PACKET SWITCHING IN RADIO CHANNELS: NEW CONFLICT-FREE MULTIPLE ACCESS SCHEMES FOR A SMALL NUMBER OF DATA USERS*

Leonard Kleinrock and Michel Scholl

Computer Science Department
University of California, Los Angeles
California 90024
(213) 825-2543

ABSTRACT

We study new access schemes for a population of geographically distributed data users who use packet-switching to communicate with each other and/or with a central station over a multiple-access broadcast ground radio channel. We introduce and analyze Alternating Priorities (AP), Round Robin (RR) and Random Order (RO) as new conflict-free methods for multiplexing a small number of buffered users without control from a central station. These methods are effective when the number of users is not too large; otherwise, a large overhead leads to performance degradation. To reduce this degradation, we consider a natural extension of AP, called Mini-Slotted Alternating Priorities (MSAP) which reduces the overhead and is superior to TDMA, FDMA, Polling and Random Access Schemes under heavy traffic conditions. At light input loads, only random access schemes outperform MSAP, when we have a large population of users. In addition, and of major importance, is the fact that MSAP does not require control from a central station.

INTRODUCTION

In this paper we focus our attention to the capacity allocation of a data communication channel to a set of data sources. The demands placed upon this resource are unpredictable and bursty [1] and are made by a population of geographically scattered and (possibly) mobile users.

One important use of a data communication channel is to connect terminal users to information processing and storage capacity available from a computer or a network of computers. The terminals are often connected to the computer by means of a centralized communication network, i.e., all demands made upon the channel are made by terminal-to-computer and by computer-to-terminal communications. In such networks, the computer controls the transmission from any user.

We also consider the communication channel as a media providing direct communication among the users (terminals, minicomputers) themselves. In such point-to-point communication networks, a central node can no longer efficiently and reliably control the transmission of all the network's users. On the contrary, control may be distributed among the users themselves. We are concerned with efficiently providing access from the users to the (centralized or point-to-point) communication network.

At the end of the sixties, the store and forward packet-switched technology emerged as a cost-effective alternative to the widespread circuit switching technology [2]. Originally, applied in distributed computer-communications networks, packet switching techniques have more recently been effectively used for radio communications (both satellite and ground radio channels) [3]. One of the first packet radio communication systems was the ALOHA system developed at the University of Hawaii [4]. The Advanced Research Projects Agency (ARPA) of the Department of Defense has undertaken a new effort whose goal is to develop a packet radio broadcast network as an interface between a point-to-point wire network (like the ARPANET) and a number of geographically scattered terminals [5]. Furthermore, there is currently an immense worldwide interest in the development of packet satellite communications systems [6, 7, 8].

In this paper, we restrict our attention to data communication over packet-switched ground[1] radio systems as an alternative for data transmission among users. For such data communications among users, broadcast radio communications are chosen as an effective alternative to conventional wire communications [3]. We consider a single broadcast high speed radio channel shared in some multi-access scheme and in a packet switched mode. The radio channel as considered in the following is characterized as a wide-band channel with a propagation delay between any source-destination pair which is only a very small fraction a of the packet transmission time[2].

The problem we are faced with is how to share and how to control access to the channel in a fashion which provides an acceptable level of performance. Several multiple access techniques which attempt to resolve some of these issues have previously been implemented or proposed. These fall into the following categories:

- Fixed Assignment: Time Division Multiple Access (FDMA) [9]
- Roll call Polling [9, 10]
- Random access schemes: ALOHA [6, 8, 11, 12] and Carrier Sense Multiple Access (CSMA) [13, 14]
- Reservation techniques: Carrier Sense Split-Reservation Multiple Access (CS SRMA) [15], Dynamic conflict-free reservation schemes [16]

With TDMA and FDMA, the performance is very sensitive to the number of users and we observe a poor delay performance at low loads due to the inherent burstiness of the traffic. On the other hand, ALOHA

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

[1] Here we exclude the study of satellite communications systems.

[2] Consider, for example, 1000 bit packets transmitted over a radio channel operating at a speed of 100k bits/s. If the maximum distance between any source-destination pair is 10 mi, then the (speed of light) packet propagation delay is of the order of 54μs. Therefore $a = .005$. If the users are less than 2 miles apart then $a = .001$. On the contrary, when one considers satellite channels [12], we usually have $a >> 1$.

and CSMA provide better delay performance at low input rates even for a very large number of users (e.g., N=1000 users) at the price of collisions increasing with the load; however, at higher load this results in a poor channel efficiency. In addition, random access modes are suitable for a distributed control of the access to the channel while Polling and Reservation schemes require special control from a master user.

We wonder if such random access techniques are optimal for the distributed channel access control of a small number ($N \leq 20$) of (possibly) buffered users. To answer this question, in this paper we recommend an approach different from the conflict prone random access modes. This approach is as follows:

1) We choose a distributed dynamic channel assignment known to each user which is conflict-free. By avoiding collisions, we insure a high channel utilization under heavy traffic conditions. By distributed assignment, we mean the following: If user i is presently transmitting a packet over the channel, an assignment scheme or priority rule common to all users designates a user j (possibly=i) to transmit next. At the end of the user's packet transmission, all users know from this priority rule to which user (j) the channel has next been assigned.

