The Scheduled-Retransmission Multiaccess (SRMA) Protocol for Packet Satellite Communication

TAK-SHING YUM, SENIOR MEMBER, IEEE, AND ERIC W. M. WONG, STUDENT MEMBER, IEEE

Abstract—An improvement of the announced retransmission random access (ARRA) protocol for packet satellite communication called the scheduled retransmission multiaccess (SRMA) protocol is introduced. Besides avoiding the collisions between new and retransmitted packets, the improved protocol can also eliminate reservation conflicts from different slots. Explicit acknowledgment is used so that traffic to other zones served by the same satellite can also be accommodated. Assuming three percent of the system capacity is used for retransmission reservation, the fixed-frame and dynamic-frame SRMA's can achieve effective maximum throughputs of 0.65 and 0.89, respectively. Both protocols can provide average delays considerably lower than that of Aloha even when the throughput is as low as 0.2.

I. INTRODUCTION

Since the introduction of the Aloha system [1], research in multiaccess protocols has flourished. For networks with very low propagation delays, the series of carrier-sensing protocols [2] can give maximum throughput close to unity. However, for a satellite channel with a large propagation delay, efficient protocols are more difficult to design. One class of techniques makes use of the reservation principle. These techniques can attain a channel capacity close to unity, but they also have in common the delay overhead of one round-trip propagation time for exchanging reservation information. Some protocols in this class [3], [4] have contention-based reservation, so that not all reservations are successful in one shot.

For networks with bursty traffic, random access techniques can offer more satisfactory delay performance. Wieselthier and Ephremides [5] designed the interleaved frame flush-out (IFFO) protocol. IFFO is basically a combined reservation and random-access protocol. To ensure that all reservations are successful, one reservation minislot is assigned to each station in the system per frame (this therefore limits the number of stations in the system). Frame length is variable but must be less than or equal to the round-trip propagation delay or 0.24 s. A frame is composed of a status slot (for transmitting reservation minipackets in time-division multiple access (TDMA) fashion), a reserved subframe, and a contention subframe. All packets make a reservation in the TDMA slots near the beginning of a frame. In addition, some packets are selected to be transmitted immediately in the contention subframe. If the transmission is successful, the reservation is canceled. Suda et al. [6] presented a dual-mode protocol for packet and circuit switched traffic using slotted-Aloha and packet reservation protocols, respectively. The mean transmission delay and optimum frame length were found. Raychaudhuri [7] proposed the announced retransmission random access (ARRA) protocol which makes use of a low-rate subchannel to announce the retransmission time so that conflicts between new and retransmitted packets are prevented. It was shown that the extended ARRA can achieve a capacity close to 0.6.

In this correspondence we introduce the scheduled retransmission multiaccess (SRMA) protocols that is somewhat similar to ARRA. In addition to avoiding the collision between new and retransmitted packets, SRMA also eliminates the reservation collisions from different slots in the frame. In contrast to ARRA, the common minislot pool at the beginning of each frame is not needed for SRMA. Two versions of SRMA are described and analyzed. At three-percent retransmission reservation overhead, the fixed-frame version (SRMA/FF) can give a maximum throughput of 0.65, and the dynamic-frame version (SRMA/DF) can attain a throughput of 0.89. Moreover, the average delay is considerably lower than that of the slotted Aloha even when the throughput is as low as 0.2. Generally speaking, SRMA behaves like Aloha under low traffic and like a reservation protocol under high traffic.

II. DESCRIPTION OF THE SRMA PROTOCOLS

A. The Fixed-Frame SRMA

Let the packet satellite channel be divided into frames of K slots each. Let each frame be divided into an Aloha subframe and a reserved subframe (Fig. 1). Each slot in the Aloha subframe has a header and a body. The header consists of M minislots, each long enough so that the three states—"idle," "success," and "collision,"—can be distinguished. The body can accommodate one packet. For each new or reattempting packet transmitted in a slot, one of the M minislots in the corresponding header is marked (by a series of bits) at random for retransmission scheduling purposes in case of collision. Note that guard times are needed between frames, slots, and minislots to assure synchronization. A discussion on this appeared in [5]. Alternatively, to alleviate the stringent synchronization requirements, a separate

REFERENCES

Manuscript received June 9, 1987; revised January 1989.

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

IEEE Log Number 8913674.
control channel can be used to accommodate the reservation information. Fig. 1(b) shows such an arrangement. Since the control channel has a very low bit rate, synchronization can easily be maintained [8].

