The Tone Sense Multiaccess Protocol with Partial Collision Detection (TSMA/PCD) for Packet Satellite Communications

Man-Keung Lo and Tak-Shing P. Yum

Abstract—**The Tone Sense Multiaccess with Partial Collision Detection (TSMA/PCD) protocol is particularly suitable for a packet satellite system serving an area with a dense population of earth stations. By incorporating a narrow-band ground radio channel for broadcasting busy tones, the earth stations are able to avoid packet collisions by sensing for the absence of busy tones before transmitting packets. Besides, partial collision detection capability can also be achieved. Single-tone TSMA/PCD gives 97% of the CSMA/CD throughput when $N = 10$ tones are used, while for multitone and slot-by-slot announcement TSMA/PCD protocols, only $N = 8$ and $N = 2$, respectively, are sufficient to drive the system to the CSMA/CD performance.**

I. Introduction

The main advantages inherent to satellite communication are the broadcasting capability, the full connectivity of stations, the flexibility of station organization, the capacity to support mobile users, and the high transmission quality [1]. It has long been thought that the CSMA-type protocols are not suitable for packet satellite systems because of the long round-trip propagation delay. However, if all the stations accessing a particular satellite channel are confined to a metropolitan area, the ability to sense the channel status before transmission can be provided by incorporating a narrow-band ground radio channel for the stations to broadcast busy tones while transmitting packets.

In this paper, we propose an uplink multiaccess protocol called the Tone Sense Multiaccess with Partial Collision Detection (TSMA/PCD) for use in a packet satellite system serving a zone with a dense population of earth stations. The downlink could use TDM with statistical multiplexing provided on board. If the satellite has multiple transponders serving a number of zones, the routing of packets to and from other zones can be done by the satellite.

The turning-on and turning-off of a busy tone make the tone signal look like an on–off keying signal. Let the shortest duration of a tone be denoted as a slot, and let its length be τ. Then the maximum bandwidth of the tone is $2/\tau$ Hz. To minimize the bandwidth occupied by the tone, τ should be as large as possible. For a slot size of 1 ms, the tone signal bandwidth is about 2 kHz. Two or three tones thus occupy about 4–6 kHz, a very small band in the VHF region. As in the slotted CSMA and CSMA/CD protocols [2], [3], the smaller the slot size, the higher the throughput efficiency. But the price to pay is larger terrestrial channel bandwidth. For a fixed terrestrial channel bandwidth, one way to increase the satellite channel efficiency is to increase the transmission duration. This can be done by grouping a number of information units together to form a larger packet. TSMA/PCD can provide better delay performance and can support variable length packets without additional complication when compared with reservation techniques.

Paper approved by the Editor for Multiple Access Strategies of the IEEE Communications Society. Manuscript received May 29, 1987; revised May 14, 1988, July 8, 1989, and August 15, 1991. This paper was presented at the IEEE International Conference on Communications, Philadelphia, PA, June 1988.

The authors are with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong.

IEEE Log Number 9210322.

II. The TSMA/PCD Protocol

For convenience of analysis, we assume the system is synchronized or slotted. We mentioned before that the slot size is taken as the shortest duration of a tone, which in turn is determined by the tone bandwidth requirement. Another constraint on the slot size which is usually nonbinding is that the slot size must be larger than the sum of the maximum ground propagation delay between any two stations in the system and the tone detection time T_{CS} [4]. With the slot size so defined, all stations can know the presence of another transmission in the system within one slot interval.

Packets could have variable lengths. But in the analysis, we only consider fixed packets of length T slots. The round-trip propagation delay on the satellite is R slots. We assume that the combined arrivals of new and retransmitted packets form the stations constitute a Poisson process with rate g packets per time slot.

We assume that the satellite channel is noise-free (the noisy channel case can be found in [6]), and in a packet collision, all packets involved are destroyed. Acknowledgments (positive or negative) are piggybacked on the downlink packets, and the transmitting station can receive the acknowledgment $R + 2T$ slots later. When a busy tone is issued, it takes at most one slot for the other stations to sense. If the other transmitting stations are not sending the same busy tone, the collision can be detected by all stations. Thus, if a collision is detected, the maximum duration from the beginning of a transmission until the channel is again idle is two slots. We also assume that the tone channel has a high SNR so that the probabilities of miss and false alarm are both zero.

