Delay bounds for packet satellite protocols

Eric W.M. Wong [a] and Tak-Shing Peter Yum [b]

[a] *Department of Electronic Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong*
E-mail: ewong@ee.cityu.edu.hk

[b] *Department of Information Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong*
E-mail: yum@ie.cuhk.edu.hk

Two simple and tight delay lower bounds are derived for packet satellite protocols with memoryless packet arrival process and single copy transmission. One bound is for protocols with contention-free reservation and the other is for protocols with contention-based reservation. The derivation indicates that for minimum delay, a protocol should strive to maintain a balance between transmitting packets immediately and making reservations before transmissions.

1. Introduction

In multiaccess communication systems, the average packet delay is bounded below by the G/G/1 queuing delay [4] with the same interarrival and service time distributions. This delay bound is very loose for packet satellite systems where the round trip propagation delay is long and carrier sensing is not possible. A tighter delay bound is desirable for assessing the possible delay improvement on existing protocols and for deciding whether a particular delay requirement can ever be satisfied.

In this paper, two new delay lower bounds are derived for packet satellite systems with contention-free and contention-based reservations, respectively. The class of protocols whose delays we are trying to bound is of the hybrid random-access/reservation type. This class of protocols includes random access protocols and reservation protocols as special cases and is sufficiently general to be of interest. The environment in which the protocols are to operate is defined by a set of conditions. We shall call this environment ξ and the delay bounds are for the protocols operating in ξ. The conditions defining ξ are:

1) *The packet arrival process is of the memoryless type.* For a finite population model this refers to the Bernoulli process and for an infinite population model, Poisson.

2) *Transmitting multiple copies of the same packet and making multiple reservations for the same packet are not allowed.* Transmitting multiple copies and making multiple reservation might give slightly smaller delay when the traffic is light. Since we have not made any investigation on this, we shall not consider this option.

3) *A single uplink channel is considered.* This condition is not really restrictive because multiple channel systems involve three kinds of inefficiencies:

- the overhead in partitioning a channel into several TDM or FDM subchannels,
- longer transmission time on lower bit rate subchannels,
- multiple reservation queues on the satellite give a longer average delay than a single reservation queue.

4) *Only the slotted channel is considered.* The unslotted channel gives slightly better delay performance only at very very low traffic conditions.

In the following, we will describe the packet satellite system and design an idealized protocol for deriving the delay lower bounds.

2. The packet satellite system

Consider a packet satellite system serving a population of users. Besides the uplink data channel, let there also be an uplink narrow-band control channel for making reservations and a downlink announcement channel for broadcasting successful reservations. In practice, the control channel and the announcement channel can be piggybacked on the up- and down-link data channels, respectively. The data channel is slotted with slot width equal to one packet transmission time. There are two types of slots. Aloha slots are for transmitting packets immediately whereas Reserved slots are for packets with successful reservations. The announcement channel broadcasts the locations of the Reserved slots so that other stations will refrain from transmitting on these slots. All non-Reserved slots are treated as Aloha slots.

3. The idealized protocol with contention-free reservation

Many protocols were proposed for the above system and an extensive survey can be found in [1]. To obtain a delay lower bound for all possible protocols in ξ, we hypothesize an idealized protocol by assuming:

(1) contention-free reservation,

(2) no reservation overflow in the reservation queue,

(3) an optimal balance of the packet traffic rate and the reservation traffic rate in the system,

(4) the traffic statistics after the balancing process is memoryless.

These idealized assumptions guarantee that no practical protocols of the hybrid random-access/reservation type will have a smaller delay than the idealized protocol. The delay of this idealized protocol is therefore a delay lower bound for all practical protocols of the hybrid random-access/reservation type in ξ.

Consider the arrival of a new packet. If it hits an Aloha slot, it will either make a *normal* reservation on the control channel for future transmission or be transmitted in the current Aloha slot with a *spare* reservation made on the control channel. This spare reservation assures that, in case of a collision in the Aloha slot, the retransmission is always successful. If the transmission is successful, the spare reservation is discarded. On the other hand, if the arriving packet hits a Reserved slot, it will either make a normal reservation right away or be transmitted in one of the future Aloha slots.

In a practical protocol, some form of strategy is needed to optimally balance the random-access traffic and the reservation traffic. Since the idealized protocol is used for deriving a delay lower bound, it need not be realizable. An optimal traffic balancing strategy can therefore be assumed as built-in.

All reservations are processed by the satellite and for each successful reservation, a Reserved slot is assigned on the uplink data channel. Since all reservations are assumed to be successful, a packet will encounter at most one collision before successful transmission. A flow chart summarizing this protocol is shown in figure 1.

In the next section, we shall derive the delay of the idealized protocol assuming a finite population model. A similar bound for an infinite population model can be obtained either by letting the population size N go to infinity or by starting from the Poisson arrival model. These bounds turn out to be expressible in closed forms. To tighten these bounds, we relax the assumption of contention-free reservation. The resulting delay lower bound for the protocols with contention-based reservation is derived in section 5.

