A Go-Back-N Protocol with Multicopy Retransmission for High Speed Satellite Communications

Heliomar M. de Lima
Banco do Brasil S. A.
R. Br. S. Francisco, 177
Rio de Janeiro - RJ - Brasil
e-mail: heliomar@coe.ufrj.br

Otto Carlos M. B. Duarte
UFRJ - COPPE/EE - PEE
CEP 21945-970 - P.O.Box 68504
Rio de Janeiro - RJ - Brasil
e-mail: otto@coe.ufrj.br

ABSTRACT This paper presents and analyses a GB(N) ARQ protocol that requires equal size buffer length in the transmitter and receiver. The recovery strategy uses the idea of multicopy retransmissions. It avoids buffer overflows, reduces the effect of the round trip delay and performs well in high speed satellite environments. The analysis uses a semi-Markov process model and includes a parameter in order to obtain the throughput as a function of the number of erroneous messages that may be processed and reordered during the recovery procedure. Numerical results show that the proposed strategy presents better performance than all the other known GB(N) schemes. Furthermore, the throughput efficiency remains in an usable range even for very high error rate conditions where it outperforms some finite buffer selective repeat schemes.

1. Introduction

An error recovery mechanism is normally employed to guarantee a reliable end to end transfer when the network residual error rate does not satisfy the quality of service required by the application. Error recovery techniques can be classified in three types: Forward Error Correction (FEC), Automatic Repeat Request (ARQ) and hybrid (FEC+ARQ). ARQ schemes are simple to implement, robust, reliable and require a return channel in order to ask the retransmissions. This kind of schemes are the most used in practice and several different strategies have been proposed [1-9] . Today, the bottleneck of high speed communications is the message processing time. In order to face this problem this paper looks for simple protocols, without correcting codes. Therefore, to overcome a noisy channel condition, this work considers that the transmitter should provide repeated transmissions of the same message.

The idea of repeated transmissions was introduced by Sastry. He proposed a GB(N) protocol [1] that continuously retransmits copies of the same message when a message is negatively acknowledged by the receiver. However, in high speed environments this scheme outperforms the conventional GB(N) only under bit error rates in which its throughput efficiency presents low and no practical values. Birrel has proposed another GB(N) strategy [7] in which the sender always transmits m copies of each message. For $m \geq 2$ the throughput is penalized even when the channel presents very good error rate, since its maximum efficiency is $1/m$.

In order to improve the performance of these protocols some authors have increased the receiver capabilities. In this class of protocols the receiver has a buffering memory, can accept and acknowledge the errorless non-sequenced messages received and reorder the retransmitted messages. Morris's strategy [2] has an improved receiver memory capacity and uses the same repeated retransmission scheme of the Sastry one. Lin&Yu proposed the SETRAN protocol [4], other improved receiver capacity GB(N) with special repeated retransmissions. The positively acknowledged messages are not retransmitted: in the correspondent time slots the emitter retransmits copies of the oldest negatively acknowledged message stored in its retransmission buffer.

Miller&Lin [6] introduced the use of repeated retransmissions in the selective repeat schemes. However, the best efficiency results concerning the "multicopy retransmission protocols" were obtained by the multiphase and multicopy selective repeat scheme proposed by Weldon [8], which uses a few additional copies and only in the retransmissions.

This work presents a GB(N) recovery strategy which does not consume too much resources (processing and memory), well suited to a high speed satellite environment where the channel is good most of the time and noisy only occasionally. In addition to using the key idea of multicopy retransmissions suggested in Weldon's scheme, the proposed scheme requires a buffering capability of $(S+1)$-messages length at the receiver to to store the good messages following an erroneously received one. The emitter retransmits m copies ($m \geq 2$) of all negatively acknowledged messages. Furthermore, the throughput analytical expressions include a parameter in order to obtain the performance as a function of the number of erroneous messages that the receiver can process and reorder during a recovery procedure.

First, the analytical model, based on discrete-time semi-Markov process, is presented. The following section describes the proposed recovery strategy and analyses its throughput efficiency. Numerical results are provided and compared with other ARQ schemes. Finally, the conclusions are presented.

2. Analytical Model

The protocol throughput efficiency is defined as the rate between the number of errorless received and accepted messages and the total number of transmitted and retransmitted messages. It will be obtained by assigning states to the receiver according to the recovery process conditions and modeling the receiver behavior as a discrete-time semi-Markov process, which state sequence is $X(0), X(1), X(2), X(3), \ldots, X(n), X(n+1), \ldots$ where $X(n), n = 0, 1, 2, \ldots$, is a random variable denoting the receiver state after the reception of the n^{th} message.