A. Nonpersistent TSMA/PCD

A station with a packet ready for transmission will first sense the ground channel. If it is busy, the station resenses the channel after a random delay which is uniform on $[1, K]$.

If it is not, the packet is transmitted, and a randomly chosen busy tone is broadcast.

If a collision is detected during the first slot, then: 1) the packet transmission is aborted immediately 2) the tone is terminated at the end of the first slot, and 3) the station resenses the channel after a random delay (uniform on $[1, K]$).

If collision is not detected, the sending station waits for the acknowledgment from the satellite. If a negative acknowledgment is received, the station repeats the entire procedure.

B. 1-Persistent TSMA/PCD

The 1-persistent TSMA/PCD is similar to the nonpersistent TSMA/PCD, except that when the ground channel is sensed busy, a ready station waits until the channel goes idle and transmits its packet immediately.

III. Analysis of Nonpersistent TSMA/PCD

A. Throughput

Let K be the maximum retransmission interval in slots, and let N be the total number of busy tones from which a station is allowed to choose randomly. A successful transmission period is $T + 1$ slots. For an unsuccessful transmission period, the length is either $T + 1$ slots if collision is undetected or two slots if collision is detected. We define a transmission period and its immediate following idle period to be a cycle. In steady state, let P_S be the probability that a transmission is successful. Then P_S, the probability that there is exactly one arrival in the slot immediately before the transmission period given that there is at least one arrival occurs, is

$$P_S = \frac{ge^{-g}}{1 - e^{-g}}.$$

Next, let P_u be the probability that a collision is undetected given that a collision has occurred, let α_k be the probability that k ready stations all choose the same tone for transmission announcement, and let β_k be the probability that k stations become ready given that the number of ready stations is at least two. Then P_u can be evaluated as

$$\begin{aligned} P_u &= \sum_{k=2}^{\infty} \alpha_k \beta_k = \sum_{k=2}^{\infty} \left[\frac{1}{N}\right]^{k-1} \frac{g^k e^{-g}}{k!(1 - e^{-g} - ge^{-g})} \\ &= \frac{e^{-g}\left(Ne^{g/N} - N - g\right)}{(1 - e^{-g} - ge^{-g})}. \end{aligned} \quad (1)$$

Hence, the expected duration of a transmission period $E[TP]$ is

$$\begin{aligned} E[TP] = (T+1)P_S &+ 2(1 - P_S)(1 - P_u) \\ &+ (T+1)(1 - P_S)P_u. \end{aligned}$$

We include the slot containing the arrival in the idle period I. Therefore,

$$E[I] = \frac{1}{1 - e^{-g}}. \quad (2)$$

The channel throughput S is given by

$$S = \frac{P_S T}{E[TP] + E[I]}. \quad (3)$$

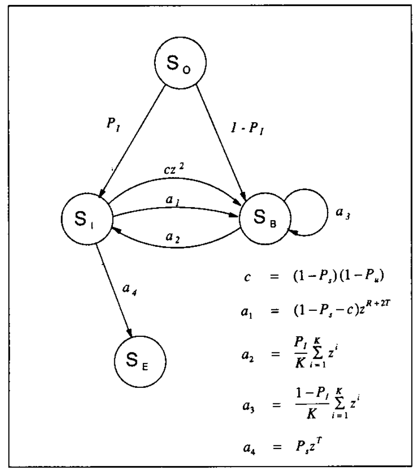

Fig. 1. Flow graph of packet delay process, nonpersistent TSMA/PCD.

B. Delay

Under the Poisson arrival assumption and in steady state, the distribution of busy and idle periods seen by an arrival is their respective equilibrium distributions. Therefore, in steady state, the probability that a packet arrives at an idle period P_I is

$$P_I = \frac{E[I]}{E[TP] + E[I]}.$$

The probability of arrival at a busy period P_B therefore is $P_B = 1 - P_I$.

