Sub-block retransmission ARQ schemes

Youngsup Kim, Saewoong Bahk*

School of Electrical Engineering, Seoul National University, Seoul, South Korea

Received 24 February 1997; accepted 9 January 1998

Abstract

In this paper we propose a sub-block retransmission scheme for ARQ and hybrid ARQ. When the channel is quiet the sub-block retransmission scheme behaves like a conventional ARQ or hybrid ARQ scheme. As the channel becomes increasingly noisy, the data block is divided into smaller sub-blocks for transmission. Each sub-block is encoded for error control by an appropriate shortened code of which the code length is adapted to the corresponding channel BER. The received block is checked for errors sub-block by sub-block. The proposed sub-block retransmission scheme provides improved throughput over conventional ARQ schemes by retransmitting only the naked sub-blocks in the occurrence of errors. An example of transferring ATM cells is considered for simulation study. © 1998 Elsevier Science B.V.

Keywords: Retransmission protocol; Hybrid ARQ; ARQ; Shortened code; Wireless ATM

1. Introduction

There are two basic categories of error control schemes: automatic repeat request (ARQ) schemes and forward error correction (FEC) schemes. ARQ schemes are simple and provide high reliability. However the throughput is not consistent and deteriorates rapidly with increasing channel error rate. FEC schemes provide constant throughput. However the system reliability falls as the channel degrades. Hybrid ARQ schemes combine the ARQ with the FEC [1,2].

This paper presents a sub-block retransmission scheme for ARQ (SRS-A) and hybrid ARQ (SRS-HA) schemes. When the channel is quiet, SRS-HA behaves like a conventional (hybrid) ARQ scheme. As the channel becomes noisy, the data block is divided into smaller sub-blocks for transmission. Each sub-block is encoded by an appropriate shortened code where the code length is adapted to the corresponding channel BER. The encoded block is then transmitted. The received block is checked for errors sub-block by sub-block. The proposed scheme provides improved throughput by retransmitting only the naked sub-blocks in the occurrence of errors.

In Section 2 we compare the proposed scheme with error control schemes that are related with our study: pure ARQ (no FEC) and hybrid ARQ schemes. Section 3 describes in detail a single-code hybrid ARQ that is best suited for the sub-block retransmission scheme. We consider the performance analysis of the proposed scheme and an example of transferring ATM cells for simulation study, followed by conclusion in Section 4.

2. Comparison of the sub-block retransmission schemes with other competitive schemes

It is well known that the throughput of the conventional ARQ and the hybrid ARQ protocols can be improved by dynamically adapting the code length for error control to respond to actual channel conditions [3,4,6,8]. Wu [3] presented a hybrid ARQ scheme using multiple shortened codes (MSC-HA) having varying degrees of error correcting capability depending on the channel condition. MSC-HA uses shortened codes obtained from a t-error-correcting (n + m,n) cyclic code C_t. A shortened code can be encoded and decoded using the same circuits as for the original code. The transmitter assembles the data to be transmitted in blocks of k = n − m bits. The transmitter system operates in one of a finite number of states denoted S_0, S_1, S_2, ..., S_r. The state is chosen according to the channel BER p. When the channel is quiet the system operates in state S_0. In state S_0 a k-bit data block is encoded into an n-bit word based on an (n,n − m) shortened code C_0 obtained from C_t. C_0 is used for error detection only and the system behaves like a
conventional ARQ system. When the channel becomes noisy the system moves into a higher-order state.

We now present the idea of SRS-HA in comparison with MSC-HA. MSC-HA corrects errors by parts, i.e., sub-blocks. It detects error, however, by checking the whole block consisting of l, decoded sub-blocks in state $S_j$. Retransmission of the block is requested if any of the decoded sub-blocks is detected in error. In the proposed scheme errors can be detected by parts as well. However the receiver requests retransmission not by blocks but by sub-blocks. Following is a brief description of SRS-HA devised to compare it with MSC-HA. It uses the same cyclic code $C_1$ as MSC-HA from which shortened codes are obtained for error control. It also uses the same number of states as MSC-HA, i.e. $v$ states. It operates exactly the same way as MSC-HA in states $S_0$ and $S_1$. In states $S_j$ for $j = 2,3,...,v$, a k-bit data block is first divided into $l_j$ parts with lengths $k_{j_1}, k_{j_2},...,k_{j_l}$. Then each part is encoded for error detection into a word of $k_{j_i} + m$ bits for $1 \leq i \leq l_j$ using a shortened code obtained from $C_1$. Then the $(k_{j_i} + m)$-bit word is encoded for error correction into another word of $k_{j_i} + 2m$ bits using a shortened code obtained from $C_i$. $l_j$ sub-blocks of such $(k_{j_i} + 2m)$-bit words form an encoded block of $k + 2lm$ bits for transmission. When the receiver receives the encoded block, it first decodes each sub-block to attempt to correct errors. The decoded sub-block is then checked for errors. Only the sub-block(s) rather than the entire block is retransmitted if it is detected in error. The receiver sends to the transmitter a feedback message containing the acknowledgment information about the received sub-blocks. A bit map consisting of $l_j$ bits is used for the feedback message. The $i$th bit of the map gives the ack/nak status of the $i$th sub-block. Using this feedback information, the transmitter retransmits only the sub-block(s) detected in error.