4. Delay lower bound for protocols with contention-free reservation

Let there be N users in the system. Let λ_a be the average number of transmissions in an Aloha slot, λ_r be the average number of normal reservations (i.e., excluding the spare reservations) per slot on the control channel and S be the throughput of the idealized protocol. It is easily seen that when $\lambda_a > 1$, $\lambda_r \geqslant 1$ or $S \geqslant 1$, the system is unstable so for simplicity, we only consider the case $\lambda_a \leqslant 1$, $\lambda_r < 1$ and $S < 1$ in the following derivation. Let the average number of successful reservations per slot be x. Since each successful reservation is assigned a Reserved slot, x is the average number of packets transmitted through reservation per slot. This also means that x is equal to the probability that a slot is of the reserved type. With the assumption that all reservations are successful, x is derived as

$$\begin{aligned} x &\equiv [\text{ av. no. of successful reservations per slot}] \\ &= \begin{bmatrix}\text{av. no. of normal}\\ \text{reservations}\\ \text{per slot}\end{bmatrix} + \begin{bmatrix}\text{av. no. of remaining}\\ \text{spare reservations}\\ \text{from an Aloha slot}\end{bmatrix} \Pr \begin{bmatrix}\text{a slot is}\\ \text{of the}\\ \text{Aloha type}\end{bmatrix} \\ &= \lambda_r + \left[\lambda_a - \lambda_a(1-\lambda_a/N)^{N-1}\right](1-x), \end{aligned} \tag{1}$$

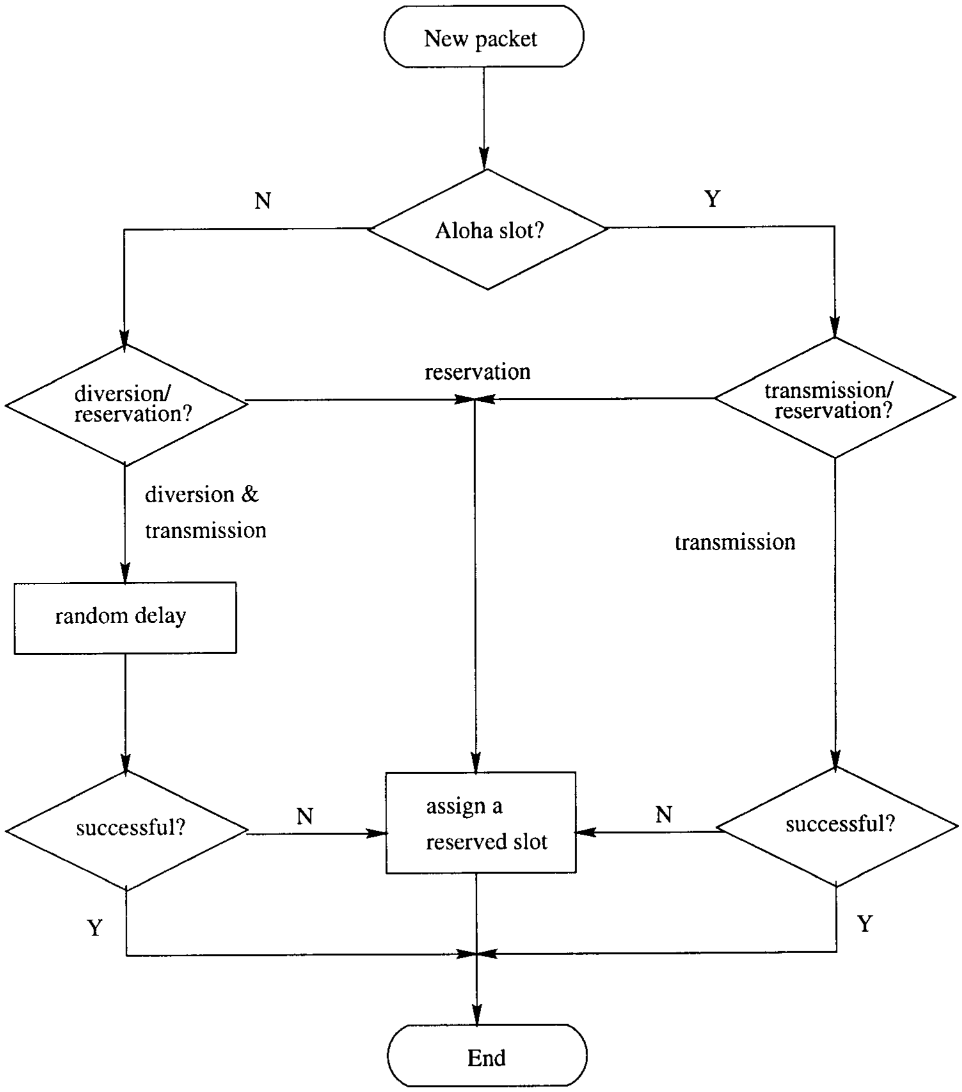

Figure 1. Flow chart of the idealized protocol with contention-free reseration.

where $\lambda_a(1-\lambda_a/N)^{N-1}$ is the average number of successful transmissions in an Aloha slot.

