#189

ON THE ANALYSIS AND SIMULATION OF BUFFERED PACKET RADIO SYSTEMS*

Fouad A. Tobagi and Leonard Kleinrock
Computer Science Department
University of California, Los Angeles 90024

Abstract

Consider a (small) number of buffered terminals communicating over a packet-switched radio channel. The allocation of bandwidth among the contending terminals can be fixed (e.g., TDMA), centrally controlled (e.g., polling) or random (e.g., ALOHA or CSMA). These allocation techniques are compared with respect to throughput and delay. With random access, the success of a packet transmission is dependent on the state of the system and this results in a very difficult analysis problem leading to the use of simulation.

I. INTRODUCTION

Numerous papers have already appeared in the literature which discuss the advantage of using radio as an alternative to wire communication for terminal-to-computer communication [1]. Here consider an environment consisting of a population of M identical user terminals wishing to communicate with a central station over a shared radio channel of limited bandwidth, say W Hz. The basic question is how to allocate this bandwidth among the contending packet-switched terminals such that the limiting communications resource is efficiently utilized and such that the terminals' delays are within an acceptable range. The various known alternatives fall into the following categories:

(i) Fixed assignment such as FDMA and TDMA.
(ii) Random Access (no assignment) such as ALOHA and Carrier Sense Multiple Access (CSMA) [1,2].
(iii) Centrally controlled assignment such as polling and reservation [1,3].

Much analysis has already been developed to determine and compare the performance of the above techniques. However, in analyzing the random access modes and in comparing the above techniques, it has often been assumed that the environment consists of a large population of terminals each of which generates packets with an infinitesimally small rate, and that each packet can be successfully transmitted in a time interval much less than the average time between successive packets generated by a given user. Each user in the large population is assumed to have at most one packet requiring transmission at any time (including any previously blocked packet). There are numerous situations in ground radio and satellite communication systems where we are in presence of terminals generating packets at higher input rates than assumed above. Good examples are: file users, concentrators, repeaters in a packet radio system, ...,etc. These terminals are few in number and require buffering capability for proper operation, i.e., they may have more than one packet in their queue ready for transmission.

In the following, we give the delay equations for both polling and TDMA, show the difficulty in analyzing slotted ALOHA with buffering capability, and finally give some simulation results which we use in comparing the above mentioned assignment techniques for small M .

II. ANALYSIS

This paper is not the first to deal with finite populations of terminals in a slotted ALOHA channel. Abramson [4] provided a model allowing the determination of channel capacity. Gitman [5] used the same model to study a broadcast network in which the origination devices cannot reach the

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

destination in one hop, thus introducing a (small) number of repeaters with buffering capabilities. However the model allows only the determination of capacity under some strong assumptions of independence, which were not verified, and hence gives no solution to the problem of queue size and packet delay. Other papers have suggested some efficient techniques to use a radio channel in an environment including large users. Roberts [3] introduced and analyzed a reservation scheme suitable to satellite ground stations where the generated messages can be single packets or multi-packet blocks; Binder [6] introduced and simulated a multi-access scheme combining a dynamic allocation technique with time-division multiplexing or Round Robin. We shall not concern ourselves with these techniques in this short presentation.

In this section, we focus on various attempts made in analyzing a slotted ALOHA channel accessed by M identical users with buffering capabilities, and the difficulty encountered because of the dependence among the queues. Before we do so, we first briefly present the delay equations for both the TDMA and polling techniques.

2.1 Fixed Assignment

A number of disadvantages of FDMA exist when compared with TDMA: wasted bandwidth for adequate frequency separation, lack of flexibility in achieving dynamic allocation of bandwidth, lack of broadcast operation. The only major disadvantage in TDMA is the need to provide rapid burst synchronization and sufficient burst separation to avoid time overlap. However, in a satellite communication environment, INTELSAT's MAT-1 experimental TDMA system has shown that guard bands of less than 200 nanosec are achievable and new operational systems are moving towards the use of TDMA. The average packet delay in a synchronized TDMA system with a Poisson arrival process can be shown to be

$$D = M\left[\frac{\lambda}{2(1-\lambda)} + 1 + \frac{1}{2}\right] \text{ (slots)} \qquad (1)$$

where λ is the average number of arrivals per slot from all users. λ is also the channel utilization. We note that for a fixed channel utilization, D is proportional to M.