Throughout this paper it is assumed that the continuous transmission of messages. The information message length is constant and equal to L, with K bits of useful information. The time slot required to transmit one information message is used as time unit. The transmission channels are characterized by their bit error rate p (independent bit errors) and their constant round trip delay S. Moreover, it is assumed that all erroneous messages are detected, the control messages (acknowledgements) are error-free (the return channel is noiseless) and the processing time at the transmitter and receiver is zero. Furthermore, the following notation is used: P_{be} is the probability that an information message arrives with at least one bit error and P_{bc} is the probability that an information message is error-free.

3. Protocol Description

The proposed strategy is a GB(N) scheme supposed to be used in a satellite environment characterized by a good bit error rate during most of the time and that occasionally presents a noisy condition. Thus, the scheme is a conventional GB(N) with a modified recovery procedure in order to achieve an improved throughput performance in the high bit error rate region. In addition to the classical (S+1)-messages retransmission buffer, the scheme requires an equal size receiving buffer. This buffer allows the receiver to process and acknowledge the received messages after the detection of an erroneous message and to store the correctly received ones, eliminating the need for successful retransmission of the messages already correctly received and stored. Further, the transmitter uses the basic idea of the Weldon's recovery scheme retransmitting m copies ($m \geq 2$) of all negatively acknowledged messages.

The scheme requires two types of messages. Information messages that are sequentially numbered and control messages. The control messages provide two acknowledgement fields. The first one is an accumulative positive acknowledgement and contains a number $N(R)$ indicating the good reception of all messages numbered until $N(R)-1$. The second one ($N_1(R)$) is an individual negative acknowledgement indicating an erroneous reception of the message numbered $N_1(R)$. Hereafter in this work, when this field is activate the control message will be called REJ $N_1(R)$ (*reject*).

The receiver begins an error recovery procedure when it detects an erroneous message. It sends a retransmission request, reserves a slot in the buffer and continues to receive, process and acknowledge the messages sent in advance. The messages correctly received are stored to posterior delivery while a buffer slot is reserved for each erroneous detected message. Henceforth, the missing messages corresponding to the buffer slots will be called messages B_{e_i} and the corresponding retransmission requests, REJ B_{e_i}, where $0 \leq i \leq S$ (note that in this case $N_1 = B_{e_i}$). During a recovery procedure, the receiver stops the message processing and storing when it detects the message B_{e_v}. In other words, it discards all the messages following the v^{th} detected error and waits for the recovery of message B_{e_0}. During a recovery procedure, the receiver waits for m erroneous copies of each message corresponding to a "hole" in the receiving buffer (messages B_{e_i}) before resending a REJ B_{e_i}.

The transmitter sends sequenced messages and stores them in a retransmission buffer while waiting for their positive acknowledgement. On the receipt of a REJ B_{e_i}, it compares the fields $N(R)$ and B_{e_i}. If $N(R)$ is equal to B_{e_i}, it goes back and retransmits continuously the retransmission buffer, resending m copies of each negatively acknowledged message. Hereafter this retransmission will called B_{e_i} go-back-N retransmission. Otherwise, if $N(R)$ is not equal to B_{e_i}, this information indicates the messages that must be repeated m times without causing a "go back procedure".

4. Performance Analysis

The proposed strategy defines two states for the receiver: normal and exceptional states. In the normal state, the receiver accepts the sequenced correctly received messages and delivers them to the user. There is no outstanding error recovery in this state. The receiver turns to the exceptional state when an erroneous message is received. It remains in this state while there is an erroneous message to be recovered.

When the receiver goes to the exceptional state, it begins the recovery procedure that it will be classified in cycles. The exceptional state has one *transmission cycle* and one or more *retransmission cycles*. The transmission cycle corresponds to the processing and storing of messages sent in advance. In this cycle the receiver may store up to S+1 messages (including the slots reserved for messages B_{e_i}). Additionally, the number of messages B_{e_i} is limited to $v+1$ ($B_{e_i}, 0 \leq i \leq v$). The retransmission cycles correspond to the number of go-back-N retransmission request sendings (REJ $B_{e_i}, 0 \leq i < v$ with $B_{e_i} = N(R)$ sendings). The final retransmission cycle ends immediately before the reception of the message B_{e_v} retransmission and the non-final ones ends after the retransmission of the message numbered $B_{e_0}+S$.