The throughput S of the idealized protocol is given by

$$S = \Pr\begin{bmatrix}\text{a slot is}\\ \text{of the}\\ \text{res. type}\end{bmatrix}\Pr\begin{bmatrix}\text{a res. slot}\\ \text{contains a}\\ \text{succ. tx'n}\end{bmatrix} + \Pr\begin{bmatrix}\text{a slot is}\\ \text{of the}\\ \text{Aloha type}\end{bmatrix}\Pr\begin{bmatrix}\text{an Aloha slot}\\ \text{contains a}\\ \text{succ. tx'n}\end{bmatrix}$$

$$= x + (1-x)\lambda_a(1-\lambda_a/N)^{N-1}. \tag{2}$$

Solving x from (1) and substituting into (2), we have

$$S = \frac{\lambda_r[1 - \lambda_a(1 - \lambda_a/N)^{N-1}] + \lambda_a}{1 + \lambda_a - \lambda_a(1 - \lambda_a/N)^{N-1}}. \tag{3}$$

By differentiating (3), we observed two properties:

Property 1. S is a monotonically increasing function of λ_a and λ_r. (See appendix A.)

Property 2. For a given S, λ_a and λ_r are inversely related functions. (See appendix A.)

The average delay D of the idealized protocol consists of the sum of five terms denoted as D_1 to D_5. The average synchronization delay D_1 is equal to 0.5 slot. The expected reservation delay D_2 is equal to the round trip propagation delay R (in units of slots) multiplied by the probability of transmission through reservation or $D_2 = (x/S)R$. The average waiting time in the reservation queue formed by the reservation traffic, denoted by D_3, is given by the average waiting time in a discrete-time queueing system with bulk arrival of rate x and service time equal to one time unit. From appendix B, we have

$$D_3 = \frac{x(1 - N^{-1})}{2(1 - x)}.$$

The combined packet transmission and propagation time D_4 is equal to $(1 + R)$. The average delay of traffic diversion from the Reserved slots to the Aloha slots is denoted as D_5. Adding up the five terms, we have

$$D = \frac{x(1 - N^{-1})}{2(1 - x)} + R(1 + x/S) + 1.5 + D_5. \tag{4}$$

For the idealized protocol, parameter x in (4) should be chosen such that D is minimum. However, as D_5 involves the specification of the traffic diversion process and is in general much smaller than the round trip propagation delay R, we shall neglect D_5 in the optimization process. In doing so, the delay obtained is only a lower bound for the idealized protocol. This bound is obviously also a lower bound for all protocols in ξ. Let

$$D_L = \frac{x(1 - N^{-1})}{2(1 - x)} + R(1 + x/S) + 1.5. \tag{5}$$

To minimize D_L for a given value of S, (5) stipulates that x should be as small as possible. From (2), x can be expressed as

$$x = 1 - \frac{1 - S}{1 - \lambda_a(1 - \lambda_a/N)^{N-1}}.$$

Differentiating x with respect to λ_a, x is found to have a single minimum at $\lambda_a = 1$. But λ_a and λ_r must also satisfy (3). Therefore, substituting $\lambda_a = 1$ into (3) and solving for λ_r, we obtain

$$\lambda_r = \frac{S[2-(1-N^{-1})^{N-1}]-1}{1-(1-N^{-1})^{N-1}}. \tag{6}$$

Since λ_r must be non-negative, this means that for the above "$\lambda_a = 1$" solution to be valid,

$$S \geqslant \frac{1}{2-(1-N^{-1})^{N-1}} \equiv S_c. \tag{7}$$

At the boundary point $S = S_c$, we have $\lambda_a = 1$ and $\lambda_r = 0$.

When $S < S_c$, the optimal solution is located at $\lambda_r = 0$ (and hence $\lambda_a < 1$ using property 1 and in comparison with the $S = S_c$ case). This can be proved by contradiction. The argument goes as follows. Suppose the optimal solution is at λ_r' where $\lambda_r' > 0$ and hence $\lambda_a < 1$ from property 2, then property 2 states that we can reduce λ_r to zero by increasing λ_a towards one. Since increasing λ_a towards one leads to lower delay, λ_r' cannot be the optimal solution for λ_r. Therefore, $\lambda_r = 0$ must be the optimal solution for $S < S_c$. Then, we set $\lambda_r = 0$ in (3) to obtain

$$S = \frac{\lambda_a}{1+\lambda_a-\lambda_a(1-\lambda_a/N)^{N-1}} \tag{8}$$

and from which the constrained optimum value of λ_a, denoted λ_a^* can be solved numerically. Substituting the optimum λ_a and λ_r into (4), we obtain the delay lower bound $D_L(S, R, N)$ of the idealized protocol as

$$D_L(S,R,N) = \begin{cases} [S-\lambda_a^*(1-\lambda_a^*/N)^{N-1}]\dfrac{(1-N^{-1})}{2(1-S)} \\ \quad + R\dfrac{2S-(1+S)\lambda_a^*(1-\lambda_a^*/N)^{N-1}}{S[1-\lambda_a^*(1-\lambda_a^*/N)^{N-1}]} + 1.5, & S < S_c, \\ \dfrac{S(1-N^{-1})-(1-N^{-1})^N}{2(1-S)} \\ \quad + R\dfrac{2S-(1+S)(1-N^{-1})^{N-1}}{S[1-(1-N^{-1})^{N-1}]} + 1.5, & S \geqslant S_c. \end{cases} \tag{9}$$

