AN ANALYSIS OF THE RESERVATION-ALOHA PROTOCOL FOR SATELLITE PACKET SWITCHING

Simon S. Lam

Department of Computer Sciences
The University of Texas at Austin
Austin, Texas 78712

ABSTRACT

A brief overview of multiaccess protocols for packet broadcasting networks is first given. The Reservation-ALOHA protocol of Crowther et al. [1] for satellite packet switching is then analyzed. It is assumed that the slotted ALOHA component of a Reservation-ALOHA channel is well-behaved and has a known equilibrium probability distribution for delay. Two user models with Poisson message arrivals are considered. (Each message consists of a group of packets with a general probability distribution for group size.) In the first model, each user handles one message at a time. In the second model, each user has infinite buffering capacity for queueing. Equilibrium probability distributions are derived for important system performance variables including message delay. Bounds on channel capacity are established for two slightly different protocols and some numerical results are illustrated.

1. INTRODUCTION

Packet broadcasting networks may be defined to be packet switching networks in which the connectivity requirements of a population of users are furnished by a broadcast medium (see figure 1). Two obvious examples of broadcast media are satellite and ground radio channels [2,3]. However, they may also be cable networks with either a loop or multipoint tree topology [4,5].

The basic operation of a packet broadcasting network can be explained as follows. The broadcast medium is analogous to a mailbox into which a sender may deposit a package of data (a packet). The destination address on the package is (somehow) visible to all users. The package will be retrieved by the proper addressee (or addressees) and ignored by the others. The mailbox, however, is big enough for one package only. Consequently, if two or more packages are placed into it at the same time (by different users, of course!), they are all destroyed.

From the above analogy and figure 1, we see that the central problem of a packet broadcasting network is <u>conflict resolution</u> among the population of users sharing use of the broadcast channel. The users are typically geographically distributed. The distances involved range from thousands of miles for a satellite network to, perhaps, tens of feet for an in-house cable network. Many multiaccess protocols for conflict resolution have been proposed and studied. We investigate in this paper the Reservation-ALOHA or R-ALOHA protocol originally proposed by Crowther et al. [1] for a satellite network. An analytical model is developed and equilibirum probability distributions of important system performance variables including message delay are derived.

Before we go into details of the analytic results for R-ALOHA we shall briefly review the class of multiaccess protocols for packet broadcasting networks.

2. MULTIACCESS PROTOCOLS

We classify multiaccess protocols for packet broadcasting networks into two major categories (see figure 2).

1) Protocols for passive users.

   Included in this category are a variety of polling protocols. The users normally keep quiet whether or not they have data to send. They are queried from time to time by a central controller. A user can use the broadcast channel only when so queried. Polling protocols have been widely implemented in many of the existing data networks. A drawback to conventional polling protocols is excessive overhead in networks with a large population of users with bursty traffic. The message delay is more a function of the time required to poll the population of <u>all</u> users than of the amount of traffic from competing users [6]. Recently, an adaptive polling protocol proposed by Hayes [7] attempts to alleviate this problem.

2) Protocols for active users.

   When an active user has data to send, he does something about it; he can either (a) sends it over the broadcast channel judiciously, trying his best to avoid conflicts, or (b) attempts to join a queue global to all users for using the broadcast channel. These two alternative actions give rise to the two subcategories of (a) contention, and (b) reservation protocols for active users in figure 2.

The pure ALOHA protocol [8] was the first of the "random access" protocols within the class of contention protocols. With contention protocols, there is no attempt to coordinate the users. Instead, each user monitors the broadcast channel and tries to transmit his data the best he can without incurring a conflict. With the pure ALOHA and slotted ALOHA protocols, users monitor the broadcast channel for the success or failure of their own transmissions*, which gives rise to a maximum channel throughput of 1/(2e) and 1/e respectively, under the assumption of a large population of very bursty users [9]. For broadcast networks with a very short propagation delay (relative to a packet transmission time), CSMA protocols [10] were devised to significantly improve the channel throughput by requiring the users to do carrier sensing (i.e. to monitor the channel to see if it is already occupied by a packet transmission) before attempting a transmission. In the special case of a loop network, a built-in priority ordering exists among the users and conflicts can be avoided completely with carrier sensing; the maximum channel throughput is thus unity [4].