New packets arriving in the Aloha subframes are transmitted immediately. Those arriving in the reserved subframes, however, are scheduled to be transmitted in one of the U upcoming Aloha slots at random. Let the round-trip propagation delay be R frames. Then R frames after the initial transmission, a station will learn on the feedback channel (a low-rate announcement subchannel from satellite to stations) whether the transmission is successful or not. An unsuccessful transmission will be assigned a dedicated slot in the next frame for retransmission if the retransmission reservation is successful. The collection of all these dedicated, or reserved, slots in a frame constitutes the reserved subframe; and packets retransmitted in the reserved subframes are always successful. If both the transmission and the retransmission reservation are unsuccessful, the packets will reattempt transmission at one of the V upcoming Aloha slots chosen at random. Note that V should be designed to be less than U to minimize the delay variance. Also, to maintain unconditional stability of the channel, V can be changed adaptively according to the channel traffic. An example of such a technique is binary exponential backoff.

Let \((x, y)\) be a status vector assigned to each packet. Here \(x \in \{1, 2, \ldots, K\}\) is the slot position in which the packet is transmitted and \(y \in \{1, 2, \ldots, M\}\) is the minislot position being set to one. After a frame of packets is received by the satellite, the satellite processor would collect the set of status vectors and process them as follows:

1) All collided packets in a particular slot with nonunique status vectors are to be scheduled for later transmission. (The \((x, y)\) values are used for scheduling the retransmission orders. Two packets with the same \((x, y)\) value cannot be differentiated and therefore cannot be scheduled unambiguously.)

2) The remaining collided packets are arranged into subsets \(X_1, X_2, \ldots, X_i\) where \(X_i = \{(x, y) | x = i\}\).

3) Each \(X_i\) is further arranged in ascending order of \(y\) values.

4) The resulting set of ordered status vectors is truncated to a size of \(K\) since in the retransmission frame at most \(K\) reservations can be made.

This set of ordered status vectors (one per frame) is broadcast to all stations together with acknowledgments of successfully received packets. A station which does not receive an acknowledgment will search for its status vector. If its status vector is at position \(b\), the station will retransmit its packet at slot \(b\) of the next frame. If, on the other hand, the status vector is not found, the retransmission reservation must be unsuccessful. The station then contends for the channel after a random delay. Fig. 2 summarizes the SRMA/FF protocol in a flowchart.

![Fig. 2. SRMA/FF protocol.](image)

To illustrate the protocol, let there be ten slots in each frame and eight minislots in the header of each slot (i.e., \(K = 10, M = 8\)). A typical frame is shown in Fig. 3a. Here, the Aloha subframe starts at slot 5, and two packets \(4\) and \(6\) collide there. The corresponding minislot positions chosen are 1 and 2, respectively. In slot 6, three packets collide. Slot 7 is empty. Two packets with status vectors \((8, 2)\) and \((8, 6)\) collide in slot 8, etc. By step 1), packets \(D, E, I,\) and \(H\) are to be rescheduled for later transmission. By steps 2) and 3) the packet status vectors are arranged as \((5, 1), (5, 2), (6, 7), (8, 2), (8, 6),\) and \((10, 3)\), which is also the desired order these packets will be transmitted in the reserved
subframe R frames later. Fig. 3(b) shows the corresponding retransmission frame.

For a satellite with on-board processing, the up-link traffic may include packets destined for other stations in different "zones" served by different transponders. If that is the case, explicit acknowledgement of a successfully received packet by the satellite is necessary since that packet may not be destined to the same zone from which it originates.

B. The Dynamic Frame SRMA

In the dynamic frame protocol, the framing pattern is different. Here an Aloha frame (not a subframe) has a fixed length of F slots, and a reserved frame has a length equal to the number of successfully reserved packets in its corresponding Aloha frame. Thus in contrast to the FF protocol, there is no truncation on the number of reservations here.

Using the DF protocol, a station having a packet to transmit would 1) transmit immediately if the channel is in the Aloha mode; and 2) wait for the beginning of the next Aloha frame, and transmit the packet in one of the F slots at random if the channel is in the reserved mode. As in the FF case, one of the minislots positions is set to one for scheduling purpose. The feedback information is similar to the FF case except that now the status vectors of all successfully scheduled packets are sent back. In other words, because we now have a variable-length reserved frame, there is no reservation overflow.

All Aloha frames are labeled with sequence numbers. The numbers range from 1 to N, where N is the maximum number of Aloha frames that can be accommodated in one round-trip propagation delay. The corresponding reserved frames are labeled with the same sequence number. Thus if a packet transmitted in Aloha frame i collides, it will be retransmitted in reserved frame i provided its reservation minislot is not garbled. After receiving the status vector of frame i, all stations will initiate reserved frame i right after the current Aloha frame if reserved frame i-1 is already sent. Otherwise, reserved frame i will start right after reserved frame i-1. To summarize, all reserved frames are sent in order of their sequence number. If no reserved frame is available, Aloha frames are sent instead. All reattempting packets are to be transmitted in the next Aloha frame.