We define the packet delay D to be the time elapsed in slots from the moment of its first attempt until it is successfully transmitted. Fig. 1 shows the flow graph [5] of the packet delay process of nonpersistent TSMA/PCD. The circles represent the states encountered by the packet being transmitted. Specifically, we have the following.

1) S_o is the starting state. With probability P_I, the packet will start from state S_I and with probability $(1 - P_I)$, it will start from S_B. The transition from S_o to either S_I or S_B is instantaneous.

2) S_I indicates that the packet arrives at an idle period.

3) S_B indicates that either

a) the packet arrives at a busy period or

b) a collision is detected or

c) acknowledgment for the packet is not received after $R + 2T$ slots (undetected collision).

4) S_E indicates that the packet is successfully transmitted.

At state S_I, the packet is transmitted immediately. If the transmission is successful, the packet will enter S_E after T slots with probability P_S. If transmission is unsuccessful and collision is *detected,* which happens with probability $(1-P_S)(1-P_u)$, the packet will enter S_B after two slots. If the transmission is unsuccessful and collision is *undetected* (with probability $(1 - P_S)P_u$), the packet will also enter S_B, but after $R + 2T$ slots. At state S_B, Scheduling for retransmission is made. Since a retransmission attempt takes place randomly at one of the next K slots and will encounter an idle period with probability P_I, the packet will enter S_I after n slots

$1 \leq n \leq K$ with probability $P_I(1/K)$. The total "flow" from S_B to S_I therefore is $(P_I/K)(z + z^2 + \cdots + z^K)$ where z^i indicates i slots of delays. The "flow" from S_B back to itself is found by a similar argument.

The flow diagram indicates that the transition probabilities from a particular state depend only on that state; hence, the delay process is a transient semi-Markov process, with S_E being the absorbing state. The duration of the packet delay D is just the first passage time from S_o to S_E. For simplicity, let $c = (1 - P_S)(1 - P_u)$. Then the following flow equations are readily obtained from Fig. 1:

$$\begin{aligned} S_I &= P_I S_o + \frac{P_I}{K} \sum_{i=1}^{K} z^i S_B \\ S_B &= cz^2 S_I + (1 - P_S - c) z^{R+2T} S_I \\ &\quad + \frac{1 - P_I}{K} \sum_{i=1}^{K} z^i S_B + (1 - P_I) S_o \\ S_E &= P_s z^T S_I . \end{aligned}$$

The generating function of packet delay $G_D(z)$ is just the transfer function from state S_o to S_E or (see equation at bottom of this page).

The delay distribution can be evaluated by numerical inversion. The mean packet delay and delay variance are obtained by differentiating $G_D(z)$. Details appear in [6].

IV. Analysis of 1-Persistent TSMA/PCD

A. Throughput

Let $P_u(X)$ be the probability that the collision is undetected given that the present transmission is unsuccessful and the length of the preceding transmission period is X, and let γ_k be the probability that k stations become ready in the preceding transmission period X given that the number of ready stations is at least 2. The throughput analysis closely follows the method given in [3], except that $P_u(X)$ is given by the following instead:

$$\begin{aligned} P_u(X) &= \sum_{k=2}^{\infty} \alpha_k \gamma_k = \sum_{k=2}^{\infty} \left[\frac{1}{N}\right]^{k-1} \frac{(gX)^k e^{-gX}}{k!(1 - e^{-gX} - gXe^{-gX})} \\ &= \frac{e^{-gX}\left(Ne^{gX/N} - N - gX\right)}{(1 - e^{-gX} - gXe^{-gX})} . \end{aligned} \tag{4}$$

We refer the reader to [6] for the remaining details.

B. Delay

If a collision is detected, all stations involved will retry after a random delay which is uniform of $[1, K]$. If the collision is not detected, whether retranmission is needed or not is determined from the acknowledgment of the satellite after $R + 2T$ slots. The eight-state flow graph of the packet delay process and the derivation of the mean and variance of the delay can be found in [6].