The analysis considers a "inferior scheme" of the protocol described in the preceding section. In the proposed scheme when the message B_{e_0} is recovered and still remains one or more unrecovered messages, the receiver sends a REJ B_{e_i} requesting a B_{e_i} go-back-N retransmission and delivers to the user the sequenced messages stored in the receiving buffer. In this analysis, if the message B_{e_0} is recovered in a retransmission cycle the corresponding released buffer space will not be taken into account. The receiver, after processing and storing the messages numbered $B_{e_0}, B_{e_0}+1, \ldots, min(B_{e_v}, B_{e_0}+S)$ (during the transmission cycle), waits for the retransmissions, asks for new retransmissions and goes to the normal state when all messages $B_{e_i}, i < v$ are recovered. Figure 1 illustrates the analysed scheme.

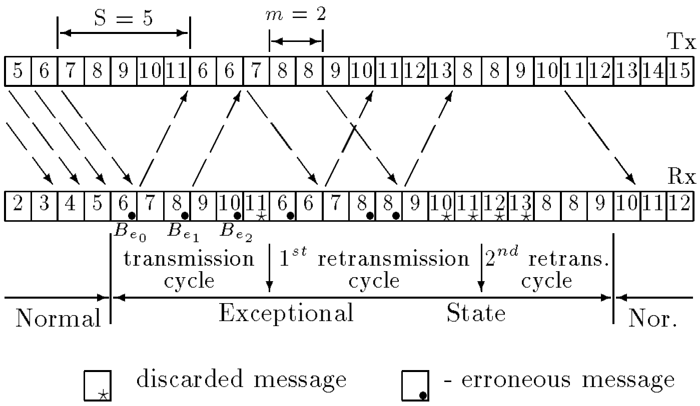

Figure 1 *Analysed Strategy ($m = 2$ and $v = 2$)*

In the following, the throughput efficiency of the this scheme will be derived. As the receiver passes from the normal to the exceptional state at the beginning of the erroneous message reception, only error-free messages are received in the normal state. Then, the throughput efficiency of the protocol in this state is K/L.

The protocol efficiency is the weighted mean value of the efficiency in each state. Then:

$$\eta = \frac{K}{L}\left(P(N) + P(E)\frac{A+1}{E[t_E]}\right) \tag{1}$$

where $P(N)$ and $P(E)$ are the probabilities of the receiver being at the normal and at the exceptional states, respectively. The expressions $(A+1)/E[t_E]$, A and $E[t_E]$ are the exceptional state efficiency, the average number of messages "sent in advance" and delivered to the user during the exceptional state and the mean waiting time at the exceptional state, respectively.

The mean number of messages (A) sent in advance and delivered during the exceptional state is equal to the mean number of messages correctly received stored in the receiving buffer, during the transmission cycle, plus the mean number of identified B_{e_i} messages. It is given by [9]:

$$A = \sum_{i=0}^{v-1}\sum_{j=i}^{S}\frac{S-j}{j+1}\binom{S}{j}P_{be}^{j}P_{bc}^{S-j} + \sum_{i=1}^{v-1}\sum_{j=i}^{S}\binom{S}{j}P_{be}^{j}P_{bc}^{S-j} \tag{2}$$

The mean waiting time at the normal state is given by:

$$E[t_N] = \sum_{i=0}^{\infty} iP_{be}P_{bc}^{i} = \frac{P_{bc}}{P_{be}} \tag{3}$$

Using the conditional transition probabilities, applying the semi-Markov process properties and combining equations 1 and 3, the throughput efficiency can be expressed as:

$$\eta = \frac{K}{L}\frac{P_{bc} + P_{be}(A+1)}{P_{bc} + P_{be}E[t_E]} \tag{4}$$

The mean waiting time at the exceptional state is given by:

$$E[t_E] = A + 1 + E[n_e](S+1) + E[t_a], \tag{5}$$

where $E[n_e]$ is the expected number of retransmission cycles and $E[t_a]$ is the additional mean waiting time due to the additional copies of messages B_{e_i} $(i < v)$.

