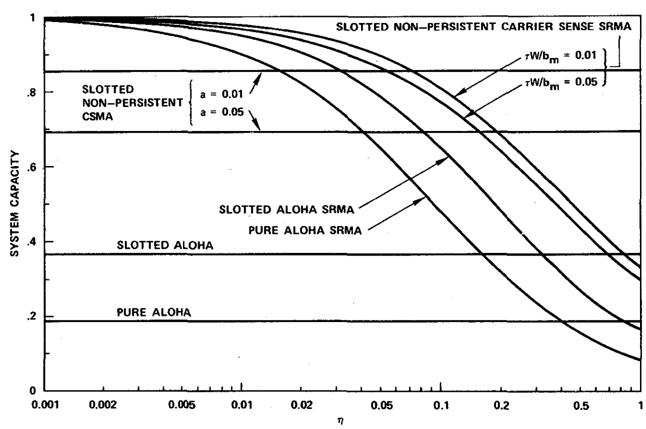

Figure 15—SRMA: Channel capacity versus η

CONCLUSION

Of interest to this paper was the consideration of packet-switched radio channels as a means of communication between terminals and a station (computer center, gate to a network, etc.). The objective of the research was to provide the communication system designer with various new access modes to the shared packet-switched radio channel, as well as the tools and conclusions necessary to select optimal solutions. Carrier Sense Multiple Access (CSMA) was introduced as a new method of multiplexing the terminals on the radio channel. Its performance was shown to be heavily affected by the ratio, a, of propagation delay to packet transmission time. In the cases of interest ($a \ll 1$), and under the major assumption that all terminals are in line-of-sight and within range, we have shown that CSMA provides improved capacity over the ALOHA modes.

However, the existence of hidden terminals can badly degrade the performance of CSMA. A good solution to the problem is provided by the Busy Tone Multiple Access (BTMA). BTMA under a Non-Persistent protocol is shown to achieve a channel capacity of 0.68 when the available bandwidth W is 100 KHz and up to 0.72 when $W=1$ MHz. Moreover, the channel capacity is shown to be insensitive to the precise setting of the system parameters.

A second alternative of multiplexing the terminals on the radio channel is the use of reservation techniques. The

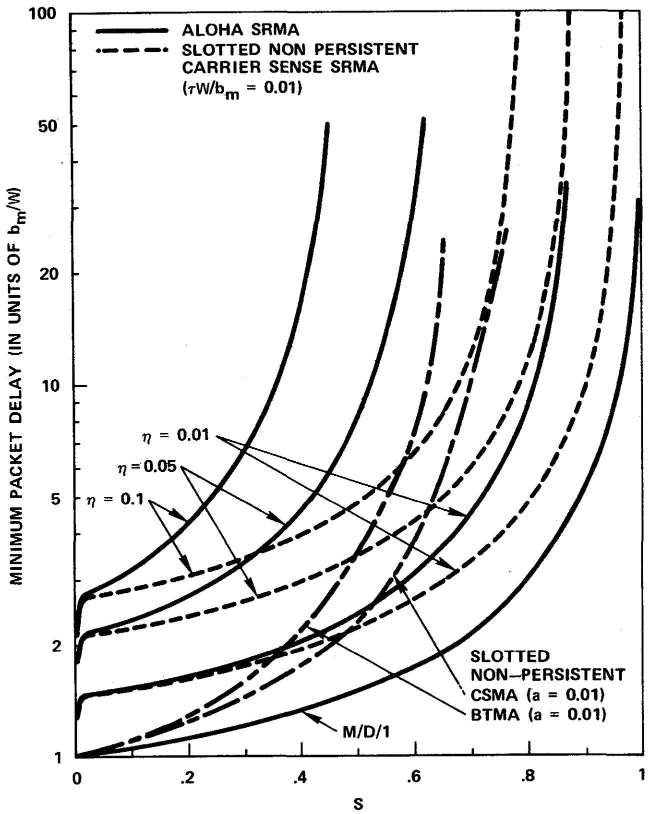

Figure 17—Minimum packet delay in SRMA

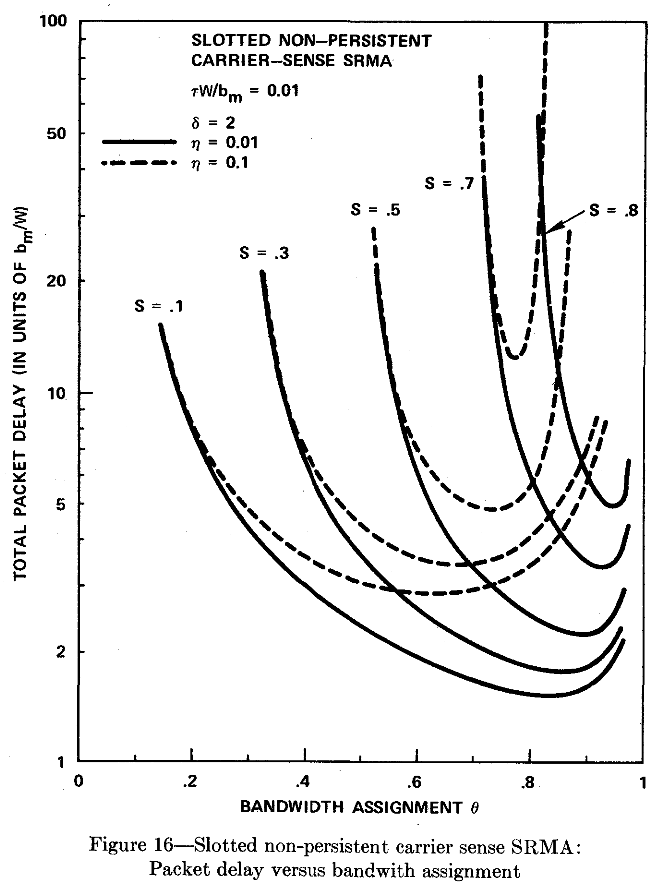

Figure 16—Slotted non-persistent carrier sense SRMA: Packet delay versus bandwith assignment