V. Multitone TSMA/PCD

In multitone TSMA/PCD, stations broadcast a group of n tones to announce their transmissions. These groups of n tones are chosen randomly from a preassigned pool of N tones. A packet collision is detected if the transmitting station senses busy tones other than its own being broadcast on the ground channel. For the same number of tones N, the multitone scheme yields a lower probability of undetected collision. For the nonpersistent case,

$$\begin{aligned} P_u &= \sum_{k=2}^{\infty} \left(\frac{1}{\beta}\right)^{k-1} \frac{g^k e^{-g}}{k!(1 - e^{-g} - ge^{-g})} \\ &= \frac{e^{-g}\left(\beta e^{g/\beta} - \beta - g\right)}{1 - e^{-g} - ge^{-g}} \end{aligned} \tag{5}$$

where β is equal to $\binom{N}{n}$. For the 1-persistent case,

$$\begin{aligned} P_u(X) &= \sum_{k=2}^{\infty} \left(\frac{1}{\beta}\right)^{k-1} \frac{(gX)^k e^{-gX}}{k!\,(1 - e^{-gX} - gXe^{-gX})} \\ &= \frac{e^{-gX}\left(\beta e^{gX/\beta} - \beta - gX\right)}{1 - e^{-gX} - gXe^{-gX}} . \end{aligned} \tag{6}$$

The value of n that minimizes P_u and $P_u(X)$ can be shown to be at $n = [N/2]$. The throughput and delay analyses follow exactly that of the single-tone case.

VI. Slot-by-Slot Announcement TSMA/PCD

The collision detection capability can also be increased by using slot-by-slot announcements. In this scheme, instead of continuously broadcasting the same busy tone for transmission announcement, a transmitting station rechooses a tone (randomly from the set of N tones) to broadcast in each slot. Thus, a collision is not detected after the nth slot if and only if all transmitting stations choose the same tone in slot 1, slot $2, \cdots$, and slot n. This probability is very small, even for very small N. As before, if collision is detected, the packet and tone transmissions are aborted immediately and will retry after a random delay uniform on $[1, K]$. Note that the minimum tone announcement duration should be equal to the tone detection time, which is less than one slot. Here, we choose the announcement duration to be exactly one slot for convenience. In the following, we derive the throughput and delay performance of the nonpersistent version.

Let $P_D(n|k)$ be the probability that a collision is detected in the nth slot from the beginning of a transmission given that

$$G_D(z) = \frac{S_E}{S_o} = \frac{KP_I P_S (1 - z) z^T}{K(1 - z) - z(1 - z^K)[1 - P_I + cP_I z^2 + P_I (1 - P_S - c) z^{R+2T}]} .$$

there are $k (k \geq 2)$ arrivals in the slot before the transmission period. Then,

$$P_D(n|k) = \left[\left(\frac{1}{N}\right)^{k-1}\right]^{n-1}\left[1-\left(\frac{1}{N}\right)^{k-1}\right]$$
$$= N^{-n(k-1)}(N^{k-1}-1) \qquad n = 1, 2, \cdots, T.$$

The probability that a collision is detected in the nth slot given that a collision has occurred therefore is

$$P_D = \sum_{k=2}^{\infty} P_D(k)\frac{g^k e^{-g}}{k!\,(1-e^{-g}-ge^{-g})}$$
$$= \frac{N^{n-1}e^{-g}\left[\exp(g/N^{n-1}) - N\exp(g/N^n) + N - 1\right]}{(1-e^{-g}-ge^{-g})}.$$

Next, the probability that the collision is undetected in the entire transmission period given that there are k transmissions $P_u(k)$ is

$$P_u(k) = \left[\left(\frac{1}{N}\right)^{k-1}\right]^{T} = N^{-T(k-1)}.$$

Hence, the probability that the collision is undetected in the entire transmission period given that a collision has occurred P_u is

$$P_u = \sum_{k=2}^{\infty} P_u(k)\frac{g^k e^{-g}}{k!\,(1-e^{-g}-ge^{-g})}$$
$$= \frac{N^T e^{-g}\left[\exp(gN^{-T}) - 1 - gN^{-T}\right]}{(1-e^{-g}-ge^{-g})}.$$

Using the same notations as in Section III, we now have

$$E[TP] = (T+1)P_S + (T+1)(1-P_S)P_u + (1-P_S)\sum_{n=1}^{T}(n+1)P_D(n).$$

As before, the throughput S is given by

$$S = \frac{P_S T}{E[TP] + E[I]}$$

and the mean and variance of the delay can be found in a similar manner [6].