The DF protocol requires all stations to keep track of the durations and locations of all untransmitted reserved frames so that they can know when to transmit their new, reserved, and reattempting packets. In contrast, the FF protocol requires the stations to keep only the length of the current reserved subframe.

Fig. 4 shows the transition between the Aloha and the reserved modes in a station. Fig. 5 summarizes the SRMA/DF protocol in a flowchart.

III. THROUGHPUT ANALYSIS

A. SRMA/FF Protocol

Packets that collided in frame I will be retransmitted after R frames if they are successfully scheduled. Consider a subchannel consisting of frames I, I + (R + 1), I + 2(R + 1), I + 3(R + 1), etc., and let us rename these frames as i, i + 1, i + 2, i + 3, etc. We shall evaluate the throughput of the subchannel and argue that, since all subchannels have the same throughput, the channel throughput is simply R + 1 times the subchannel throughput.

Let the combined arrivals of new and reattempting packets to the Aloha subframes be a Poisson process with rate g packet/slot.
Let $L_i$ be the number of “collided packets with successful reservations” in the $i$th slot of the Aloha subframe, and let $N$ be the total number of packets transmitted in that slot. Then

\[ P[ L_i = r | N = n ] = P[ r \text{ out of } n \text{ packets have their reservations not in conflict with others}. ] \]

For $n = 0$ and 1, it is readily seen that $P[ L_i = 0 | N = 0 ] = P[ L_i = 0 | N = 1 ] = 1$. For $n = 2, 3, 4, \ldots$, we use the result in [9] to obtain

\[ P[ L_i = r | N = n ] = P[ r \text{ cells have exactly one ball given that } n \text{ balls are tossed in } M \text{ cells}] . \]

\[ = \frac{ (M-1) \cdot n! }{ r! (M-n)! } \sum_{k=r}^{M} \binom{M}{k} (-1)^{k} \frac{ (M-k)^n (n-k)! }{ (k-r)! (M-k)!(n-k)! } . \]

Since $N$ has a Poisson distribution, we have

\[ P[ L_i = r | N = n ] = \frac{ g^n e^{-g} }{ n! } \sum_{k=r}^{M} \binom{M}{k} (-1)^{k} \frac{ (M-k)^n (n-k)! }{ (k-r)! (M-k)!(n-k)! } . \]

Let random variables $X_i$ and $Y_i$ denote the length (in slots) of the $i$th reserved and the $i$th Aloha subframes, respectively. Let $W_i$ be the total number of successful reservations for the collided packets in frame $i$. Then

\[ W_i = L_1 + L_2 + \cdots + L_i . \]

Since all $L_i$ in (2) are independent identically distributed random variables, the generating function of $W_i$ is

\[ G_{W_i}(z) = \sum_{k=0}^{\infty} [ G_{L_i}(z) ]^k P[ Y_i = k ] . \]

(3)

where $G_{L_i}(z)$ is the common generating function of the $L_i$.

The protocol specifies that at most $K$ slots can be reserved in a frame; thus the length of the $(i+1)$th reserved subframe must be

\[ X_{i+1} = \min[ W_i, K ] . \]

Its distribution is

\[ P[ X_{i+1} = j ] = \begin{cases} P[ W_i = j ], & j = 0, 1, 2, \ldots, K-1 \\ \sum_{k=j}^{K} P[ W_i = k ], & j = K. \end{cases} \]

(5)

Also, since $X_i + Y_i = K$,

\[ P[ Y_i = k ] = P[ X_i = K-k ], \quad k = 0, 1, 2, \ldots, K . \]

(6)

At steady state, $\{ W_i \}$ and hence $\{ X_i \}$ and $\{ Y_i \}$ shall all have distributions that are independent of $i$. Substituting (5) into (6) and then into (3), we have

\[ \sum_{j=0}^{MK} P[ W = j ] z^j = \sum_{n=K}^{MK} P[ W = n ] + \sum_{k=1}^{K} \sum_{m=0}^{M} P[ L = m ] z^m \]

\[ \cdot \sum_{k=K}^{MK} P[ W = K-k ] z^{K-k} . \]

Equating the coefficients of $z^j$, we arrive at a system of simultaneous linear equations whereby $\{ P[ W = j ] \}$ and hence $\{ P[ X = j ] \}$ can be solved.