It can be shown that

$$D_L(S,R,N) < D_L(S,R,N+1), \quad N = 1,2,\ldots.$$

In the limit $N \to \infty$, (9) becomes

$$D_L(S,R,\infty) = \begin{cases} \dfrac{S-e^{-\lambda_a^*}}{2(1-S)} + R\dfrac{2S-(1+S)\lambda_a^* e^{-\lambda_a^*}}{S(1-\lambda_a^* e^{-\lambda_a^*})} + 1.5, & S < \dfrac{e}{2e-1}, \\ \dfrac{S-e^{-1}}{2(1-S)} + R\dfrac{2S-(1+S)e^{-1}}{S(1-e^{-1})} + 1.5, & S \geqslant \dfrac{e}{2e-1}, \end{cases} \tag{10}$$

which can be independently derived by assuming a Poisson arrival process.

The control channel may be regarded as a pure overhead because it is not used for transmitting data packets. For protocols with control channels consuming a fixed ratio w of the total bandwidth, the effective throughput S becomes

$$S|_{\text{with overhead}} = (1 - w)S|_{\text{without overhead}}.$$

Figure 2 shows the average delay of the UCA protocol [3] with contention-free reservation, the average delay of the C-MA (Controlled Multiaccess) protocol [5] (20 minislots per slot and a maximum of 10 reservations in the reservation queue) and the delay lower bound. Poisson arrival process and zero control channel overhead are assumed in all three cases. We see that both UCA and C-MA protocols have very good delay performance because at most 5% delay reduction can be hoped for. As both UCA and C-MA are not the minimum delay protocol, the difference between the lower bound and the delay of the unknown minimum delay protocol is less than 5% for $R = 100$. Figure 2 also shows that for R large, the M/D/1 bound (the perfect scheduling with fixed size packets and Poisson sources [4]) is too loose to be of any use.

5. Delay lower bound for protocols with contention-based reservation

In section 4, we derived the delay lower bound assuming a contention-free control channel. Here, we relax this assumption by choosing the control channel to be of the slotted Aloha type. Let the control channel be divided into minislots and let there be M minislots to a slot. Let there be an infinite population and the arrival of input packets be a Poisson process. As before, we first design an idealized protocol and derive its average delay. This delay is therefore a lower bound for all hybrid protocols in ξ with an infinite population and contention-based reservations.

For the idealized protocol under contention-based reservation, we made three more assumptions in addition to assumptions 2–4 in section 3. First, we assume that all packets which are successfully transmitted in Aloha slots did not make any spare reservations. This "noncausal" assumption guarantees that there is no spare reservation from successful packets to interfere with the other reservations and hence a smaller delay will result. Second, we assume that all collided packets have made spare reservations because doing so will provide an extra chance of obtaining a Reserved slot for retransmission. When a reservation collides with the other reservations, the stations concerned will reattempt the channel after a random delay. Third, we assume that the combined arrival of normal and spare reservations to the control channel is given by a Poisson process. This is an idealized assumption because packets collided on the Aloha slots will tend to have their spare reservations aggregated together on the control channel. Assuming these reservations to be uncorrelated and modeling them as Poisson arrivals will give an underestimated delay. But for obtaining a delay lower bound this is acceptable.