VII. Numerical Result and Discussion

Fig. 2 shows throughput S as a function of N. Here, $N = 1$ and $N = \infty$ correspond to CSMA and CSMA/CD, respectively. For the nonpersistent case, N as small as 2 can already attain 92% of the maximum throughput. The throughput curve of the 1-persistent version for $N = 2$ is a little strange. An explanation of this can be found in [6]. Fig. 3 shows the delay characteristics. We choose the slot size to be 0.167 ms, and so $R = 1617$ slots. Here, even for $N = 1$ or 2, the nonpersistent protocol can already give a high throughput with very low delay.

Fig. 4 shows the coefficient of variation C_D of packet delay as a function of S with N as a parameter. When $S = 0$, C_D is obviously zero. But when $S \to S_{\max}$, the number of stations involved in a collision is usually fairly large. So, even for small N, the collision detection probability is high. Therefore,

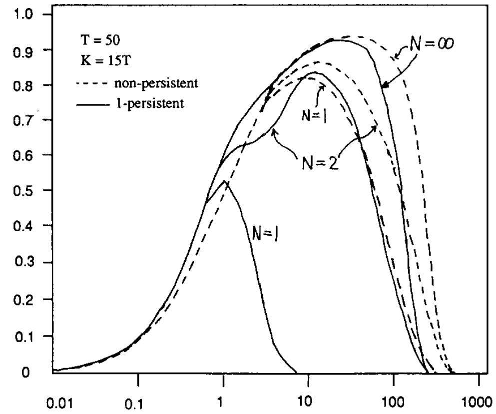

Fig. 2. The throughput of TSMA/PCD with N as a parameter.

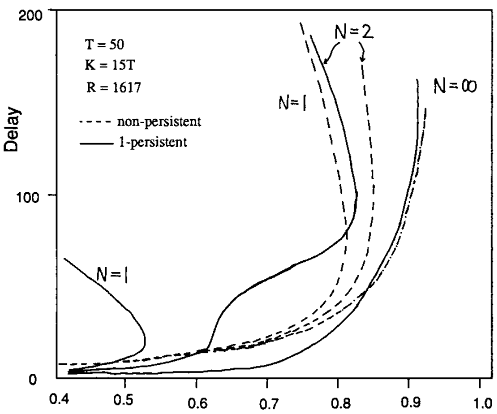

Fig. 3. Throughput–delay performance for nonpersistent and 1-persistent TSMA/PCD.

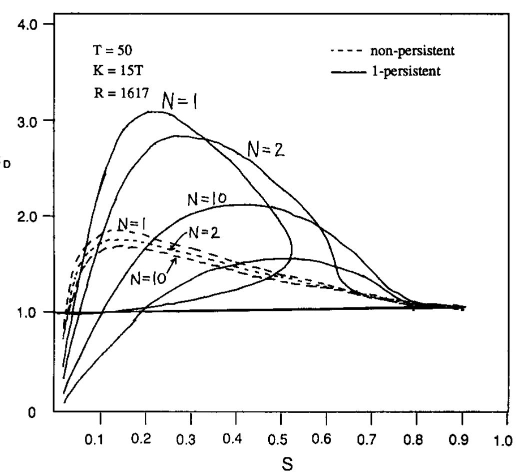

Fig. 4. Coefficient of variation of packet delay.

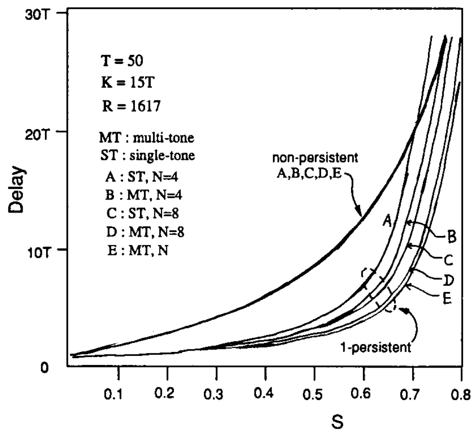

Fig. 5. Comparison of packet delay between multitone and single-tone TSMA/PCD.