Let $E[ X ]$ be the mean of $X$. Then $E[ Y ] = K - E[ X ]$. The probability of exactly one packet arrival in an Aloha slot is, $g e^{-g}$, hence the average number of successful packet transmissions in an Aloha subframe is $E[ Y ] g e^{-g}$. We now let $h$ be the length of a minislot. Then $h M$ is the control overhead per packet. The packet size excluding overhead is therefore $1 - h M$. The subchannel throughput $S'$ and the subchannel traffic $G'$ are, respectively,

\[ S' = \frac{ E[ X ] + E[ Y ] g e^{-g} }{ (R+1)K } , \]

\[ G' = \frac{ E[ X ] + g E[ Y ] }{ (R+1)K } . \]

The channel throughput $S$ and the channel traffic $G$ are therefore $(R+1)S'$ and $(R+1)G'$, respectively.

B. SRMA/DF protocol

The throughput deriviation of the DF protocol is much simpler than that of the FF protocol. Here $Y_i = F_i$, a constant, and $X_i = W_i, \forall i$. Hence $X_i = L_1 + L_2 + \cdots + L_i$ and $E[ X ] = F^*E[ L ]$. The channel throughput and the channel traffic can be derived in a similar fashion as

\[ S = \frac{ E[ L ] + g e^{-g} }{ E[ L ] + 1 } , \]

\[ G = \frac{ E[ L ] + g }{ E[ L ] + 1 } , \]

which are, as expected, independent of $F$.

IV. DELAY ANALYSIS

The following delay analysis follows the method in [10]. We shall derive the expected delay for SRMA/FF only. The delay expression for the DF protocol is the same as the FF protocol with the replacement of $E[ X ]$, $E[ Y ]$, and $K$ by $F^*E[ L ]$, $F$, and $F^*E[ L ] + 1$, respectively. Let $b$ be the probability of successful transmission or successful retransmission reservation for a packet.

Then

\[ q = \frac{ \text{average number of successful transmissions and reservations in a frame} }{ \frac{ E[ X ] + E[ Y ] g e^{-g} }{ E[ Y ] g } } . \]

Then the average number of retransmission attempts before successful transmission or successful retransmission reservation is

\[ \sum_{n=1}^{\infty} n q(1-q)^n = 1/q - 1 . \]

Next, let

\[ f = \Pr[\text{a successful ts'ed is through retransmission reservation}] = \frac{ E[ X ] }{ (E[ X ] + E[ Y ] g e^{-g}) } . \]

and

\[ D_1 = E[ X ] / 2 + \{ U/2E[ Y ] \} E[ X ] = E[ X ] (1 + U/2E[ Y ] ) / 2, \]

the expected initial delay for a packet,

\[ D_2 = E[ Y ] / 2 + RK + E[ X ] + \{ U(2E[ Y ]) \} K, \]

the expected addition delay for each unsuccessful retransmission reservation,

\[ D_3 = E[ Y ] / 2 + RK + E[ X ] / 2 = (R + 0.5)K, \]

the expected duration between the successful reservation of a packet and its actual retransmission.
The expected delay $D$ in slots therefore is

$$D = (RK + 1) + D_i + (1/q - 1) D_j + fD,$$

where $RK + 1$ accounts for the packet propagation and transmission delays.

V. RESULTS AND DISCUSSIONS

Fig. 6 compares the throughput of slotted Aloha, FF, and DF protocols. For FF we choose the frame size $K$ equal to 20 slots, and for both FF and DF, the size of the control minislot $h = 0.001$ and $M = 10$. We see that all three protocols are unstable as the channel traffic $G$ increases. While the throughput $S$ of slotted Aloha peaks at $G = 1$, $S$ of FF and DF have peaks at $G = 1.4$ and $3.0$, respectively. The maximum throughputs are 0.65 and 0.78. Notice that the $G$ given here for DF is the average channel traffic over the entire frame. Since there is contention at the Aloha frames only, the actual channel traffic on the Aloha frames is 10.4 packets/slot at maximum throughput. Thus, DF is much more stable than slotted Aloha.

Fig. 7 compares the average delays (including transmission, contention and propagation delays) of four protocols. A round-trip propagation delay of 250 slots is assumed. The slotted-Aloha delay performance is well-known. For both FF and DF protocols, $h = 0.001$, $M = 10$, $U = 12$, and $V = 10$ are assumed. The extended ARRA delay is obtained from [7] with the same number of minislots per slot and slots per frame. The figure shows that the delays of FF and DF are comparable up to about $S = 0.6$ and are both significantly smaller than that of ARRA. A different traffic model was used for the analysis of JFFO [5] and so direct comparison is not possible.