Let $n_{REJB_{e_i}}$ be the number of B_{e_i} go-back-N retransmissions $(0 \leq i < v)$ needed to recover this message, given that at least a B_{e_i} go-back-N retransmission occurs (in response to a REJ B_{e_i} with $B_{e_i} = N(R)$). Then:

$$E[n_{REJB_{e_i}}] = \sum_{i=1}^{\infty} i(1-P_{be}^{m})P_{be}^{m(i-1)} = \frac{1}{1-P_{be}^{m}} \tag{6}$$

A REJ B_{e_j} with $B_{e_j} = N(R)(1 \leq j < v)$ is sent in the n^{th} retransmission cycle when within this cycle, the receiver completes the recovery of all messages $B_{e_i}, i < j$ (not all B_{e_i} ones have been recovered within the $(n-1)^{th}$ retransmission cycle) and all m copies of the message B_{e_j} are received in error (the message B_{e_j} has not yet been recovered). Let $p_{n,j}$ be the probability of such situation occurs, given that in the transmission cycle there are at least j erroneously received messages. Then:

$$p_{n,j} = \left((1-P_{be}^{n.m})^{j} - (1-P_{be}^{(n-1).m})^{j}\right)P_{be}^{n.m} \tag{7}$$

The unconditional probability of occurring a B_{e_i} go-back-N retransmission $(1 \leq i < v)$ includes the probability that in the transmission cycle there are at least i erroneously received messages. Moreover, it is necessary to compute the probability of the first REJ B_{e_i} with $B_{e_i} = N(R)$ occurring during the n^{th} retransmission cycle, $n = 1, 2, 3, \ldots$. Thus:

$$Pr[REJB_{e_i}] = \sum_{n=1}^{\infty}\sum_{k=i}^{S}\binom{S}{k}P_{be}^{k}P_{bc}^{S-k}p_{n,i} \tag{8}$$

The mean number of exceptional state retransmission cycles is given by:

$$E[n_e] = \frac{1}{1-P_{be}^{m}} + \sum_{i=1}^{v-1} Pr[REJB_{e_i}].E[n_{REJB_{e_i}}]. \tag{9}$$

To calculate $E[t_a]$ it is necessary to compute the additional waiting time due to the additional message copies. Each message $B_{e_i}(0 \leq i < v)$ increases in $m-1$ time slots the waiting time of the first retransmission cycle. In following cycles, each message not recovered in the previous cycles increases this waiting time in up to m–1 time slots. Thus, the mean waiting time due to the additional copies of message $B_{e_i}(0 \leq i < v)$ is:

$$E[t_{a_i}] \leq (m-1).\frac{1}{1-P_{be}^m}\sum_{j=i}^{S}\binom{S}{j}P_{be}^{j}P_{bc}^{S-j} \qquad (10)$$

where $1/(1-P_{be}^m)$ is the average number of retransmission cycles until message B_{e_i} being successfully recovered. It corresponds to the average number of REJ B_{e_i} sent by the receiver until recovering this message (with B_{ei} equal to $N(R)$ or not).

From the above, the additional mean waiting time due to the additional copies of messages B_{e_i} is

$$E[t_a] = \sum_{j=0}^{v-1}E[t_{a_j}] \leq \frac{m-1}{1-P_{be}^m}\sum_{j=0}^{v-1}\sum_{i=j}^{S}\binom{S}{i}P_{be}^{i}P_{bc}^{S-i}. \quad (11)$$

Therefore, substituting equations 11, 9 and 5 in equation 4, a lower bound on the throughput efficiency of the proposed scheme is obtained.

5. Numerical Results

This section discusses numerical results of the model presented and compares them with other published results. Figure 2 shows the throughput efficiency of the proposed scheme for various m values over a satellite channel with a $10Mbps$ data rate and a message length of 6832 bits ($L = 6832$ and $S = 1024$). Receiver buffering and multicopy retransmissions improve the throughput efficiency for a large range of bit error rates (BER). As expected, greater values of m imply larger BER regions where the scheme works better than the conventional GB(N). Nevertheless, the incremental improvement is not linear and decreases when the m value increases. This can be explained considering that a given m value may be too much for a low BER region and not enough for a high BER region. In other words, when the channel is fairly good there will be a waste of bandwidth when the recovery is accomplished with the c^{th} $(c < m)$ retransmission copy (m–c messages will be discarded). On the other hand, there will also be a waste of bandwidth if all the m retransmitted copies are received in error and a new retransmission procedure must be initiated. The results suggest the use of low m values.