2) As in CSMA, we use the carrier sensing capability of each user to listen to the carrier due to another user's transmission[3]. After a propagation time τ, all other users may start detecting the presence or absence of the carrier due to user j's transmission. In case the carrier is absent (user j had no packet to transmit), they all know from the priority rule which user (k) is chosen to start transmission immediately. They all listen to the carrier for the next τ seconds, after which, if the carrier is absent, a third user may start transmitting a packet, etc... Therefore, even though only one user has a packet to transmit, after a worst case of N attempts, this user will be chosen again. Two classes of protocols are presented and analyzed in this paper. Alternating Priorities (AP), Round Robin (RR) and Random Order (RO) belong to the first class and are the subject of Section II in which we present the protocols, discuss the assumptions, establish the throughput-delay performance and finally compare them to the performance of existing multiple access modes. AP, RR and RO are shown to be suitable for the multiple access of a small number of (possibly) buffered users without the control of a central station. However, their performance degrades badly as the number of users increases. As an example of the second class of protocols, in Section III we introduce and analyze Minislotted Alternating Priorities (MSAP) and compare its performance to that of the other competitive access schemes. MSAP is shown to be very well adapted to the distributed multiple access to a ground radio channel with a small number of buffered users and to accept a larger number of users ($N \leq 50$) without serious performance degradation. Moreover, MSAP is shown to perform better than roll call Polling, and it is one of the few schemes known which, under heavy traffic conditions, behaves like M/D/1 (perfect scheduling) to within a multiplicative constant.

II. AP, RR, and RO protocols and their Throughput-Delay Characteristics

A) Transmission Protocols and System Assumptions

[3] In the context of packet radio channels, sensing carrier prior to transmission was originally suggested by D. Wax of the University of Hawaii in an internal memorandum dated March 4, 1971. The practical problems of feasibility and implementation of carrier sensing are not addressed here.

At any instant, a user is said to be ready if he has a packet ready for transmission (otherwise he is said to be idle). The various protocols considered below differ by the priority rule, common to all users. This rule, based on which user has transmitted the last packet, establishes a priority order among the N users for the next packet transmission. That user which is i^{th} in priority order ($1 \leq i \leq N$), (assuming he is ready) transmits the next packet if and only if all higher priority users are idle.

All users are assumed to be in line-of-sight (LOS) and within range of each other. Therefore, we assume that any user has the ability to sense the carrier of any other transmission on the channel. Furthermore, the time required to detect the carrier due to a packet transmission is considered to be negligible. All packets are of constant length and are transmitted

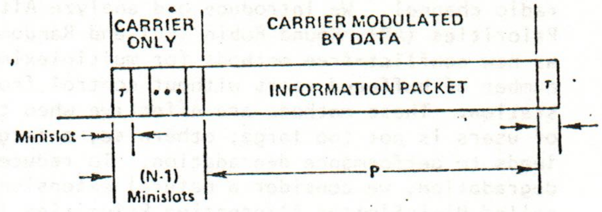

Figure 1. Slot Configuration.

over an assumed noiseless channel. The system assumes no multipath effects. The time axis is slotted. A slot consists of three parts (see Fig. 1):

1) an overhead of (N-1) "minislots", each of duration τ, where τ is the maximum propagation time between any source-destination pair and N is the number of users;

2) the packet transmission time of duration P;

3) one minislot which accounts for the (propagation) time between the end of transmission and the end of reception.

The N users are ordered in each slot by the priority rule which characterizes the protocol. For all priority rules (and thus for all protocols) the N users are synchronized[4] as follows in each slot:

1) If the highest priority user is ready, he need not sense the channel and synchronizes his packet's transmission as follows:

(i) At the beginning of the slot, he begins transmission of the carrier (with no modulation).[5]

(ii) (N-1) minislots later he transmits his packet. Otherwise, (if he is idle) he remains quiet until the end of the slot.

2) If the i^{th} user in priority ($1 < i \leq N$) is ready, he senses the channel for (i-1) minislots.

(i) If no carrier is detected after (i-1) minislots, then at the beginning of the i^{th} minislot, he

[4] The practical problems involved in synchronizing user are not addressed in this paper. In particular, we assume that when the users are not an equal distance apart, the synchronization is feasible, although not necessarily simple.

[5] After one minislot at most, all other users know whether the slot is reserved (carrier detected) or no (carrier absent) since the carrier detection time was assumed to be negligible. If this is too strong an assumption, one can increase the duration of a minislot (previously chosen as equal to τ), and use a larger value of the parameter a (previously defined a τ/P).

transmits his carrier and (N-i) minislots later, he transmits the packet.

(ii) Otherwise (idle user or carrier detected earlier) he waits for the next slot and the process is repeated (with a possibly different priority order).

Under all protocols, a slot is unused if and only if all users are idle.

We first consider Alternating Priorities (AP) named after the priority queueing system studied in [17]. AP obeys the following rule. The N users are ordered in a given sequence (say 1, 2, ..., N).

1) Assign the slot[6] to that user (say user i) who transmitted the last packet. If user i is ready, he transmits a packet in this slot. Otherwise (if there are no more packets from this user),

2) Assign the slot to the next user in sequence (i.e., user (i mod N)+1).

(i) If this next user is ready, he transmits a packet in this slot, and in the following slot, operate as above.