For networks with a short propagation delay, other conflict-free protocols have also been proposed and studied [11]. The relatively large propagation delay (= 0.27 second) of a satellite channel renders carrier sensing impractical except in the case of an extremely long packet or low speed channel. The R-ALOHA protocol (to be described below) was proposed [1] to improve the throughput of a satellite channel beyond that of slotted ALOHA. We show below that the improvement can be substantial when users have several packets of data to send at a time or when they have steady streams of packet arrivals and buffering capability. Although it was invented for a satellite channel, R-ALOHA can be used for any of the other broadcast media.

Reservation protocols, also for active users, maintain a queue global to all users for channel access. Each user when he has data to send generates a request packet to reserve for a place in the queue. A fraction of the broadcast channel capacity is used

* Alternatively, a positive acknowledgement protocol is used instead [8].

for such request packets. Implementation of the queue may be centralized or distributed. With a global queue, the transmission of data packets is orderly. However, a multiaccess protocol is now needed for the request packets. In most proposals, the queue implementation is distributed; either a contention of fixed TDMA protocol is used for conflict resolution of request packets [9].

3. THE R-ALOHA PROTOCOL

The broadcast channel is assumed to be slotted in time, and the slots are organized into frames with M slots in each frame (see figure 3). Each time slot is long enough for the transmission of a packet of data. The duration T of a frame is assumed to be greater than the maximum channel propagation delay in the broadcast network. Consequently, each user is aware of the usage status of time slots one frame ago. The network operates without any central control but requires each user to obey the same set of rules for transmitting packets into time slots depending upon what happened in the previous frame. A time slot in the previous frame may be:

- unused, which means that either (a) it was empty, or (b) two or more packets were transmitted into it (a collision) and thus none could be received correctly;
- used, which means that exactly one packet was transmitted into it and the packet was successfully received. (We shall assume that the channel is error-free except for collisions.)

The transmission rules are:

1) If slot m (say) had a successful transmission by user X (say) in the previous frame, slot m is off limits to everyone except user X in the current frame. Slot m is said to be reserved by user X. (Note that user X has exclusive access to slot m as long as he continues to transmit a packet into it in every frame.)
2) Those slots in the last frame which were unused are available for contention by all users according to a slotted ALOHA protocol (the details of which are not specified).

We further differentiate between two protocols depending on whether an end-of-use flag is included in the last packet before a user gives up his reserved slot:

(P1) End-of-use flag not included, and
(P2) End-of-use flag included.

4. USER MODELS

A population of N users is considered with identical behavior and message arrival statistics. Messages arrive to each user according to a stationary Poisson process with rate λ messages/second. Each message consists of a group of h packets, with the first two moments $\overline{h}$ and $\overline{h^2}$ and probability generating function H(z).

We shall require that each user can reserve at most one time slot in a frame at a time. (The more general model which permits a user to reserve multiple time slots in a frame may be approximated by regarding each such user as multiple users in our model here.) With this requirement, the problem is interesting only if N >M, where N may possibly be infinite as in the infinite population model of slotted ALOHA. Another consequence is that the number of users who may access a nonreserved slot is N - m, where m is the number of users holding reserved slots.

We shall consider the following user models:

1) Single-message users
   Each user handles one message at a time. (The Poisson source shuts itself off until all packets of the current message have been successfully transmitted.)
2) Queued users
   Each user has infinite buffering capacity; a queue is maintained with Poisson arrivals at the constant rate of λ messages/second.

We note that both user models are much more general than user models previously considered for "random access" protocols [10,12].