Figure 3 plots the throughput efficiency for proposed scheme and some GB(N) ones (Sastry's[1], SETRAN[4], Birrel,s[7] and Weldon's[8] with parameters set to a multicopy GB(N) protocol ($q = 0$ and $n_{q+1} = m$)). It can be seen that the use of repeated copies even for the first transmission (Birrel) results in poor performance for regions with low BER. On the other hand, the drawback of the recovery strategies that retransmit continuous repeated copies is the bandwidth waste of S messages for each negatively acknowledged message. This technique performs well when utilized with a few message copies and only during retransmissions (Weldon). Further, the use of a receiving buffer (SETRAN) can improve the throughput results in high BER region. The proposed scheme, which combines multicopy retransmission with receiving buffer, presents results that largely overcomes the others.

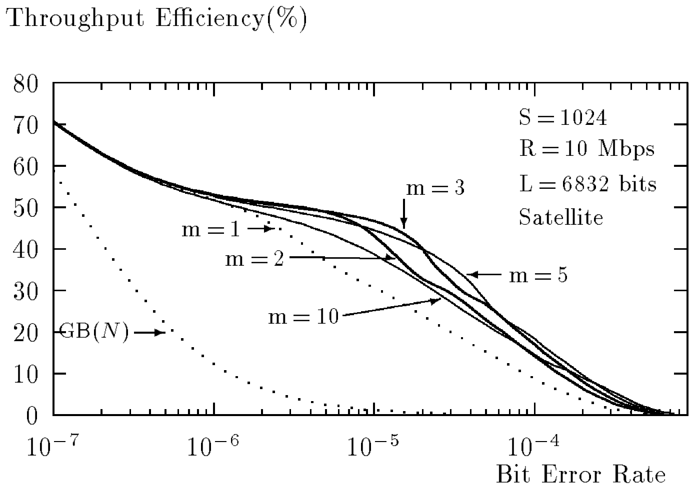

Figure 2 *Proposed Scheme Throughput Efficiency*

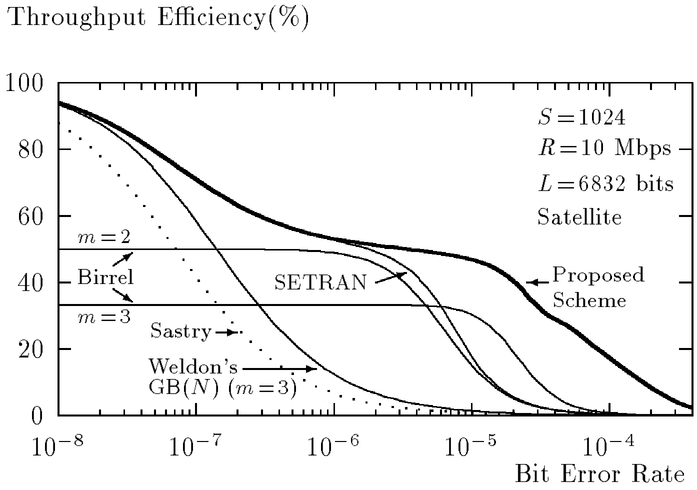

Figure 3 *Comparison of the Proposed Scheme (m=3) and Other GB(N) Protocols*

Figure 4 compares the proposed scheme with some SR ARQ strategies (Easton's basic SREJ+REJ[3], SR+ST[6], Yu&Lin's[5] and Weldon's[8]). It can be seen that even with twice less memory in the transmitter the proposed scheme performs better in the high error rate region.

Figure 5 plots the throughput efficiency as function of v, the number of erroneous messages that may be processed and reordered during a recovery procedure. The incremental improvement decreases with the increase of v. It can be seen that the improvement due to higher values of v is located in a region where the throughput efficiency is very low, and thus, of no practical use. Therefore, these results suggest that the parameter v can be used in a protocol configuration in order to establish a tradeoff between the efficiency and the receiver processing time, considering that this last one is a function of the number of erroneous messages which are received during the recovery cycle.