2.2 Polling

In the polling technique, the central station asks the terminals one by one in sequence whether they have anything to transmit. 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. Message packets arriving at a terminal are queued in its buffer until the terminal is polled, at which time the buffer is completely emptied. Konheim and Meister [7] analyzed this polling technique deriving stationary distributions for queue lengths and waiting times. In application to packet radio, the expected packet delay is given by (see reference [1])

$$D = \frac{1}{2}\,\frac{\lambda}{1-\lambda} + 1 + \frac{\alpha}{2}\left(1-\frac{\lambda}{M}\right)\left[1 + \frac{Mr}{1-\lambda}\right]$$

where

$\alpha = \tau/T_m$ and $r = T_p/\tau + 2$ if $T_p/\tau \geq 1$

$\alpha = T_p/T_m$ and $r = 1 + 2\tau/T_p$ if $T_p/\tau < 1$

τ = one way propagation delay from terminal to station (or satellite transponder)

T_m = transmission time of a message packet

T_p = transmission time of a polling packet

2.3 Slotted ALOHA

The protocol and the model:
The time axis is divided into slots of equal size and equal to the transmission time of a packet. A terminal which has a non-empty queue will avoid repeated conflicts by transmitting the packet at the head of the queue in the next slot with probability p. This corresponds to there being eligible slots for a given terminal, and the intervals between these slots being geometrically distributed. The arrival process to each queue is assumed, in this section only and for simplicity of analysis, to be Bernouilli with parameter λ. It is to be noted that the service time of a packet at a terminal is dependent on the state of the remaining queues. Let n_i denote the number of packets accumulated at queue i at a given slot. The state of the system is the vector $(n_1, n_2, \ldots, n_i, \ldots, n_M)$. If there are exactly k non empty queues in the system, then the probability of successfully transmitting a packet from one given such terminal over the slot is $p(1-p)^{k-1}$. For simplicity, let us consider the case $M=2$ and infinite buffer size. Since the state of the system over a slot depends only on the state of the system over the previous slot, we have a Markov chain. We can then determine the transition probabilities and write the equilibrium equations relating the state probabilities $P(n_1, n_2)$. Let $G(z_1,z_2)$ denote the generating function defined as

$$G(z_1, z_2) = \sum_{n_1} \sum_{n_2} P(n_1,n_2)\, z_1^{n_1} z_2^{n_2}$$

The marginal distribution for the number in queue can be shown to have the following generating function

$$G(z_1,1) = \frac{\left[\lambda z_1 + 1 - \lambda\right]\left[p^2 G(z_1,0) - p(1-p)G(0,1) - p^2 P(0,0)\right]}{\left[1-p(1-p)\right]\lambda z_1 - p(1-p)(1-\lambda)}$$

The analyticity condition and the condition $G(1,1)=1$ are not sufficient to solve for all the unknowns. The joint probabilities $P(n_1,0)$ remain undetermined. Thus a closed form solution has not been obtained and other approaches have to be investigated. Numerical procedures can be employed for finite buffer size; however, because of the rapidly increasing size of the state space when either M or the buffer size increase, these techniques are not particularly attractive. What about approximations? One approximate solution is given by assuming each terminal to be alone in the system. The solution is easily given by solving the Bernouilli/geometric/1 queueing system. However, this approximation is valid only for very small values of buffer content and channel utilization. A heavy traffic approximation is possible by assuming the other queue to be always busy. Such a solution is valid only for high values of utilization. A third approximation consists of assuming that the state of the queues is independent over time; that is, the probability that queue 2 is non-empty is a constant ρ (over all slots). Therefore the probability of successfully transmitting from queue 1 is (assuming identical queues)

$$P_s = p(1-\rho p)$$

where $\rho = \lambda/P_s$. From this equation we derive $\lambda = \rho p(1-\rho p)$ which, denoting ρp by G, is identical to Abramson's model in [4]. This approximation was used in this last reference to predict the maximum channel throughput. However, this model is unsatisfactory in estimating the throughput-delay performance of the system. For example, for fixed λ, ρp is constant. Increasing p would result in decreasing ρ, which is wrong since when p increases beyond a certain value, the interference level increases and produces a negative effect (see the simulation results in the following section). Thus no approximation was found that could give a good prediction of the system performance over a reasonably wide range of the parameters (and this range cannot be easily determined). Simulation appears once more to be the only recourse in studying slotted ALOHA in finite population environments.

III. RESULTS

The emphasis in this paper is on small values of M ($1 \leq M \leq 10$). In Fig. 1, we plot the normalized delay in slotted ALOHA, obtained by simulation, versus p for M=2 and various values of λ. Similar curves can be plotted for other values of M. For each value of λ there is an optimum value of p, the choice of which is critical especially for large λ. Minimum delay versus throughput can be plotted for various values of M. Simulation results have shown that the throughput-delay performance deteriorates as M increases; with M=5, our simulation shows that we already reach the infinite population performance. In Fig. 2 we plot the normalized delay versus throughput for the following cases with Poisson arrivals:

(1) TDMA, M= 2,5,10

(2) Slotted ALOHA, M=2,∞

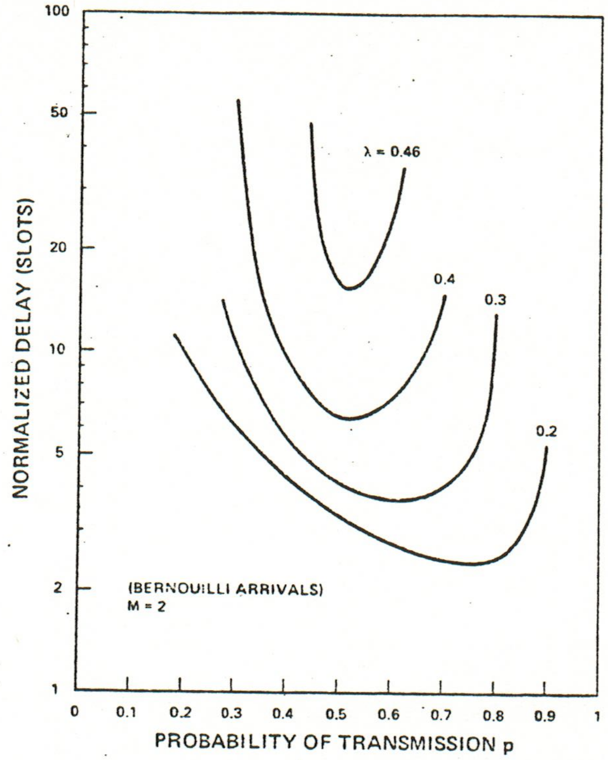

Fig. 1. Slotted ALOHA: Delay versus p for M=2

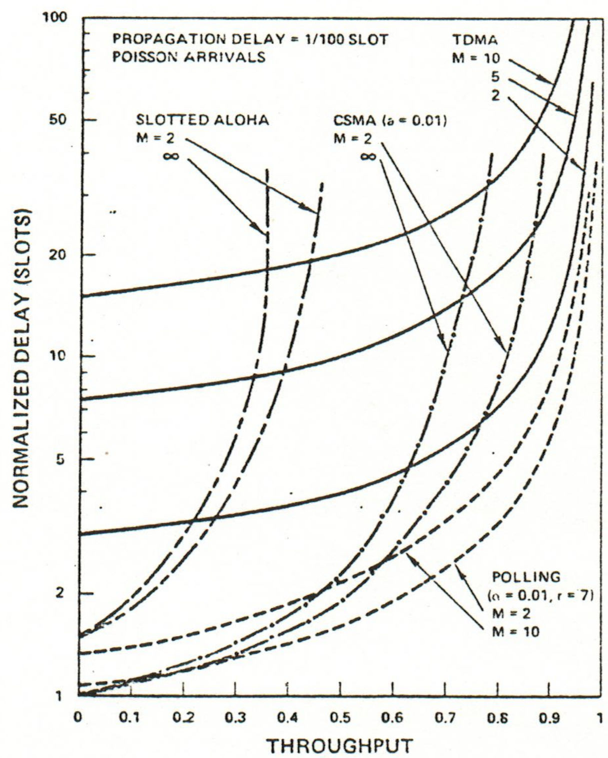

Fig. 2. Delay versus Throughput in ground radio

(3) Non-persistent CSMA, $M=2,\infty$

(4) Polling, $M=2,10$; $T_p/\tau=5$

Fig. 2 corresponds to ground radio systems for which the propagation delay is a negligible fraction of the transmission time of a packet on the channel. For the polling scheme, we assume $\tau/T_m=0.01$ and $T_p/T_m=0.05$. (This corresponds to 1000 bit message packets and 50 bit polling packets). It is evident from this graph that polling provides a high channel utilization with low delays. In satellite systems, however, polling and CSMA are inefficient as the propagation delays are very long. In Fig. 3 we compare slotted ALOHA and TDMA for such an environment, and show that when the number of stations is small, TDMA is more efficient than random access.

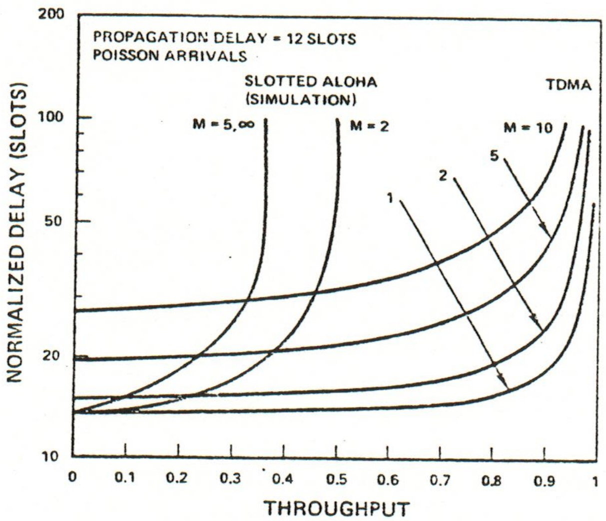

Fig. 3. Delay versus Throughput in satellites

In real problems, the question is: what is the bandwidth required to achieve an average packet delay below an admissible maximum for a given traffic requirement? The bandwidth required, of course, is a function of the assignment technique considered. The answer to this question is given by plotting packet delay (in seconds) versus W, the bandwidth used, for various values of λ (packets per second). An example is given in Fig. 4 which corresponds to ground radio systems (small propagation delay) and M=10. It shows the saving in cost obtained by selecting the proper technique.