(ii) If this next user is idle, then repeat step 2 until either a ready user is found or the N users have been scanned. In this latter case (all users idle), the slot is unused and in the following slot, operate as above. (The following slot is assigned to user i).

The second protocol is called Round Robin (RR). As in TDMA, each user is pre-assigned one slot in a round robin (i.e., cyclic) fashion according to a given sequence of say, 1, 2, ..., N, 1, 2,...).

1) If the user (say i) to whom the current slot is assigned is ready, he transmits a packet in this slot.

2) Otherwise (user i idle) assign the slot to the next user in sequence (i.e., user (i mod N)+1).

(i) If this next user is ready, he transmits a packet in this slot.

(ii) If this next user is idle, then repeat step 2 until either a ready user is found or the N users have been scanned. In this latter case (all users idle), the slot is unused.

3) No matter who uses the current slot (assigned to user i), the next slot is assigned to user (i mod N)+1.[7]

In the third protocol, called Random Order, the priority order of the N users is chosen at random, i.e., each user generates the same pseudo-random permutation of 1, 2, ..., N which gives the priority order of the N users for the current slot. No matter who uses the current slot, each user generates a new permutation (the same for all users) which gives the priority order of the N users for the next slot.

B) Traffic Model and Channel Capacity

Here we characterize the traffic source, define some variables and give the first important performance measure, namely the channel capacity.

We assume that our traffic source consists of a finite number N of buffered users, with unlimited buffer size. Each user generates packets independently of the others according to a homogeneous Poisson point process. We assume that the full packet is instantaneously generated at those points. The aggregate packet generation rate is denoted by λ (packets/second). If N is not too large, each user may generate packets frequently enough so that the interarrival time between successive packets at a given user is less than the delay incurred by a packet from arrival to the end of transmission. Thus, each user may have more than one packet requiring transmission at any time, which will be transmitted on a first-come-first-served basis within his queue.

In addition, we characterize the traffic as follows. Each packet (of constant length) requires P seconds for transmission. Let $S = \lambda P$. S is the average number of packets generated per transmission time, i.e., it is the input rate normalized with respect to P. In equilibrium, S can also be referred to as the channel utilization [12, 14]. Indeed, if we were able to perfectly schedule the packets into the available channel space with absolutely no control overhead, we could achieve a maximum throughput equal to 1. Because there are N minislots wasted (for sensing the carrier) between the successive transmission of two packets (see Fig. 1), the maximum achievable throughput (the maximum channel utilization), called the channel capacity [12, 14] under a given protocol, and denoted by C, is less than one; the maximum rate of packets transmitted per slot is always equal to one, but the slot size increases with N. Since within each slot $N\tau$ seconds are lost for control, the channel capacity of AP, RR and RO is

$$C = \frac{1}{1 + Na} \tag{1}$$

As N increases, the capacity of AP, RR and RO decays very quickly below the slotted non-persistant CSMA capacity [14] when a is not too small (N=44 for a=.001, N=18 for a=.01, N=10 for a=.05) and is worse than the slotted ALOHA mode [12] for N > 172 if a=.01 (or N > 34 if a=.05). However, when a is very small, say a=.001, the capacity is large (> 90%) for N < 110.

As a increases, for a small number of users (N=10), the protocols studied in this section give a higher channel capacity than all CSMA protocols for values of a not larger than .038, and is fairly good when a < .02 (> 83%). When the number of users is larger, C quickly decays as a increases; if N=50 for values of a > .35, the capacity drops below that of slotted ALOHA.

Together with the channel capacity, the expected packet delay T is an important performance measure. T is defined as the average time, normalized with respect to P, elapsing from the generation of a packet until the end of its transmission.

C) Delay Analysis

We model our multiple access schemes as an M/D/1 priority queueing system with rest period [17] where the service time and the rest period have the same deterministic distribution of length one slot and where, to each user, there corresponds a priority with a total of N priority groups. In addition, our queueing disciplines are work-conserving (the server never stands idle in the face of a non-empty queue, i.e., a slot is unused if and only if all users are idle), non-preemptive and service independent. We may then apply Kleinrock's conservation law [18] which has been extended in [19] to include the case of priority queueing systems with rest period.

Let us consider first the case where all users have the same input rate: $\lambda_i = \lambda/N$ for all i. Denoting by D_i the expected packet delay (waiting time and transmission time) of a packet generated at user i normalized with respect to one slot, it is shown in [19] that:

[6] By "assigning the slot to user i" we mean: "the (highest) priority to transmit in this slot is given to user i".

[7] Variant: If the current slot is assigned to user i and used by user j (j = i, i+1, ..., N, 1, ..., i-1), then the next slot is assigned to user(j mod N)+1. If the current slot is not used (all users idle) then the next slot is assigned to user(i mod N)+1.

$$D_i = \frac{1}{2(1-\rho)} + 1 \qquad \text{for all } i \text{ and for all protocols} \tag{2}$$

where $\rho=\lambda[1+Na]P$ is the total normalized input rate (packet/slot). Thus we see that when the packet generation rate is the same at all users, then the expected packet delay is independent of the protocol.

When the input rate λ_i is not the same at all users, the problem of solving for D_i is not easy. An analysis of the expected packet delay under AP for the special case of N=2 can be found in [19].