Fig. 8 shows the maximum throughput $S_{\text{max}}$ achievable by FF as a function of $K$ for various $M$ with $h = 0.001$. We see that $S_{\text{max}}$ is insensitive to $K$ for $K \geq 5$. In fact, the choice of $K$ could be quite arbitrary. For example, for round-trip propagation delay equal to 250 slots and $M = 10$, the average delay at maximum throughput for $K = 10$ and $K = 30$ are both equal to 271 slots. For $M = 5$, the difference in maximum throughput between $K = 5$ and $K = 30$ is only 3.6 percent (0.614 and 0.637). The average delays are both equal to 350 slots.

Fig. 9. Maximum throughput of FF as function of $M$, $K = 10$. 
On the other hand, the effect of reservation overhead on maximum throughput is more profound. Fig. 9 shows $S_{\text{max}}$ versus $M$ with $h$ as a parameter. We see that initially $S_{\text{max}}$ increases with $M$ (which is expected). However, as $M$ increases beyond some critical value (depending on $h$), the reservation overhead gets so large that $S_{\text{max}}$ starts to drop. Thus, for a given value of $h$, there exists an $M$ that maximizes the throughput. In general, for $h$ not too large, the throughput is maximized for a broad range of $M$. Fig. 10 shows the same for the DF protocol. We see that for $h$ not too large, say equal to 0.1 percent of the packet length, a maximum throughput of 0.89 can be obtained.

**New Lower Bounds for Constant Weight Codes**

CORNELIS L. M. VAN PUL AND TUVI ETZION, MEMBER, IEEE

**Abstract** — Some new lower bounds are given for $A(n,4,w)$, the maximum number of codewords in a binary code of length $n$, minimum distance 4, and constant weight $w$. In a number of cases the results significantly improve on the best bounds previously known.

**I. INTRODUCTION**

In this correspondence we present a method of finding lower bounds for $A(n,4,w)$, the maximum number of codewords in a binary code of length $n$, minimum distance 4, and constant weight $w$. Graham and Sloane [1] gave the first lower bound for $A(n,4,w)$:

$$A(n,4,w) \geq \begin{pmatrix} n \\ w \end{pmatrix}$$

Their proof is based on a mapping $T: F^n_2 \rightarrow Z_n$, where $F^n_2$ denotes the set of $\binom{n}{w}$ binary vectors of length $n$ and weight $w$ and $Z_n = \mathbb{Z}/n\mathbb{Z}$ denotes the residue classes modulo $n$. The mapping is $T(\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_k) = \sum_{i=1}^{k} \mathbf{b}_i (\mod n)$, and for $1 \leq i \leq n$ the code $C_i = T^{-1}(i-1)$ is a constant weight code with distance 4. Clearly,

$$A(n,4,w) \geq \max_{1 \leq i \leq n} |C_i|$$

and (1) is an immediate consequence of (2).

Klove [2] observed that we can take other additive groups instead of $Z_n$ and found larger codes in a few cases. Other bounds for $A(n,4,w)$ were provided by Kibler [3], Brouwer [4], and Delise and Piret [5].

The codes given by (2) and by Klove [2] are not significantly better than the ones obtained from (1). In the present correspondence we show that in many cases we can partition $F^n_2$ into $n$ classes so that each class $i$ is a constant weight code, $C_i$, with distance 4, $\bigcup_{i=1}^{n} C_i = F^n_2$, and $\max_{1 \leq i \leq n} |C_i|$ is fairly large, and that in many cases we can partition $F^n_2$ into $n-1$ classes so that each class $i$ is a constant weight code, $C_i$ with distance 4, $\bigcup_{i=1}^{n-1} C_i = F^n_2$, and $\max_{1 \leq i \leq n-1} |C_i|$ is large. Using these partitions we were in many cases able to improve on the lower bounds obtained from (1) and (2).

In Section II we present the new method of obtaining lower bounds for $A(n,4,w)$, a table of new lower bounds, and constructions for partitioning $F^n_2$. In the Appendix we present partitions of $F^n_2$ which were used to obtain the new lower bounds.

**II. THE NEW LOWER BOUNDS**

For the representation of our results we need some definitions. An $(n,d,w)$ code is a code of length $n$, constant weight $w$, and distance $d$. A set of codes will be called disjoint if the intersection of any two different members of the set is empty. Let $F^*$ denote the set of all binary $n$-tuples. A partition $\Pi(n)$ of $F^*$ is a set of $k$ subsets (called classes), $A_1, A_2, \cdots, A_k$, such that each...