Fig. 1 compares the throughput of the two hybrid ARQ schemes for selective repeat retransmission protocol with infinite buffer. Due to the sub-block status bits the proposed scheme requires slightly more overhead bits than MSC-HA. However it shows better performance than MSC-HA by retransmitting only the sub-blocks detected in error. For SRS-HA the 983-bit data block is divided into two (one 492-bit and one 491-bit), four, eight, and sixteen parts in states $S_2$, $S_4$, $S_8$, and $S_{16}$, respectively. SRS-HA is also more reliable than MSC-HA since a shortened code with smaller size has, in general, lower probability of undetected error [5].

Now consider the case of pure ARQ schemes (no FEC). The idea of error-detection by parts may be found in an ARQ protocol with adaptive block size as well if the block sizes depending on the channel condition are used so that the size of each block is a multiple of the minimum base size [4]. Martins [4] uses optimum block size for estimated channel BER. Once the block size is selected, Martins’ ARQ behaves like a conventional ARQ. As usual errors can be detected by some check sum calculated from the header and the data field. For a noisy channel the header may require more powerful capability of error protection than the data field. This is because packets may be delivered to wrong destination(s) in the occurrence of header error in many protocols such as Internet Protocol and ATM. When error control scheme is applied to the data field under such header error protecting environment, it is obvious that SRS-A performs better than Martins’ ARQ. This is because SRS-A transmits only one header for a block consisting of several sub-blocks while Martins’ ARQ transmits one header per sub-block. Fig. 2 shows that SRS-A operates very well over a wide range of BERs compared to Martins’. Here we let the header and CRC length be 32 and 16 bits respectively.

The size of the data block in SRS-A is 1024 bits equal to the size of the largest data block in Martins’ ARQ used when the channel is quiet. Five block sizes of 1024, 512, 256, 128, and 64 bits are used in Martins’ ARQ. The ideal throughput curve of Martins’ ARQ is the envelope of the throughput curves for the five block sizes. For SRS-A the data block of 1024 bits is divided into two, four, eight, and sixteen parts of equal sizes and each part is encoded by a 16-bit CRC. The ideal throughput curve is the envelope of the throughput curves derived later. It is assumed that the header is received correctly and that the feedback channel is error-free. The two schemes have exactly the same reliability.
3. Single-code hybrid ARQ scheme with sub-block retransmission

In this section we describe hybrid ARQ with sub-block retransmission scheme using a single code (SRS-SCHA). Single-code hybrid ARQ is best suited for the sub-block retransmission scheme because the single code can be used for both error correction and detection required by each sub-block [2]. The transmitter system operates in one of a finite number of states denoted $S_1, S_2, \ldots, S_n$, (no $S_0$, state). The operating state is chosen according to the channel BER $p$ as before. Let $C_r(t)$ be an $(n_0 + m - l, n_0 - l)$ shortened code obtained from a $t$-error-correcting $(n_0 + m, n_0)$ cyclic code $C$ with minimum distance of $d_{\text{min}}$. The transmitter assembles the data to be transmitted in blocks of $K$ bits. Let $d_1, d_2, \ldots, d_r$ be integer divisors of $K$ with $d_1 < d_2 < \ldots < d_r$. The data block of $K$ bits is divided into $d_i$ parts $(1 \leq j \leq r)$ in state $S_j$. Each part is encoded into a sub-block of $n_0 + m - l$ bits using the $(n_0 + m - l, n_0 - l)$ shortened code $C_r(t)$, where $l$ is determined by

$$n_0 - l = K/d_i.$$  

(1)

The encoded block consisting of $d_i$ sub-blocks is then transmitted. The procedures for the proposed SRS-SCHA are as follows.

- When a block arrives at the receiver, each sub-block is decoded using a $t'(\leq t)$-error-correcting bounded-distance decoder.

- The receiver sends the transmitter a feedback message indicating the receiver state. The receiver state message contains ack/nack fields for the block and the sub-blocks. The block ack/nack field indicates whether the transmitted block is wholly accepted at the receiver. This field plays the role of acknowledgment as in conventional ARQ schemes. The sub-block ack/nack field is a bit map consisting of $d_i$ bits. The $ith$ bit of the map represents the state of the $ith$ sub-block. A sub-block is unacknowledged if the decoder experiences a decoding failure. This occurs when the sub-block contains a detectable, but uncorrectable error pattern.