We define the random variable v to be the total number of packets that a user transmits before he gives up a reserved time slot. For the model of single-message users, v is just the number h of packets in a message. For the model of queued users, v is the number of packets that arrive within a busy period of the user queue. We shall use $\overline{v}$ to denote the mean value of v.

5. THE ANALYSIS

In order to characterize the performance of the R-ALOHA protocol, we are interested in the probability distributions of the number of used slots within a frame and message delay. Our derivation of these results is based upon the following assumption concerning the usage of nonreserved time slots under a slotted ALOHA protocol.

Assumption (A1) A successful packet transmission occurs in each nonreserved time slot with a constant probability (denoted by S).

It is well known that a slotted ALOHA channel suffers from instability behavior and needs to be adaptively controlled [12,13]. Adaptive control algorithms have been proposed and studied extensively in the past [12-16]. Our approach here is to divorce ourselves from this problem. Instead, we shall consider it to be solved satisfactorily and separately such that the above assumption is valid. (This idea was suggested by Ferguson in [15].) This is why details of the slotted ALOHA protocol were omitted in our earlier specification of the R-ALOHA protocol.

We next define the equilibrium probabilities

$$P_i = \text{Prob [i slots in a frame are used]}.$$

Our first major result follows.

Theorem 1

$$P_i = \binom{M}{i} U^i (1-U)^{M-i} \tag{1}$$

where

$$U = \frac{S}{S + (1/\overline{v})} \tag{2}$$

under protocol (P1), and

$$U = \frac{S}{S + [(1-S)/\overline{v}]} \tag{3}$$

under protocol (P2).

Proof: The key of the proof is to consider each of the M channels in figure 3 separately. Given assumption (A1) and the assumption that users have independent identical arrival statistics, the channels are statistically independent of each other. Also, each channel has alternating idle and busy periods which are statistically independent and constitute an alternating renewal process. Let t_{idle} and t_{busy} be the idle and busy period duration respectively. The probability that the channel is busy is from Cox [17]

$$U = \frac{E[t_{busy}]}{E[t_{idle}] + E[t_{busy}]}$$

Since the M channels are statistically independent, Eq.(1) follows.

Given assumption (A1), we have

$$\text{Prob}[t_{idle} = k \text{ slots}] = S(1-S)^{k-1} \quad k=1,2,...$$

under (P1), and

$$\text{Prob}[t_{idle} = k \text{ slots}] = S(1-S)^{k} \quad k=0,1,...$$

under (P2). Hence,

$$E[t_{idle}] = \begin{cases} 1/S & \text{under (P1)} \\ (1/S)-1 & \text{under (P2)} \end{cases}$$

Under both (P1) and (P2),

$$E[t_{busy}] = \bar{v}$$

Hence,

$$U = \frac{\bar{v}}{\frac{1}{S} + \bar{v}} = \frac{S}{S+(1/\bar{v})}$$

under (P1), and

$$U = \frac{\bar{v}}{[(1/S)-1] + \bar{v}} = \frac{S}{S+[(1-S)/(\bar{v})]}$$

under (P2). Q.E.D.

From Eq.(1), the probability generating function of P_i is

$$Q(z) = (1-U+Uz)^M \tag{4}$$

with mean

$$\bar{m} = MU \tag{5}$$

variance

$$\sigma_m^2 = MU(1-U)$$

and coefficient of variation

$$C_m = \sigma_m/\bar{m} = \sqrt{(1-U)/(MU)}$$

Note that in the above results, v can have a general probability distribution. However, if we restrict v to be geometrically distributed, i.e.

$$\text{Prob}[v=i] = r(1-r)^{i-1} \qquad i=1,2,...$$

for some parameter r, the following equations can be derived for Q(z) using a Markov chain approach.

$$Q(z) = (1-S+Sz)^M \, Q(\frac{z+r(1-z)}{1-S+Sz}) \tag{6}$$

under (P1), and

$$Q(z) = (1-S+Sz)^M \, Q(\frac{(1-r)z}{1-S+Sz} + r) \tag{7}$$

under (P2). These equations are the result of a different solution approach to our problem (and under the more restrictive assumption of v being geometrically distributed). It can be easily verified that Q(z) given by Eq.(4) with the appropriate expression from either Eq.(2) or (3) is a solution to Eqs.(6) and (7) above as it should be.