Intuitively, one expects that packets generated at users with lower input rates should have smaller mean waits than should packets generated at users with higher input rates under RR as compared to AP. This is verified by simulation in [19]. From the conservation law [18, 19], we see that AP favors the packets generated at users with higher input rates.

Since the average packet delay is constant for these protocols in the case of identically loaded users $\left(\lambda_i = \frac{\lambda}{N} \text{ for all } i\right)$, it is necessary to compare the higher moments of the delay distribution under AP, RR and RO. The delay variance was simulated under AP, RR and RO for various values of the number of users. As N increases [19], the smallest variance is to be expected under AP and the largest under RR, the difference between the two being rather small (less than $1(\text{slot})^2$ for $\rho=.6$).

Thus, we may conclude that the three protocols are quite equivalent in terms of the mean throughput delay performance, and differ only slightly with respect to variance.

D) Discussion

Let us compare this performance to that of the best among the existing techniques of multiple access over a ground radio channel which have been mentioned in the introduction. In Figs. 2 and 3 we plot the average packet delay normalized with respect to the packet transmission time P versus the throughput for CSMA, POLLING, CS SRMA and for our new schemes (AP, RR and RO).

We plot the performance of CS SRMA as predicted by the infinite population model [15]. This performance (unstable channel: no steady state over an infinite time horizon) is likely an upper bound for the steady state performance (stable channel) for $N \leq 100$ users, whose performance has not been studied. The ratio of the request packet length over the information packet length in CS SRMA has been chosen to be $\eta = .01$.

An analysis of (roll call) Polling can be found in [10] where stationary distributions for queue lengths and waiting times are derived. These results are applied in [15] to packet radio. The expected packet delay (normalized with respect to P) is given by [15]:

$$T = 1 + \frac{S}{2(1-S)} + \frac{a}{2}\left(1 - \frac{S}{N}\right)\left(1 + \frac{Nr}{1-S}\right) \tag{3}$$

where r is chosen to be r=3.

Fig. 2 depicts an example of performance for N=10 users. The performance of TDMA is plotted on the same figure. The expected delay under TDMA [14, footnote #2], [20][8] is given by

$$T = 1 + N\left[\frac{S}{2(1-S)} + \frac{1}{2}\right] \tag{4}$$

For these parameters, we find the channel capacity is equal to .91 (see Eq. [illegible]) for AP, RR and RO, and the delay under those protocols is by far lower than with TDMA and slightly larger than with Polling. Let us now compare AP, RR and RO to CSMA. The (slotted non-persistent) CSMA performance is plotted as predicted by the infinite population model [14]. It was shown in [20] that the performance predicted by this model is a very good approximation of the performance of N=10 buffered users contending for the channel under CSMA. We note from Fig. 2 that at light traffic, CSMA provides the best delays. But when S is greater than .5, the new schemes perform much better than CSMA. CS SRMA performance has not been plotted in Fig. 2, since the model that predicted this performance is not suitable for a small number of users. If a=.001, even for a more

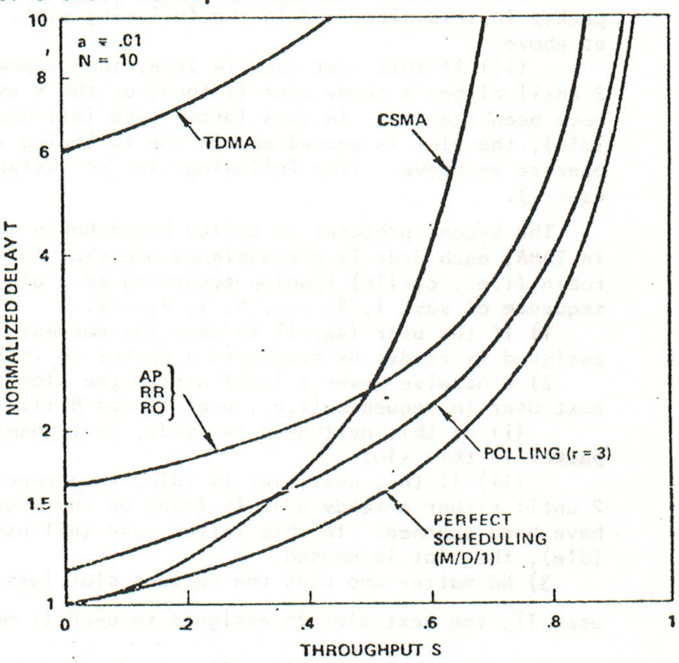

Figure 2. AP, RR and RO, Average Packet Delay vs Throughput: Comparison to Existing Techniques (N = 10; a = .01).