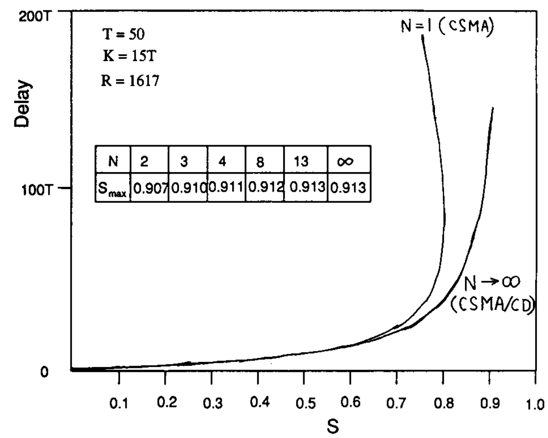

Fig. 6. Throughput–delay characteristics of slot-by-slot announcement scheme for nonpersistent TSMA/PCD.

the major cause of delay in this case is the random scheduling delay after the detection of a collision. In other words, for $G \gg 1$, $D = Y_1 + Y_2 + \cdots + Y_M$ where Y_i is the ith scheduling delay and M is the total number of schedulings before successful transmission and is geometrically distributed. The mean and variance of D are derived in [7] to be

$$E[D] = E[Y]E[M]$$
$$\sigma_D^2 = E[M]\sigma_Y^2 + E^2[Y]\sigma_M^2 .$$

Therefore the squared coefficient of variation is

$$c_D^2 = \frac{\sigma_D^2}{E^2[D]} = 1 + \frac{\sigma_Y^2}{E[M]E^2[Y]} .$$

Here, $E[Y]$ and σ_Y^2 are independent of G, and as $G \to \infty$, $E[M] \to \infty$. Therefore, as $G \to \infty$, $c_D \to 1$. Under light traffic conditions, the delay variance is due primarily to Y. Under moderate traffic conditions, the variances of both M and Y are significant, and hence a peak occurs for c_D.

Fig. 5 compares the single-tone and multitone TSMA/PCD. For the nonpersistent case, the multitone and single-tone versions offer almost identical delay performance. But for 1-persistent TSMA/PCD, the multitone version offers a significant reduction of average delay. For the slot-by-slot announcement nonpersistent protocol, we see from Fig. 6 that for N as small as 2, the maximum throughput (0.907) is already 99.3% of the CSMA/CD maximum throughput. Moreover, the delay is identical to that of the CSMA/CD.

Acknowledgment

The authors would like to thank the reviewers and the Editor for their many valuable comments. The technical content and readability of this paper have considerably improved with their help.

References

[1] I. N. Knight and M. T. Neibert, "An overview of satellite transmission issues and the ISDN," in *Proc. IEEE Int. Conf. Commun.*, Toronto, Canada, June 1986, pp. 1073–1076.
[2] L. Kleinrock and F. A. Tobagi, "Packet switching in radio channels: Part I—Carrier sense multiple access modes and their throughput-delay characteristics," *IEEE Trans. Commun.*, vol. COM-23, pp. 1400–1416, Dec. 1975.
[3] F. A. Tobagi and V. B. Hunt, "Performance analysis of carrier sense multiple access with collision detection," *Comput. Networks*, vol. 4, no. 5, pp. 245–259, 1980.
[4] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels: Part II—The hidden terminal problem in carrier sense multiple-access and the busy-tone solution," *IEEE Trans. Commun.*, vol. COM-23, pp. 1417–1433, Dec. 1975.
[5] R. A. Howard, *Dynamic Probabilistics Systems, Vol. 1: Markov Models.* Wiley, 1971.
[6] M. K. Lo and T. S. Yum, "The tone sense multiaccess protocol with partial collision detection (TSMA/PCD) for packet satellite communications," in *Proc. IEEE ICC,* Philadelphia, PA, June 1988.
[7] L. Kleinrock, *Queueing Systems, Vol. 1: Theory.* New York: Wiley Interscience, 1975.