Assuming that v is geometrically distributed, Kanehira [18] derived an expression equivalent to Eq.(2) for U under protocol (P1). Our model and results in this paper are much more general than his. In particular, v can have an arbitrary probability distribution. In the model of queued users, v corresponds to the number of packets served in a busy period, which is generally not geometrically distributed even if h is.

R-ALOHA channel capacity

At this point, let us investigate the maximum possible throughput of a channel that employs the R-ALOHA protocol. The throughput of a channel is defined to be the fraction of time slots in which data packets are successfully transmitted and is equal to U above for R-ALOHA. The maximum possible throughput of a channel under a specific multiaccess protocol is said to be the channel capacity of that protocol.

Let C_{RA} And C_{SA} denote the channel capacities of R-ALOHA and slotted ALOHA respectively. We then have under protocol (P1)

$$U \le C_{RA} = \frac{C_{SA}}{C_{SA} + (1/\bar{v})} \tag{8}$$

Hence,

$$\frac{C_{SA}}{1+C_{SA}} \le C_{RA} \le 1 \tag{9}$$

for $\bar{v}$ ranging from 1 to ∞.

Similarly, under protocol (P2)

$$U \le C_{RA} = \frac{C_{SA}}{C_{SA} + [(1-C_{SA})/\bar{v}]} \tag{10}$$

Hence,

$$C_{SA} \le C_{RA} \le 1 \tag{11}$$

for $\bar{v}$ ranging from 1 to ∞.

For a large population of users, we know that [9,12]

$$C_{SA} = 1/e$$

The channel capacity of R-ALOHA is shown in figure 4 for a large population of users for both protocols (P1) and (P2).

So far we have derived some very useful results in terms of S and $\bar{v}$. We know that $\bar{v} = \bar{h}$ for the model of single-message users. However, $\bar{v}$ is still unknown for the model of queued users and S is unknown for both user models. Their derivations will be given below. The equilibrium channel throughput, however, is easily obtained using the following argument without first determining S and $\bar{v}$.

Since the system is assumed to be in equilibrium, the channel throughput rate must be equal to the channel input rate. We therefore have for the model of single-message users

$$\begin{aligned} U &= \text{channel input rate in packets/slot} \\ &= (N-\bar{m})\,\lambda\,\bar{h}\ T/M \\ &= (N-MU)\lambda\bar{h}\ T/M \end{aligned}$$

from which we get

$$U = \frac{(N\lambda\bar{h}T)}{M(1+\lambda\bar{h}T)} \tag{12}$$

For the model of queued users, we simply have

$$U = N\lambda\bar{h}T/M \tag{13}$$

Message delay analysis

We shall assume that the delay d_A incurred by a user to successfully transmit a packet into a non-reserved time slot has a known probability density function (pdf) with the Laplace transform $D_A^*(s)$ and mean value $\bar{d}_A$. Note, however, that although the delay incurred by a packet in a slotted ALOHA channel has known results [12,15,19], such results must be modified for our use here. The slotted ALOHA time slots imbedded within a R-ALOHA channel are not contiguous. The number of time slots in between slotted ALOHA slots is a random variable J with

$$\text{Prob}[J=j] = U^j(1-U) \qquad j=0,1,2,...$$

under our earlier assumptions.