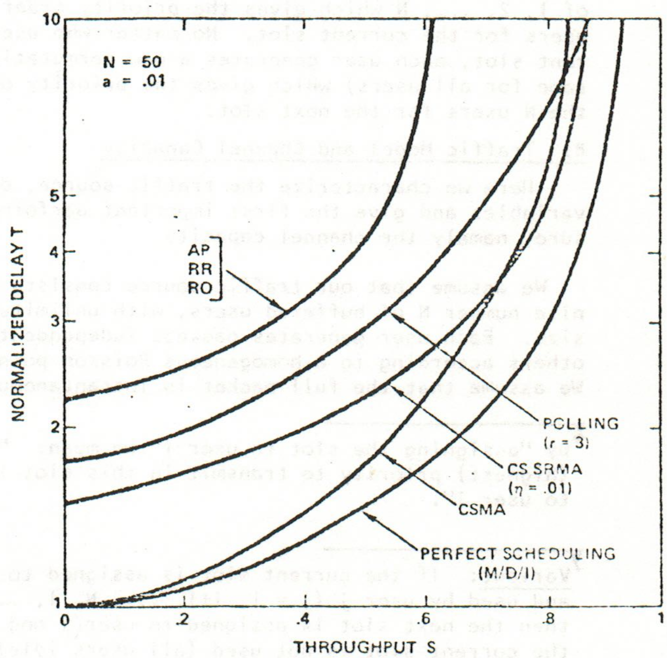

Figure 3. AP, RR, RO: T vs S. Comparison to Existing Techniques (N = 50, a = .01)

[8] As a matter of fact, the expected packet delay under TDMA, as given in [20, Eq. (1)] is incorrect, and should be reduced by (N-1).

significant number of users (N=50), AP, RR and RO produce a performance comparable to that under Polling or CS SRMA [19]. Under heavy traffic conditions ($S > .6$), the new schemes provide lower delays than CSMA, although they achieve the same channel capacity: $C \simeq .95$ [19].

In the second example we choose N=50 and a=.01. The steady state performance (stable channel) of CSMA is plotted in Fig. 3, as predicted by the finite population model studied in [21]. The delay is significantly higher with the new protocols. Even at very light traffic ($S \simeq 0$), the delay is 2.25. The capacity of the channel is only 2/3 under the new protocols (Eq. (1)), while it is .84 for CSMA [21], greater than .9 for CS SRMA [15], and 1 for Polling.

In summary, for a small number of users, AP, RR and RO provide a good channel capacity ($C=1/(1+Na) > .9$) and a delay-throughput performance close to that of Polling, or CS SRMA (Fig. 2). When all users are very close to each other (a small, e.g., a=.001), AP, RR and RO accept a significant number of users ($N \leq 50$) without performance degradation. Under heavy traffic conditions they perform better than CSMA. But, as with CSMA and CS SRMA, they require all users to be in line-of-sight and within range of each other, while Polling does not have such a requirement. However, the new schemes (and also CSMA) have the advantage of not requiring control from a master user (central station) while Polling and CS SRMA do. Thus, we conclude that the new protocols are particularly suitable for multiple access from a small number of buffered users without control from a central station. When a is not too small (e.g., a=.01) the performance degrades with the number of users (Fig. 3). Indeed, in each slot the overhead is proportional to the number of users.

In the next section, we consider a natural extension of AP, which reduces the control overhead.

III. MINI-SLOTTED ALTERNATING PRIORITIES (MSAP)

In this section we introduce and analyze a conflict-free scheme referred to as Mini-Slotted Alternating Priorities (MSAP), which also allows buffering capabilities and also does not require control from a central station. We solve for the average packet delay under MSAP and show that MSAP performs better than CSMA (and CS SRMA) under heavy traffic conditions for all numbers of users, and performs better than Polling for all traffic levels and all numbers of users.

A) Protocol

The major difference between MSAP and the schemes AP, RR and RO studied in Section II, comes from the slot size which is now taken as equal to the maximum propagation delay τ, (what we called a minislot in Section II is now referred to as a slot). As with the former schemes, we use the carrier sense capability of each user. However, we now reduce the channel time lost in control, i.e., the overhead due to carrier sensing in order to "steal" a slot assigned to an idle user. Without loss of generality, we assume that users grasp the channel according to a fixed order, say 1, 2, ..., N. The protocol obeys the Alternating Priorities rule as follows:

1) Assign the channel to that user (say i) who transmitted the last packet, if possible; otherwise,
2) Assign the channel to the next user in sequence.

By carrier sensing, at most one (mini-)slot later, all users detect the end of transmission of user i (absence of carrier [9]); in particular, so does the next user in sequence (user(i mod N)+1. Then

i) either: User (i mod N)+1 starts transmitting a packet; in this case, one slot after the beginning of his transmission, all others detect the carrier. They wait until the end of this packet's transmission and then operate as above.

ii) or: User (i mod N)+1 is idle; in this case, one slot later, all other users do not detect the carrier; they know that it is the turn of the next user in sequence, i.e., user (i mod N)+2 and operate as above.

When all users are idle, the "turn" keeps changing at each slot until it is the turn of a non-idle user.

In Fig. 4 we consider an example with four users. Two slots after the end of user 3's transmission, user 1 starts transmission since he detected that user 4 was idle. He transmits three packets, followed by user 2 and then user 3.

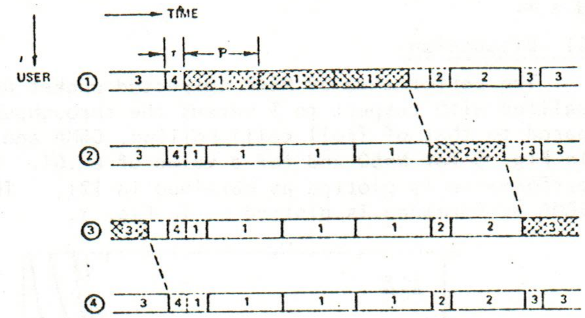

Figure 4. Minislotted Alternating Priorities (MSAP): Example of 4 Users (Cross-Hatching Indicates a Transmission).