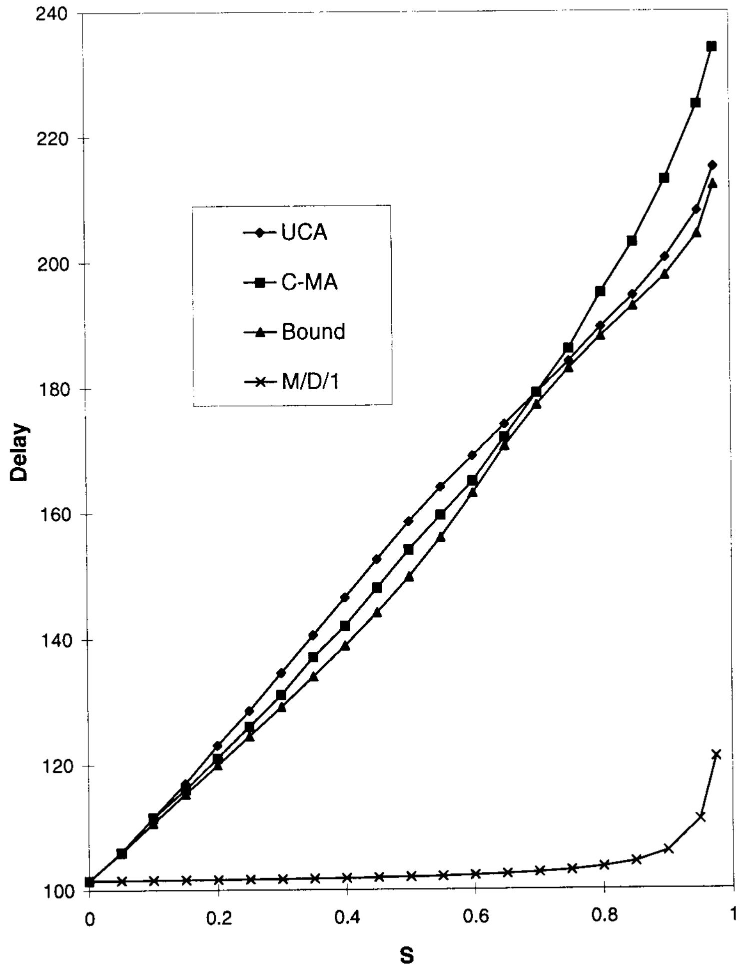

Figure 2. Delay and delay bounds for $R = 100$.

Let the combined arrival rate of normal and spare reservations to the control channel be λ_m per *minislot*, where

$$\lambda_m = \frac{\lambda_r + (\lambda_a - \lambda_a e^{-\lambda_a})(1 - x)}{M}. \tag{11}$$

Here as before x is the average number of successful reservations per slot and is given by

$$x = M[\text{av. no. of successful reservations in a minislot}] = M\lambda_m e^{-\lambda_m}. \tag{12}$$

As a check, by setting $M \to \infty$ (12) degenerates to the case of no reservation contention. Substituting x from (12) into (2) and letting $N \to \infty$, we obtain the throughput S of the idealized protocol as

$$S = M\lambda_{\rm m}{\rm e}^{-\lambda_{\rm m}} + \left(1 - M\lambda_{\rm m}{\rm e}^{-\lambda_{\rm m}}\right)\lambda_{\rm a}{\rm e}^{-\lambda_{\rm a}}. \tag{13}$$

Similar to section 4, for system stability we only consider the case where $\lambda_{\rm a} \leqslant 1$, $\lambda_{\rm m} < 1$ and $S < 1$ in the following derivation. From (2), $S < 1$ implies $x < 1$ and hence $M\lambda_{\rm m}{\rm e}^{-\lambda_{\rm m}} < 1$ from (12). By differentiating (13) with respect to $\lambda_{\rm m}$, we obtain:

Property 3. For a given value of S, $\lambda_{\rm m}$ and $\lambda_{\rm a}$ are inversely related. (See appendix A.)

Substituting (12) into (11) and solving for $\lambda_{\rm r}$, we have

$$\lambda_{\rm r} = M\lambda_{\rm m} - \left(1 - M\lambda_{\rm m}{\rm e}^{-\lambda_{\rm m}}\right)\left(\lambda_{\rm a} - \lambda_{\rm a}{\rm e}^{-\lambda_{\rm a}}\right). \tag{14}$$

By differentiating $\lambda_{\rm r}$ with respect to $\lambda_{\rm m}$ and $\lambda_{\rm a}$, respectively, we obtain:

Property 4. $\lambda_{\rm r}$ is a monotonically increasing function of $\lambda_{\rm m}$ for a fixed $\lambda_{\rm a}$. (See appendix A.)

Property 5. $\lambda_{\rm r}$ is a monotonically decreasing function of $\lambda_{\rm a}$ for a fixed $\lambda_{\rm m}$. (See appendix A.)

Property 3 states that for a fixed S, $\lambda_{\rm a}$ will decrease when we increase $\lambda_{\rm m}$. On the other hand, property 4 states that increasing $\lambda_{\rm m}$ causes a corresponding increase of $\lambda_{\rm r}$. Also, the decrease of $\lambda_{\rm a}$ causes an increase of $\lambda_{\rm r}$ according to property 5. Therefore, we conclude:

Property 6. $\lambda_{\rm r}$ is a monotonically increasing function of $\lambda_{\rm m}$ for a given S.

The average packet delay consists of the sum of seven terms denoted as D_1 to D_7. D_1, D_2, D_4 and D_5 are the same as that in section 3. D_3 is the mean waiting time in the satellite reservation queue and is given by the waiting time in a discrete-time queueing system with the distribution of bulk size U given by

$$\Pr[U = k] = \binom{M}{k}\left(\frac{x}{M}\right)^k\left(\frac{M-x}{M}\right)^{M-k}.$$