For the model of single-message users, the overall delay d of a message consists of d_A and the transmission delay of the rest of the message (if more than 1 packet long). The pdf of message delay has the Laplace transform

$$D^*(s) = D_A^*(s)H(e^{-sT})e^{sT} \tag{14}$$

with mean

$$\bar{d} = \bar{d}_A + (\bar{h}-1)T \tag{15}$$

For the model of queued users, the overall delay d of a message can be obtained by considering each user queue as a generalized M|G|1 queue in which the first customer of each busy period receives exceptional service [20]. In that context, the service time pdf of customers who initiates busy periods has the Laplace transform

$$B_0^*(s) = D_A^*(s)H(e^{-sT})e^{sT} \quad (16)$$

with mean

$$\bar{x}_0 = \bar{d}_A + (\bar{h}-1)T$$

The service time pdf of customers who arrive to find the queue busy has the Laplace transform

$$B^*(s) = H(e^{-sT}) \quad (17)$$

with mean

$$\bar{x} = \bar{h}T$$

The pdf of message delay has the Laplace transform [20]

$$D^*(s) = \frac{P_0[(\lambda-s)B_0^*(s) - \lambda B^*(s)]}{(\lambda-s) - \lambda B^*(s)} \quad (18)$$

where

$$P_0 = \frac{1-\lambda\bar{x}}{1-\lambda(\bar{x}-\bar{x}_0)}$$

$$= \frac{1-\lambda\bar{h}T}{1+\lambda(\bar{d}_A-T)} \quad (19)$$

The average message delay is

$$\bar{d} = \frac{\overline{x_0}}{1-\lambda(\bar{x}-\bar{x}_0)} + \frac{\lambda(\overline{x_0^2} - \overline{x^2})}{2[1-\lambda(\bar{x}-\bar{x}_0)]} + \frac{\lambda\overline{x^2}}{2(1-\lambda\bar{x})} \quad (20)$$

where $\overline{x^2}$ and $\overline{x_0^2}$ are second moments of service times which can be obtained from $B^*(s)$ and $B_0^*(s)$ respectively.

Solution for S and $\bar{v}$

Most of the analytic results above were derived in terms of the variables S and $\bar{v}$. For the model of single-message users, they are given by

$$S = \frac{(N-\bar{n})\lambda T}{(1-U)M}$$

$$= \frac{\lambda T(N-MU)}{(1-U)M} \quad (21)$$

and

$$\bar{v} = \bar{h} \quad (22)$$

For the model of queued users, the Laplace transform $G^*(s)$ of the busy period pdf is first obtained using a delay cycle analysis [21] to be

$$G^*(s) = B_0^*(s+\lambda-\lambda Y^*(s)) \quad (23)$$

where

$$Y^*(s) = B^*(s+\lambda-\lambda Y^*(s)) \quad (24)$$

Let $\bar{g}$ and $\bar{y}$ be the mean values obtained from $G^*(s)$ and $Y^*(s)$ respectively. We have

$$\bar{g} = \bar{x}_0(1+\lambda\bar{y}) \quad (25)$$

where

$$\bar{y} = \frac{\bar{x}}{1-\lambda\bar{x}} = \frac{\bar{h}T}{1-\lambda\bar{h}T} \quad (26)$$

In this user model, $\bar{v}$ is the total number of packets that arrive within a busy period (including those of the initial message). We then have

$$\bar{v} = (1+\lambda\bar{x}_0\bar{k})\bar{h} \quad (27)$$

where

$$\bar{k} = 1/(1-\lambda\bar{x}) \quad (28)$$

Thus

$$\bar{v} = (1 + \frac{\lambda\,\overline{x_0}}{1-\lambda\bar{x}})\bar{h}$$

$$= (1 + \frac{\lambda[\bar{d}_A+(\bar{h}-1)T]}{1-\lambda\bar{h}T})\bar{h} \quad (29)$$

At this point, if we know the slotted ALOHA throughput-delay relationship for the nonreserved time slots (i.e. $\bar{d}_A$ as a function of S), $\bar{v}$ and S can be solved numerically using Eq.(29) together with the equation

$$U = \frac{S}{S+(1/\bar{v})} = N\lambda\bar{h}T/M \quad (30)$$

obtained from Eqs.(2) and (13) for protocol (P1), or

$$U = \frac{S}{S+[(1-S)/\bar{v}]} = N\lambda\bar{h}T/M \quad (31)$$

obtained from Eqs.(3) and (13) for protocol (P2).