Two remarks are noteworthy:- a) We could have chosen the Random Order Rule or the Round Robin rule. The latter is suitable for unbalanced traffic; i.e., when some users have a smaller input rate than others, then the Round Robin rule provides to users with small input rates a more frequent access to the channel than the Alternating Priorities rule does. However, the Alternation Priorities rule is chosen here, in order to minimize the "changeover" time between users. This changeover time, which is lost for packet transmission, is shorter with Alternating Priorities than it is with Round Robin (or Random Order). This overhead (one slot per switchover from one user to another) is very small, compared to that incurred with the schemes studied in Section II (for which N minislots are lost at each packet transmission time). The maximum channel utilization is obtained with the Alternating Priorities rule which allows the system to achieve full utilization of the channel. When one queue is saturated and keeps the channel for its own use, there is no changeover and therefore the throughput is S=1 packet/packet transmission time. Thus, the capacity of MSAP is equal to 1. b) In roll call Polling, the channel is assigned to the users according to the same rule. The only difference is that the polling time or changeover time between the two users is equal to the polling message transmission time, of length $b(\geq 1)$ slots, plus twice the propagation time between users and station [15].

If we denote this changeover time by r, we then have for

$$\text{a) Polling} \quad r = b + 2 \qquad \text{b) MSAP} \quad r = 1 \tag{5}$$

Since the polling message contains the identification of the user which is polled, b will increase with N (in particular, it must grow in proportion to $\mathrm{Log}_2 N$). [10]

---

[9] The carrier detection time is assumed to be negligible (see Section II).

[10] An alternative method of Polling, called "Hub Go-ahead" Polling, used is wire communications, is not readily applicable to our radio system, but is advantageous on long lines communications: the main advantage lies in lessening the number of line turnarounds [9].

particular, it must grow in proportion to $Log_2 N$).[10] Also, b depends on the parameter a. If a increases, b will decrease down to a minimum of 1. From the last statement and Eq. (5) it is evident that the changeover time is much smaller with MSAP than with Polling.

B) Expected Delay

We may apply the results of Konheim and Meister [10] for (roll call) Polling to MSAP by choosing the "polling" time r equal to 1 in Eq. (3). This equation gives the expected normalized delay in ground radio Polling [15]. Then, with MSAP, the expected normalized packet delay is given by

$$T = 1 + \frac{S}{2(1-S)} + \frac{a}{2}\left(1 - \frac{S}{N}\right)\left(1 + \frac{N}{1-S}\right) \qquad (6)$$

In particular, at very light traffic ($S \approx 0$), the ratio of the expected packet delay under Polling, over the expected packet delay under MSAP increases as $Log_2 N$ as $N \to \infty$.

C) Discussion

The performance of MSAP (expected packet delay normalized with respect to T versus the throughput) is compared to that of (roll call) Polling, CSMA and CS SRMA in Fig. 5 for N=50 and for a value of a=.01. The CSMA performance is plotted as obtained in [21]. The CS SRMA performance is plotted as in Fig. 3.

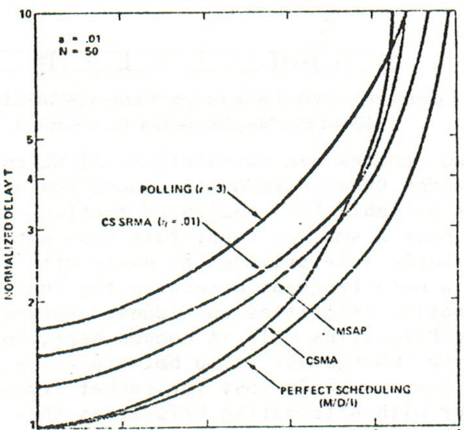

Figure 5. MSAP, T vs S: Comparison to Existing Techniques (N = 50)

In comparing MSAP to CSMA and CS SRMA, we note from Fig. 5 that at light traffic, the larger N is, the better is CSMA (and CS SRMA) performance as compared to the performance of MSAP. In particular, at very light traffic ($S \approx 0$) the expected packet delay under MSAP (see Eq. (6)) is $\left[1 + \frac{a}{2}(1+N)\right]$, whereas for CSMA it is 1.

However, under heavy traffic conditions, MSAP always performs better than CSMA (and CS SRMA). For N=50 the delay with MSAP is lower than that obtained with CSMA, for an input rate S_0 equal to .7 or higher.

In summary, the delay-throughput performance of MSAP exceeds that of Polling for all values of N and a. Under heavy traffic conditions, MSAP provides the best performance of all existing techniques for all values of N and a. (However, the reservation scheme introduced in [16] may perform better than MSAP (depending on the value of a). Under light traffic conditions, random access techniques outperform MSAP (by far if N is large). However, MSAP is more suitable than CSMA for a small number of (possibly) buffered users, since then even when the traffic is very light, the expected delay under MSAP is only slightly larger than that observed with CSMA (at S=0, the difference is equal to Na/2).