This queueing system is exactly the same as that analyzed in section 3. Therefore, we have

$$D_3 = \frac{x(1 - M^{-1})}{2(1 - x)}.$$

The additional propagation delay due to retransmissions D_6 is given by

$$D_6 = R\left[\frac{\lambda_{\rm r} + \lambda_{\rm a}(1 - x)}{S} - 1\right], \tag{15}$$

where $[\cdot]$ is the expected number of retransmissions. Eliminating λ_r and λ_a with the use of (2) and (11), we have

$$D_6 = RS^{-1}(M\lambda_m - x).$$

Neglecting D_5 and the randomization delay for retransmission D_7 for the delay lower bound of the idealized protocol, we obtain

$$\begin{aligned} D_L &= 0.5 + \frac{x(1-M^{-1})}{2(1-x)} + \frac{xR}{S} + 1 + R + \frac{M\lambda_m - x}{S}R \\ &= \frac{M-1}{2\lambda_m^{-1}e^{\lambda_m} - 2M} + R\left(\frac{M\lambda_m}{S} + 1\right) + 1.5, \end{aligned} \tag{16}$$

where λ_m is to be chosen for minimum D_L. As a check, by setting $M \to \infty$ (16) degenerates to the contention-free case. To minimize (16) for a given S, λ_m should be as small as possible. To minimize λ_m, property 3 states that λ_a should be as close to one as possible. Following the approach in section 3 and using (13) and (14), we obtain

$$S_c \equiv S|_{\lambda_a=1,\lambda_r=0}. \tag{17}$$

For $S \geqslant S_c$, we choose $\lambda_a = 1$ and solve for λ_m from (13). For $S < S_c$, we choose $\lambda_r = 0$ and solve for λ_a and λ_m simultaneously from (13) and (14). This choice of $\lambda_r = 0$ results in minimum delay because any other choice of λ_r will cause an increase of λ_m by property 6. Substituting the computed λ_m into (16), a delay lower bound for the idealized protocol can be explicitly evaluated. This lower bound is also a bound for all protocols with contention-based reservation operating in environment ξ as defined in section 1.

Figure 3 shows the average delays of the UCA and C-MA protocols with $R = 100$. Here, with only 3 minislots per slot, the contention-based reservation bound is much tighter than that for the contention-free reservation.

Figure 4 shows the delay lower bounds for $M = 3$, 5, 10 and ∞ with $R = 100$. We see that the bounds are very close for $S \leqslant 0.5$. We also notice that for $M \geqslant 5$, the bound for contention-free reservation (i.e., $M = \infty$) is a good approximation to that for contention-based reservation.

6. Conclusions

Two delay lower bounds are derived for packet satellite protocols under a set of operating conditions. They are shown to be very simple and very tight. They can be used for assessing the possible delay improvement of existing protocols and for deciding whether a particular delay requirement can ever be satisfied.

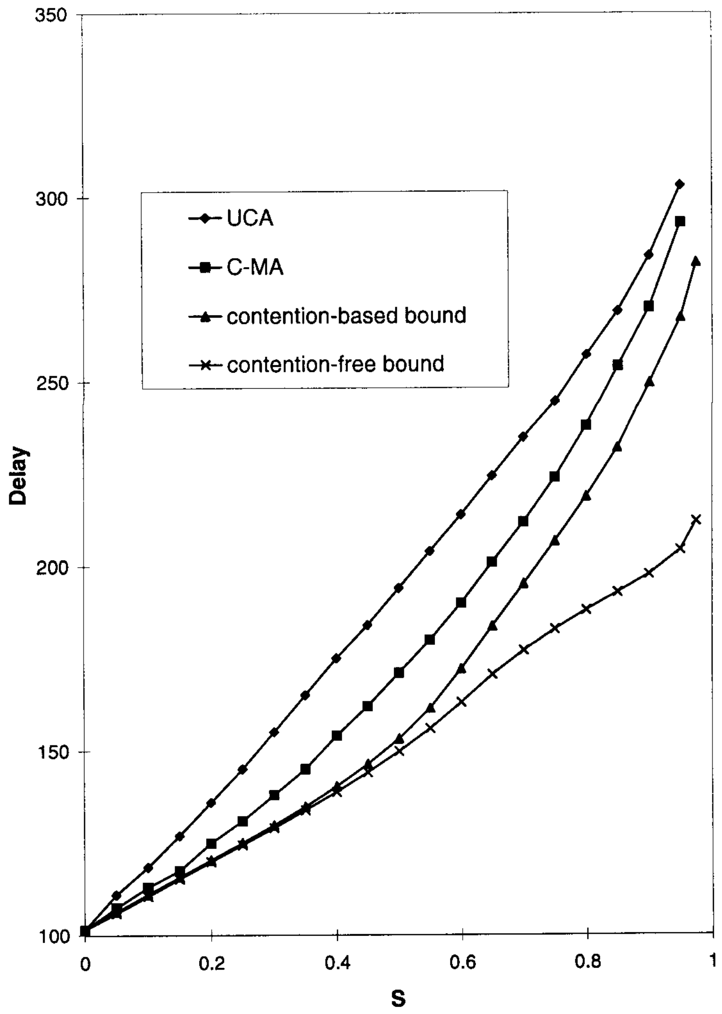

Figure 3. Comparison of delay bounds for contention-based and contention-free control channels.