6. CONCLUSIONS

Following a brief overview of multiaccess protocols for packet broadcasting networks, R-ALOHA was introduced as a protocol that can substantially improve the throughput of a broadcast channel over that of slotted ALOHA for users with several packets to send at a time and/or steady arrival streams and buffering capability. It was shown that the R-ALOHA channel capacity ranges from that of slotted ALOHA to 1 (corresponding to the capacity of fixed assigned TDMA channels) depending upon the nature of the input traffic and the user model.

Our analysis approach is based upon the assumption that the nonreserved time slots within the channel which are accessed under a slotted ALOHA protocol is well-behaved (i.e. adaptively controlled) and has known delay characteristics. Although the slotted ALOHA protocol has been studied extensively in the past, this assumption needs to be further investigated.

Two user models with Poisson arrivals were considered. Each arrival is a message consisting of a group of packets. In the first model, each user handles one message at a time. In the second model, each user has infinite buffering capacity for queueing. An equilibrium probability distribution for the number of used slots in a frame was obtained. This result leads to a number of formulas for channel throughput and channel capacity. Laplace transforms of the probability density functions of message delay and busy period were also derived. Some numerical results were illustrated. Additional numerical results are being done to fully understand the performance tradeoffs of R-ALOHA given by analytic results obtained in this paper.

Acknowledgement

The author thanks Debbie Davis for her typing assistance.

REFERENCES

1. Crowther, W. et al., "A System for Broadcast Communication: Reservation-ALOHA," Proc. 6th HICSS, University of Hawaii, Honolulu, January 1973.
2. Abramson, N. and E. R. Cacciamani, "Satellites: not just a big cable in the sky," IEEE Spectrum, Vol. 12, Sept. 1975.
3. Kahn, R. E., "The Organization of Computer Resources into a Packet Radio Network," IEEE Trans. on Commun., Vol. COM-25, Jan. 1977.
4. West, L. P., "Loop-Transmission Control Structures," IEEE Trans. on Commun., Vol. COM-20, June 1972.
5. Metcalfe, R. M. and D. R. Boggs, "Ethernet: Distributed Packet Switching for Local Computer Networks," Commun. of the ACM, Vol. 19, July 1976.
6. Schwartz, M., Computer-Communication Network Design and Analysis, Prentice-Hall, Englewood Cliffs, N. J., 1977, chapter 12.
7. Hayes, J. F., "An Adaptive Technique for Local Distribution," Conf. Rec. Int. Conf. on Communications, Chicago, June 1977.
8. Abramson, N., "THE ALOHA SYSTEM-Another alternative for computer communications," AFIPS Conf. Proc., Vol. 37, 1970, pp. 281-285.