Delay Considerations

Let us restrict ourselves here to $\tau W/b_m=0.01$. For given η and S, the total message delay $\mathfrak{D}$ is a function of θ, the bandwidth assignment. As an example, we show slotted Non-Persistent Carrier Sense SRMA with fixed message length (packet) and $\tau W/b_m=0.01$ in Figure 16. Similar plots can be obtained for other random access modes used for the request channel. For each value of S, θ must lie in a feasible range denoted as $[\theta_{\min}, \theta_{\max}]$; $\theta_{\min}$ and $\theta_{\max}$ are determined by the saturation of the message channel and the request channel respectively. For small values of θ (θ close to $\theta_{\min}$), the major part of delay is due to $\mathfrak{D}_2$; for θ close to $\theta_{\max}$, it is due to $\mathfrak{D}_1$. The optimal bandwidth assignment is defined as the value of θ which minimizes total delay. We note that the higher the load is, the more critical is the choice of θ_{opt}. The minimum delay for ALOHA-SRMA and Slotted Non-Persistent Carrier Sense SRMA is shown in Figure 17 as a function of S for various values of η. First, in comparing the two systems between themselves, we note again an important improvement in using CSMA for the request channel. The improvement is more important when larger values of η are involved.

In comparing Carrier Sense SRMA with CSMA or BTMA, we note that, unless η is large (0.1 and above), there is a value of S below which CSMA or BTMA performs better than SRMA and above which the opposite is true.

Split-Channel Reservation Multiple Access (SRMA) was considered which employs random access techniques for the request channel. The capacity of the channel under SRMA is heavily affected by the level of overhead introduced. Moreover, the throughput delay performance is significantly dependent on the performance of the random access mode used on the request channel: a Non-Persistent Carrier Sense SRMA provides better performance than ALOHA-SRMA.

In all these comparisons we note that most of the channel capacity which was unavailable with pure and slotted ALOHA may be recovered by use of these more sophisticated access schemes.

REFERENCES

1. Kahn, R. E., "The Organization of Computer Resources into a Packet Radio Network," *NCC 1975 Proceedings.*
2. Abramson, N., R. Binder, F. Kuo, A. Okinaka, and D. Wax, "ALOHA Packet Broadcasting—A Retrospect," *NCC 1975 Proceedings.*
3. Garrett, J. and S. C. Fralick, "A Technology for Packet Radio," *NCC 1975 Proceedings.*
4. Frank, H., R. Van Slyke, and I. Gitman, "Packet Radio Network Design—System Considerations," *NCC 1975 Proceedings.*
5. Fralick, S. C., D. Brandin, F. Kuo, and C. Harrison, "Digital Portable Terminals," *NCC 1975 Proceedings.*
6. Burchfiel, J., R. Tomlinson and M. Beeler, "Functions and Structure of a Packet Radio Station," *NCC 1975 Proceedings.*
7. Abramson, N., "THE ALOHA SYSTEM—Another Alternative for Computer Communications," Fall Joint Computer Conference, *AFIPS Conference Proceedings,* 1970, Vol. 37, pp. 281-285.
8. Kleinrock, L. and S. S. Lam, "Packet-Switching in a Slotted Satellite Channel," National Computer Conference, New York, June 4-8, 1973, AFIPS Conference Proceedings, 1973, Vol. 42, pp. 703-710.
9. Abramson, N., "Packet Switching with Satellites," National Computer Conference, New York, June 4-8, 1973, *AFIPS Conference Proceedings,* 1973, Vol. 42, pp. 695-702.
10. Roberts, L. G., "Dynamic Allocation of Satellite Capacity Through Packet Reservation," National Computer Conference, New York, June 4-8, 1973, *AFIPS Conference Proceedings,* 1973, Vol. 42, pp. 711-716.
11. Spring Joint Computer Conference, *AFIPS Conference Proceedings,* 1970, Vol. 36, pp. 543-597; 1972, Vol. 40, pp. 243-298.
12. Kleinrock, L., *Queueing Systems, Vol. I, Theory, Vol. II, Computer Applications,* Wiley Interscience (1975).
13. Kleinrock, L. and F. A. Tobagi, "Packet Switching in Radio Channels: Carrier Sense Multiple Access Modes and their Throughput-Delay Characteristics," to appear in *IEEE Transactions on Communications.*
14. Tobagi, F. A. and L. Kleinrock, "Packet Switching in Radio Channels: The Hidden Terminal Problem and the Carrier Sense Multiple Access Mode with a Busy Tone," to appear in *IEEE Transactions on Communications.*
15. Tobagi, F. A., *Random Access Techniques for Data Transmission Over Packet Switched Radio Networks,* School of Engineering and Applied Science, University of California, Los Angeles, UCLA-ENGR 7499, December 1974.
16. Wainstein, L. A. and V. D. Zubakov, *Extraction of Signals from Noise,* Prentice Hall, Englewood Cliffs, New Jersey, 1962.
17. Chu, W. W., "A Study of Asynchronous Time Division Multiplexing for Time-Sharing Computer Systems," Spring Joint Computer Conference, *AFIPS Conference Proceedings,* 1969, Vol. 35, pp. 669-678.