Appendix A

Proof of property 1. First, we prove that S is a monotonically increasing function of λ_a. It can be proved by showing that $dS/d\lambda_a > 0$. Rearranging equation (3), we have

$$S = \lambda_r + \frac{1 - \lambda_r}{1/\lambda_a + 1 - (1 - \lambda_a/N)^{N-1}}. \tag{A.1}$$

Differentiating it with respect to λ_a, we have

$$\frac{dS}{d\lambda_a} = \frac{(1 - \lambda_r)[1 - (1 - 1/N)\lambda_a^2(1 - \lambda_a/N)^{N-2}]}{[1 + \lambda_a - \lambda_a(1 - \lambda_a/N)^{N-1}]^2} > 0. \tag{A.2}$$

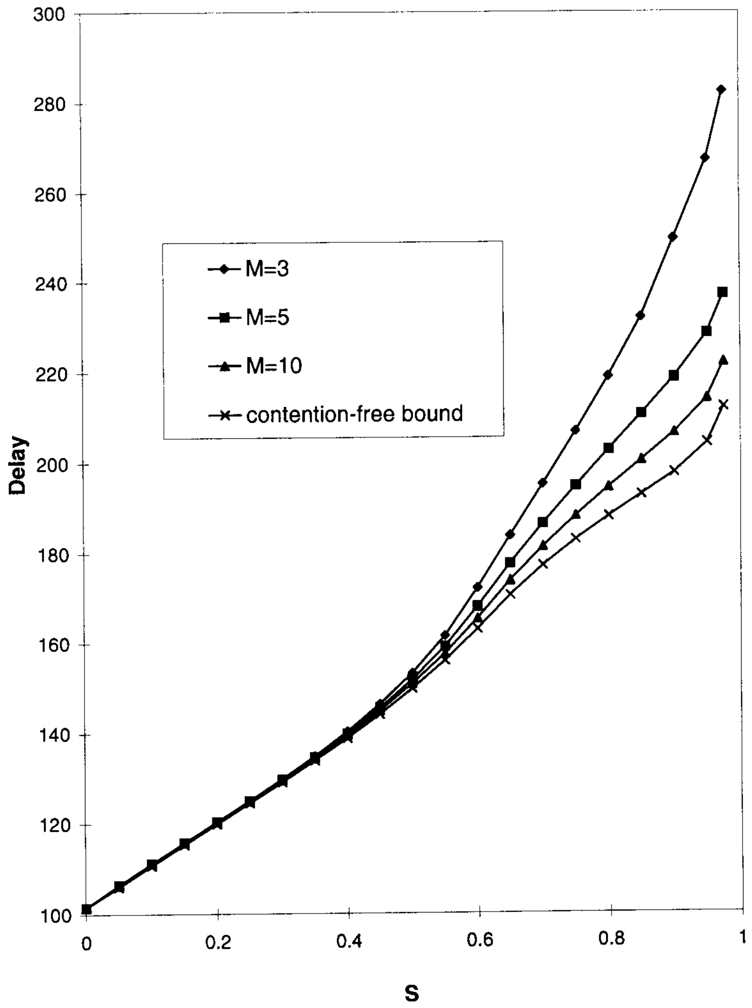

Figure 4. Delay bounds for various M values.

Next, we prove that S is a monotonically increasing function of λ_r by showing that $dS/d\lambda_r > 0$. Differentiating equation (3) with respect to λ_r, we have

$$\frac{dS}{d\lambda_r} = \frac{1 - \lambda_a(1 - \lambda_a/N)^{N-1}}{1 + \lambda_a - \lambda_a(1 - \lambda_a/N)^{N-1}} > 0. \tag{A.3}$$

□

Proof of property 2. We prove this property by showing $d\lambda_r/d\lambda_a < 0$. Rearranging equation (3), we have

$$\lambda_r = S + \frac{S - 1}{1/\lambda_a - (1 - \lambda_a/N)^{N-1}}. \tag{A.4}$$

Differentiating it with respect λ_a, we have

$$\frac{d\lambda_r}{d\lambda_a} = \frac{-(1-S)[1-\lambda_a^2(1-1/N)(1-\lambda_a/N)^{N-2}]}{[1-\lambda_a(1-\lambda_a/N)^{N-1}]^2} < 0. \tag{A.5}$$

□

Proof of property 3. We prove this property by showing $d\lambda_a d\lambda_m < 0$. Rearranging equation (13), we have

$$\lambda_a e^{-\lambda_a} = \frac{S-1}{1-M\lambda_m e^{-\lambda_m}} + 1. \tag{A.6}$$

Differentiating with respect to λ_m and rearranging the result, we obtain

$$\frac{d\lambda_a}{d\lambda_m} = \frac{-(1-S)Me^{-\lambda_m}(1-\lambda_m)}{e^{-\lambda_a}(1-\lambda_a)(1-M\lambda_m e^{-\lambda_m})^2} < 0. \tag{A.7}$$

□

Proof of property 4. We prove this property by showing $d\lambda_r/d\lambda_m > 0$. Differentiating equation (14) with respect to λ_m, we obtain

$$\frac{d\lambda_r}{d\lambda_m} = M + M\lambda_a e^{-\lambda_m}(1-\lambda_m)\left(1-e^{-\lambda_a}\right) > 0. \tag{A.8}$$