9. Lam, S. S., "Satellite Multiaccess Schemes for Data Traffic," Conf. Rec. of the Int. Conf. on Communications, Chicago, June 1977.
10. Kleinrock, L. and F. A. Tobagi, "Packet Switching in Radio Channels: Part 1--Carrier Sense Multiple Access Modes and their Throughput--Delay Characteristics," IEEE Trans. on Comm., Vol. COM-23, December 1975.
11. Kleinrock, L. and M. Scholl, "Packet Switching in Radio Channels: New Conflict-Free Multiple Access Schemes for a Small Number of Data Users," Conf. Rec. Int. Conf. on Communications, Chicago, June 1977.
12. Lam, S. S., "Packet Switching in a Multi-Access Broadcast Channel with Application to Satellite Communication in a Computer Network," Ph. D. Dissertation, Computer Science Department, Univ. of California, Los Angeles, March 1974.
13. Kleinrock, L. and S. S. Lam, "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation," IEEE Trans. on Commun., Vol. COM-23 April 1975.
14. Lam, S. S. and L. Kleinrock, "Packet Switching in a Multiaccess Broadcast Channel: Dynamic Control Procedures," IEEE Trans. on Commun., Vol. COM-23, Sept. 1975.
15. Ferguson, M. J., "On the Control, Stability, and Waiting time in a Slotted ALOHA Random-Access System," IEEE Trans. on Commun., Vol. COM-23, November 1975.
16. Gerla, M. and L. Kleinrock, "Closed Loop Stabiliity Controls for S-Aloha Satellite Communication," Proc. Fifth Data Commun. Symp., Snowbird, Utah, September 1977.
17. Cox, D. R., Renewal Theory, Methuan, London, 1962, p. 84.
18. Kanehira, E. M., "Transient and Steady-State Analysis of Reservation ALOHA," ARPA Network Information Center, Stanford Research Institute, ASS Note 35 (NIC 14180), February, 1973.
19. Kleinrock, L. and S. S. Lam, "Packet-Switching in a slotted satellite channel," AFIPS Conf. Proc., Vol. 42, AFIPS Press, Montvale, N. J., 1973, pp. 703-710.
20. Welch, P. D., "On a Generalized M/G/1 Queueing Process in which the First Customer of Each Busy Period Receives Exceptional Service," Operations Research, Vol. 12, 1964, pp. 736-752.
21. Kleinrock, L., Queueing Systems Vol. 2: Computer Applications, Wiley-Interscience, New York, 1976, p. 113.

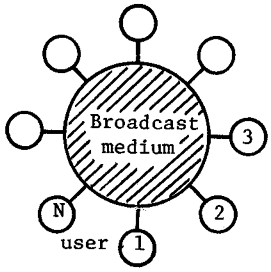

Fig. 1. Logical view of a broadcast network.

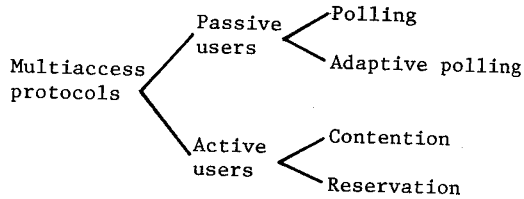

Fig. 2. A classification of multiaccess protocols.

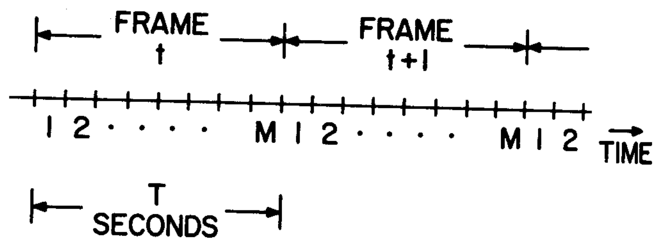

Fig. 3. Frame structure.

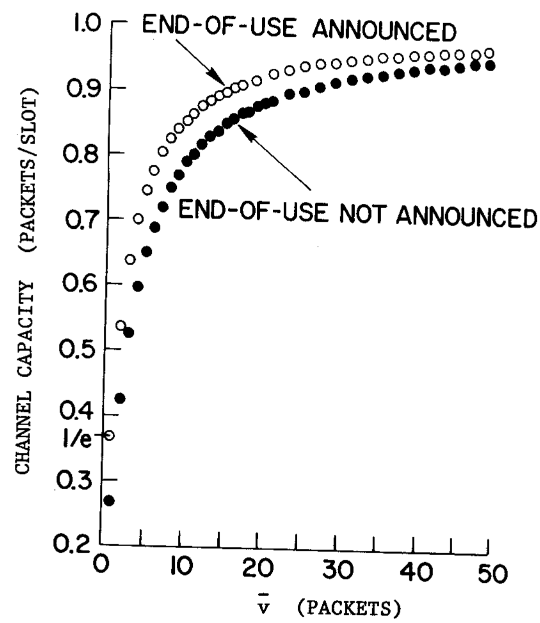

Fig. 4. R-ALOHA channel capacity versus $\bar{v}$.