To conclude, we note that when comparing the various assignment techniques, in addition to the throughput-delay-cost performance, considerations such as feasibility, simplicity, flexibility and stability must be taken into account.

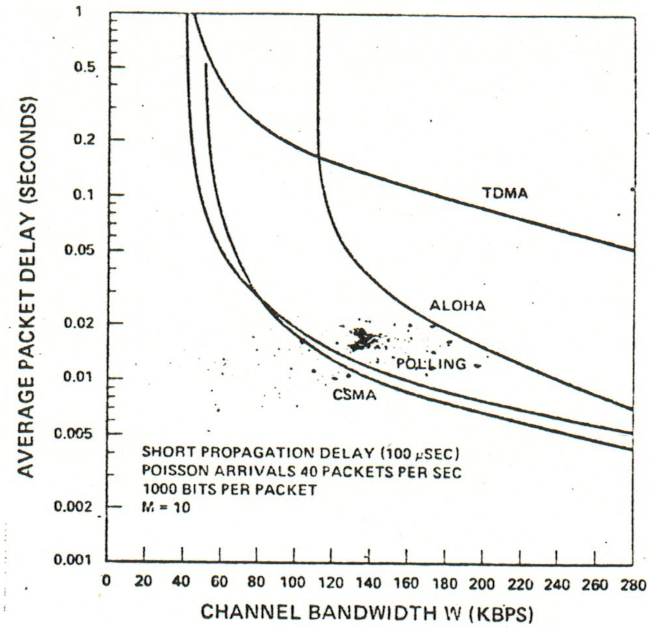

Fig. 4 Delay versus bandwidth

REFERENCES

[1] F. Tobagi, "Random Access Techniques for Data Transmission Over Packet Switched Radio Networks," Comput. Sci. Dep.,UCLA-ENG 7499,Dec.'74.

[2] L. Kleinrock and F. Tobagi, "Packet Switching in Radio Channels: Part I—Carrier Sense Multiple Access Modes and Their Throughput Delay Characteristics," IEEE Trans. Commun., Dec.'75.

[3] L.G. Roberts, "Dynamic Allocation of Satellite Capacity Through Packet Reservation," National Comput. Conf., N.Y., June 4-8, 1973, AFIPS Conf. Proc., 1973, Vol. 42, pp. 711-716.

[4] N. Abramson, "Packet Switching with Satellites," Nat'l. Comput. Conf., AFIPS Conf. Proc., Vol. 42, 1973, pp. 695-702.

[5] I. Gitman, "On the Capacity of Slotted ALOHA Networks and Some Design Problems," IEEE Trans. Commun., Vol. COM-23, March 1975.

[6] R. Binder, "A Dynamic Packet-Switching System for Satellite Broadcast Channels," Proc. of the IEEE Internat'l Conf. on Commun., June 1975.

[7] A.G. Konheim and B. Meister, "Waiting Lines and Times in a System with Polling," IBM J. Watson Research Center, Yorktown Heights, New York, Report RC 3841, May 1972.

BIOGRAPHIES

Fouad A. Tobagi was born in Beirut, Lebanon on July 18, 1947. He received the Engineering Degree from Ecole Centrale des Arts of Manufactures, Paris, France, in 1970 and the M.S. and Ph.D. degrees in computer science from UCLA in 1971 and 1974, respectively. Since December 1974 he has been a Research Staff Project Manager with the APRA project, Computer Science Department, UCLA. His current research interests include computer communication networks, packet switching over radio and satellite networks.

Leonard Kleinrock, Professor, Computer Science Department, UCLA. B.E.E., 1957, CCNY; M.S.E.E., 1959, and Ph.D., 1963, M.I.T. Assistant Engineer, Photobell Co., N.Y. from 1951-1957; Staff Associate, the M.I.T. Lincoln Laboratory, 1957-1963. In 1963 he joined the faculty at UCLA. Referee, and consultant for research and governmental organizations. Principal Investigator of a large contract with ARPA. Published over 60 papers; author, Communication Nets; Stochastic Message Flow and Delay; Queueing Systems, Vol. I-Theory (1975), and Vol. II-Computer Applications (in press). Guggenheim Fellow, 1971-1972. IEEE Fellow.