So far, in our comparison between MSAP and random access techniques (CSMA and CS SRMA), we have considered a small value of the parameter a: $a \leq .01$. As a increases, the performance of CSMA - and therefore that of CS SRMA - is known to decay below that of slotted ALOHA [14]. How does MSAP perform, compared to CSMA and slotted ALOHA when a increases? In addition, when a is small ($a \leq 0.1$) MSAP performs better than TDMA even for a small number of users (see Eq. (4) and (6)). This is no longer true when a is large ($a > .5$) and N small.

To examine these issues, for a given value of the input rate S, let us consider the regions of the Nxa plane in which either TDMA (Eq.(4)) or slotted ALOHA [12] or (slotted non-persistent) CSMA or MSAP (Eq.(5)) provides the lowest expected delay. (For all values of N, the performance of CSMA used in our comparison has been chosen as that predicted by the infinite population model studied in [14]. As mentioned earlier, this is only a lower bound on the performance predicted by more realistic models.) When the traffic is light (Fig. 6: S=.3), four regions are delimited in which respectively CSMA, ALOHA (N and a large), TDMA (N very small and a large) and MSAP produce the lowest delay. As might have been expected for a large population of users, random access techniques perform better, and when a increases, CSMA's performance decays below that of ALOHA. It is clear also from Fig. 6 that for most values of a(<.6), MSAP performs the best when N is not too large.

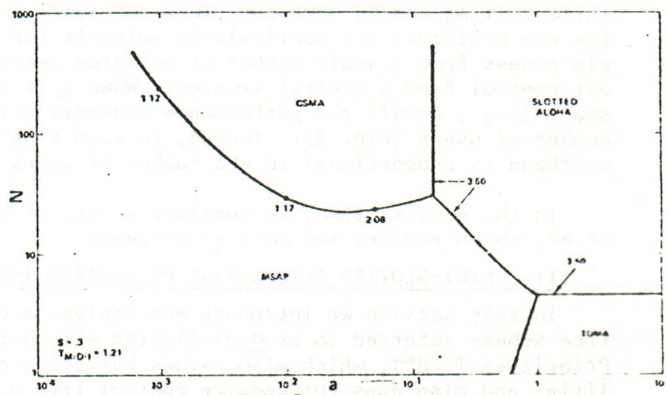

Figure 6. N vs a (S = .3)

Under heavier traffic conditions (Fig. 7: S=.6), the ALOHA region disappears since the maximum achievable throughput under slotted ALOHA is S=1/e. When a is large, TDMA is the best scheme. It is interesting to note that the bound on a beyond which TDMA performs better is a=1 for most values of N(N > 10) and that MSAP performs better than CSMA with a significantly larger number of users than under light traffic conditions. In particular, observe that for a > .1, MSAP is always better than CSMA since one cannot achieve a throughput of S=.6 with CSMA if a > .1.

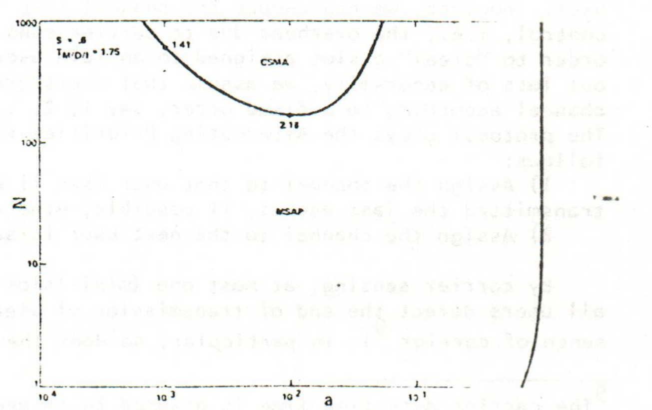

Figure 7. N vs a (S = .6)

On the boundary line between two regions, two schemes produce the same delay performance. The ratio of this delay performance over the (M/D/1) perfect scheduling's performance is represented at various points by the numerical labels in the figures along the boundary. As an example, consider Fig. 7 (S=.6). For a < 1, two regions are delimited. Above the contour (N large, a < .1), CSMA provides the lowest delay. Below the contour, MSAP performs better. For N ≈ 170 and a=.01, both schemes produce the same expected delay which is 2.18 times the expected delay in an M/D/1 queue.

IV. SUMMARY

When we are in the presence of a small number of users requiring buffer space for more than one packet, random access techniques (e.g., CSMA) are not necessarily optimal for the multiple access to a single ground radio channel. AP, RR and RO were introduced as new conflict-free multiplexing techniques suitable for a small number of (possibly) buffered users (Section II). These new schemes, as random access schemes, do not require control from a central station. When the number of users is small (e.g., N < 20 is a=.01) the new protocols produce a delay-throughput performance comparable to that of roll-call Polling and better than CSMA under heavy traffic conditions (Fig. 2). However, the major assumption made was that all users are in LOS and the performance of AP, RR and RO degrades badly as N increases (Fig. 3). To reduce this degradation, we considered a natural extension of AP, called MSAP (Section III), which is also conflict-free, reduces the overhead, provides a very good performance under heavy traffic conditions and is superior to AP, RR and RO. MSAP achieves a capacity of 1 and was shown to perform better than Polling (Fig. 5). MSAP is less sensitive than CSMA to an increase of the ratio of the propagation time to the packet transmission time (Figs. 6 and 7); however, a degradation of the delay was observed at light traffic as N increases. As with CSMA, MSAP assumes that all users are in LOS. However, MSAP presents the advantage of not requiring the control from a central station (so, too, with CSMA) while Polling does.