□

Proof of property 5. We prove this property by showing $d\lambda_r/d\lambda_a < 0$. Differentiating equation (14) with respect to λ_a, we obtain

$$\frac{d\lambda_r}{d\lambda_a} = -\left(1-M\lambda_m e^{-\lambda_m}\right)\left(1-e^{-\lambda_a}+\lambda_a e^{-\lambda_a}\right) < 0. \tag{A.9}$$

□

Appendix B

Consider a discrete-time independent bulk arrival queueing system with service time equal to one time unit. This system can be analyzed following the method in [2] for a similar continuous-time system. The distribution of the bulk size is given by

$$g_i = \binom{N}{i}\left(\frac{x}{N}\right)^i\left(\frac{N-x}{N}\right)^{N-i}, \quad 0 \leqslant i \leqslant N, \tag{B.1}$$

with the understanding that $g_i = 0$ elsewhere. The state transition probability is p_{ij}, given by

$$p_{ij} = \begin{cases} g_j, & i=0,\ j \geqslant 0, \\ g_{j-i+1}, & i>0,\ j \geqslant i-1, \\ 0, & \text{otherwise.} \end{cases} \tag{B.2}$$

Let $\{P_k\}$ be the set of equilibrium state probabilities. Then, we can write down the balance equations as follows:

$$P_k = g_0 P_{k+1} + (g_k - g_{k+1})P_0 + \sum_{i=0}^{k} P_i g_{k-i+1}, \quad k \geqslant 1, \tag{B.3}$$

$$P_0 = g_0(P_0 + P_1). \tag{B.4}$$

We solve these equation using the method of z-transforms and obtain

$$\begin{aligned}\sum_{k=1}^{\infty} P_k z^k = \frac{g_0}{z}\sum_{k=1}^{\infty} P_{k+1} z^{k+1} + P_0 \sum_{k=1}^{\infty} g_k z^k - \frac{P_0}{z}\sum_{k=1}^{\infty} g_{k+1} z^{k+1} \\ + \sum_{k=1}^{\infty}\sum_{i=0}^{k} P_i g_{k-i+1} z^k. \end{aligned}\tag{B.5}$$

Interchange the order of summation in the last term and rearranging, we obtain

$$\begin{aligned}\sum_{k=1}^{\infty}\sum_{i=0}^{k} P_i g_{k-i+1} z^k &= \sum_{i=0}^{\infty} P_i z^i \sum_{k=i}^{\infty} g_{k-i+1} z^{k-i} - g_1 P_0 \\ &= \frac{1}{z}\sum_{i=0}^{\infty} P_i z^i \sum_{j=0}^{\infty} g_{j+1} z^{j+1} - g_1 P_0. \end{aligned}\tag{B.6}$$

Substituting into (B.5), we have

$$\begin{aligned}P(z) - P_0 = \frac{g_0}{z}\left[P(z) - P_0 - P_1 z\right] + P_0\left[G(z) - g_0\right] - \frac{P_0}{z}\left[G(z) - g_0 - g_1 z\right] \\ + \frac{1}{z}P(z)\left[G(z) - g_0\right] - g_1 P_0, \end{aligned}\tag{B.7}$$

where $P(z)$ and $G(z)$ are the z-transforms of P_i and g_i, respectively. Rearranging the equation, we have

$$P(z) = \frac{P_0 G(z)(1-z)}{G(z) - z}. \tag{B.8}$$

Using $P(1) = 1$ and L'Hospital's rule, we obtain $P_0 = 1 - G'(z)$. In order to calculate the average number of customers in the system $\bar{N}$ which is equal to $P'(1)$, applying L'Hospital's rule twice to (B.8) we obtain

$$\bar{N} = G'(1)\frac{G''(1)}{2[1 - G'(1)]} = x + \frac{x^2(1 - 1/N)}{2(1-x)}. \tag{B.9}$$

Applying Little's result and using the relationship with the average system time T and the average waiting time spent in the queue W (i.e., $T = 1 + W$), we finally have

$$W = \frac{x(1 - 1/N)}{2(1-x)}. \tag{B.10}$$

References

[1] D.M. Chite and J.S. McCoskey, VSAT networks: architectures, protocols, and management, IEEE Communications Magazine 26(7) (July 1988).
[2] L. Kleinrock, *Queueing Systems, Vol. 1: Theory* (Wiley, New York 1975), p. 134.
[3] H.W. Lee and J.W. Mark, Combined random/reservation access for packet switched transmission over a satellite with on-board processing: Part I – Global beam satellite, IEEE Trans. Commun. 31 (October 1983) 1161–1171.
[4] F.A. Tobagi, Multiaccess protocols in packet communication systems, IEEE Trans. Commun. 28(4) (April 1980).
[5] E.W.M. Wong and T.S. Yum, The controlled multiaccess protocol for packet satellite communication, IEEE Trans. Commun. 39(7) (July 1991).