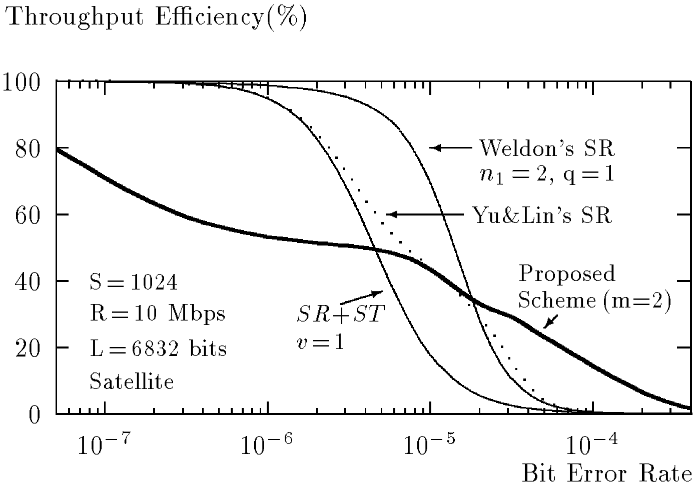

Figure 4 *New Strategy and some SR Schemes Efficiencies*

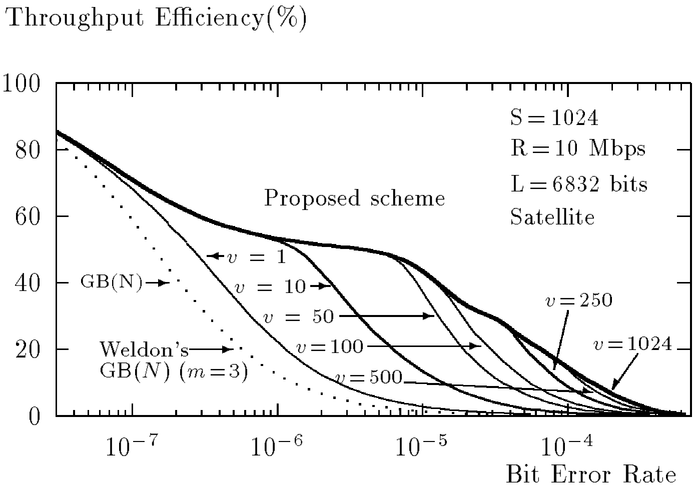

Figure 5 *Proposed Scheme* ($m=2$) *Throughput Efficiency*

6. Conclusion

This paper has presented and analysed a GB(N) scheme with multicopy retransmissions suitable for high speed satellite environments. The proposed strategy makes a best effort to recover the negatively acknowledged messages, reduces the effect of the round trip delay and performs well even under high error rate channel condition.

The throughput efficiency was obtained as a function of v, the number of erroneous messages that the receiver can process during a recovery procedure. The results show that the incremental gain in performance decreases with v. Hence, there is a tradeoff to be considered in the implementations where the receiver processing and reordering time is a function of this number.

The analytical results show that the new strategy presents a better performance than the best known GB(N) results and suggest low values for the number of repeated messages to be retransmitted. Furthermore, the proposed strategy outperforms somy finite buffer selective-repeat schemes, over a wide range of high bit error rate.

References

[1] A. R. K. Sastry, "Improving repeat-request (ARQ) performance on satellite channels under high error rate conditions", *IEEE Transactions on Communications*, vol. COM-23, no. 4, pp. 436–439, Apr. 1975.

[2] J. M. Morris, "On another Go-Back-N ARQ technique for high error rate conditions", *IEEE Transactions on Communications*, vol. COM-26, no. 1, pp. 187–189, Jan. 1978.

[3] M. C. Easton, "Batch throughput efficiency of ADCCP/HDLC/SDLC selective repeat protocols", *IEEE Transactions on Communications*, vol. COM-28, no. 2, pp. 187–195, Feb. 1980.

[4] S. Lin and P. S. Yu, "An effective error control scheme for satellite communications", *IEEE Transactions on Communications*, vol. COM-28, no. 3, pp. 395–401, Mar. 1980.

[5] P. S. Yu and S. Lin, "An efficient selective repeat ARQ scheme for satellite channels and its throughput analysis", *IEEE Transactions on Communications*, vol. COM-29, no. 3, pp. 353–363, Mar. 1981.

[6] M. J. Miller and S. Lin, "The analysis of some selective repeat ARQ schemes with finite receiver buffer", *IEEE Transactions on Communications*, vol. COM-29, no. 9, pp. 1307–1315, Sept. 1981.

[7] N. D. Birrel, "Pre-emptive retransmission for communication over noisy channels", *IEE Proceedings, Part F*, vol. 128, pp. 393–400, Nov. 1981.

[8] E. Weldon, "An improved selective repeat ARQ strategy", *IEEE Transactions on Communications*, vol. COM-30, no. 3, pp. 480–486, Mar. 1982.

[9] O. C. Duarte and H. M. de Lima, "A new selective repeat scheme - actual environment analysis", in *IEEE International Conference on Communications ICC88*, Philadelphia, pp. 22.2.1–22.2.6, June 1988.