In conclusion, MSAP is one of the few schemes known which, under heavy traffic conditions, behaves like M/D/1 (perfect scheduling) to within a multiplicative constant (Eq.(6)). MSAP is very well adapted to the distributed access to a ground radio channel with a small number of buffered users and accepts a larger number of users (e.g., N=50 if a=.01) without serious degradation.

ACKNOWLEDGEMENT

The authors take great pleasure in acknowledging the assistance of Ch. W. Tseng who created Figs. 6 and 7.

REFERENCES

[1] Jackson, P.E. and C.D. Stubbs., "A Study of Multi-access Computer Communications," *Spring Joint Comp. Conf. Proc.*, Vol. 34, pp. 491-504, 1969.

[2] Roberts, L.G., "Data by the Packet," *IEEE Spectrum*, Vol. 11, No. 2, pp. 46-51, February 1974.

[3] Kleinrock, L., *Queueing Systems, Vol. II, Computer Applications*, Wiley-Interscience, New York, 1976.

[4] Abramson, N. "The Aloha System - Another Alternative for Computer Communications," Fall Joint Comp. Conf., *AFIPS Conf. Proc.*, Vol. 37, pp. 281-285, 1970.

[5] Kahn, R.E. "The Organization of Computer Resources into a Packet Radio Network," Nat'l. Comp. Conf., *AFIPS Conf. Proc.*, Vol. 44, pp. 177-186, 1975.

[6] Abramson, N. "Packet Switching with Satellites," Nat'l. Comp. Conf., New York, June 4-8, 1973, *AFIPS Conf. Proc.*, Vol. 42, pp. 695-702, 1973.

[7] Roberts, L.G. "Dynamic Allocation of Satellite Capacity Through Packet Reservation," Nat'l. Comp. Conf., New York, June 4-8, 1973, *AFIPS Conf. Proc.*, Vol. 42, pp. 711-716, 1973.

[8] Lam, S.S. "Packet Switching in a Multi-Access Broadcast Channel with Application to Satellite Communication in a Computer Network," Sch. of Engr. and Appl. Sci., Univ. of California, Los Angeles, CA., Rep. No. UCLA-ENG-7429, April 1974.

[9] Martin, J. *Teleprocessing Network Organization*, Prentice-Hall, Englewood Cliffs, New Jersey, 1970.

[10] Konheim, A.G. and B. Meister. "Waiting Lines and Times in a System with Polling," *J. Of the ACM*, Vol. 21, No. 3, pp. 470-490, July 1974.

[11] Roberts, L.G. "ALOHA Packet System With and Without Slots and Capture," ARPA Network Info. Center, Stanford Research Institute, Menlo Park, CA., ASS Note 8 (NIC 11290), June 1972.

[12] Kleinrock, L. and S.S. Lam. "Packet Switching in a Multi-access Broadcast Channel: Performance Evaluation," *IEEE Trans., on Comm.*, Vol. COM.23, pp. 410-423, April 1975.

[13] Tobagi, F.A. "Random Access Techniques for Data Transmission over Packet Switched Radio Networks," Sch. of Engr. and Appl. Sci., Univ. of California, Los Angeles, CA., Rep. No. UCLA-ENG-7499, December 1974.

[14] Kleinrock, L. and F.A. Tobagi. "Packet Switching in Radio Channels: Part I - Carrier Sense Multiple Access Modes and their Throughput - Delay Characteristics," *IEEE Trans. on Comm.*, Vol. COM.23, No. 12, pp. 1400-1416, December 1975.

[15] Tobagi, F.A. and L. Kleinrock. "Packet Switching in Radio Channels: Part III - Polling and (Dynamic) Split-Channel Reservation Multiple Access," *IEEE Trans. of Comm.*, Vol. COM-24, No. 8, pp. 832-845, August 1976.

[16] Kleinrock, L. "Performance of Distributed Multi-Access Computer-Communication Systems," accepted for publication, IFIP Congress '77, Canada.

[17] Miller, L.W. "Alternating Priorities in Multi-class Queues, Ph.D. Dissertation, Cornell Univ., Ithaca, New York, 1964.

[18] Kleinrock, L. "A Conservation Law for a Wide Class of Queueing Disciplines," *Naval Res. Logis. Quar.*, Vol. 12, pp. 181-192, 1965.

[19] Scholl, M. "Multiplexing Techniques for Data Transmission over Packet Switched Radio Systems," Ph.D. Dissertation, Univ. of California, Comp. Sci. Dept. Los Angeles, California, September 1976.

[20] Tobagi, F.A. and L. Kleinrock. "On the Analysis and Simulation of Buffered Packet Radio Systems," Ninth Hawaii Internat'l Conf. on Sys. Sci., Univ. of Hawaii, Honolulu, January 1976, *Proc. of the Special Subconference on Computer Nets*, 1976.

[21] Tobagi, F.A. and L. Kleinrock. "Packet Switching in Radio Channels: Part IV - Stability Considerations and Dynamic Control in Carrier Sense Multiple Access," to appear in *IEEE Trans. on Comm.*