- If any of the sub-blocks within the received block is unacknowledged, the receiver state message and the acknowledged sub-blocks are saved in the receiver's buffer.

- Suppose $r(d_i - r)$ is the number of unacknowledged sub-blocks from the transmitted block. Receiving the feedback message, the transmitter retransmits only the $r$ unacknowledged sub-blocks in order.

- If the $r$ retransmitted sub-blocks are decoded without any decoding failure, the $r$ sub-blocks and the previously received $(d_i - r)$ acknowledged sub-blocks in the receiver's buffer are released to the user in order.

- If there are still any unacknowledged sub-block(s), the retransmission request is issued again in the same manner.

We now analyze the performance of the proposed scheme. Consider transferring of ATM cells over ATM-based wireless networks. An ATM-cell consists of a 5-byte header and a 48-byte data field. The ATM header may be replaced with a 4-byte wireless ATM header for wireless links as suggested previously [7]. The header for wireless networks contains a field carrying the sequence number, $SW_n$, of the transmitted cell. The sub-block retransmission scheme is applied to the 48-byte data field. The header for wireless links is protected separately by an error control scheme. Assume that the 4-byte header is encoded into a word of $h \geq 32$ bits and that the header information is correctly accepted at the receiver. Consider an $(n,k)$ shortened code $C_r(l)$ where $n = n_0 + m - l$ and $k = n_0 - l$. A $t'(\leq t)$-error-correcting bounded-distance decoder is used. The throughput of SRS-SCHA derived in Appendix A is given by

$$\eta = K \left[ nM + h + \frac{MP_r}{1 - P_r} + h \sum_{L=1}^{M} \left\{ \sum_{r=0}^{M-L} \sum_{t' = 0}^{t'} \binom{M-L}{t'} \right. \right.$$  

$$\times \left. (1 - P_r)^t \right\} \right]^{-1},$$  

(2)

where $M$ is the number of sub-blocks into which the $K$-bit data block is divided and $P_r$ is the probability of sub-block retransmission.

The feedback channel was assumed to be error-free. The accepted block error rate, $P(E)$, is defined as the ratio of the erroneous blocks accepted by the receiver to the total accepted blocks. It is given by

$$P(E) = 1 - \left( \frac{P_r}{P_r + P_{ud}} \right)^M$$  

(3)

where $P_r$ is the probability that the received word is correctly decoded and $P_{ud}$ is the probability of an undetected decoding error. The following terms in Eqs. (2) and (3) are given in the analysis of block codes [2]:

$$P_r = 1 - (P_{ud} + P_{ud})$$  

(4)

$$P_r = \sum_{i=0}^{t'} \binom{n}{i} P_r^{i}(1 - P_r)^{n-i}$$  

(5)

$$P_{ud} = \sum_{i=0}^{n} A_i P_{ud}$$  

(6)

where $A_i$ is the number of code words of weight $i$ and $P_{ud}$ is the probability that the received word falls within the Hamming sphere of radius $t'$ centered on a word which Hamming distance is $i$ from the transmitted code word. The value of $P_{ud}$ is given by

$$P_{ud} = \sum_{i=0}^{t'} \sum_{r=0}^{t'} \binom{n}{i} P_r^{i-r} (1 - P_r)^{n-r}$$  

(7)

Consider an example of using four states to illustrate this.
performance compared to the other competitive schemes by retransmitting only the naked sub-blocks in the occurrence of errors. The hybrid ARQ with sub-block retransmission was applied to the ATM cell as an example. The simulation results showed that the simulated throughput points follow the ideal throughput curve very closely. The proposed sub-block retransmission scheme improved the throughput and the reliability by using the shortened codes and dynamically adapting the code length to the varying channel BER.

Together with the FEC scheme and the hardware of a previous study [8], the proposed scheme can provide a variety of error control capabilities that are required for wireless ATM links. The sub-block retransmission scheme uses slightly longer feedback messages compared to the conventional ARQ schemes because the messages include sub-block ack/nak field. The increased size may be acceptable considering the overall performance improvement.

Appendix A Derivation of Eq. (2)

The throughput of the pure ARQ and the hybrid ARQ schemes can be expressed as

$$\eta = \frac{K}{E[T]}$$  \hspace{1cm} (A.1)

where $K$ is the number of information bits within the given block, $T$ is a random variable representing the number of transmitted bits required for the given block, and $E[\cdot]$ denotes the expectation of the indicated quantity. We can express $T$ as

$$T = (Mn + h) + \sum_{i=1}^{Z} T_i$$  \hspace{1cm} (A.2)

where $T_i$ is a random variable representing the number of transmitted bits for the $i$th retransmission. Then the average number of transmitted bits required for the given block can be expressed as

$$E[T] = (Mn + h) + \sum_{i=1}^{Z} E[T_i]$$  \hspace{1cm} (A.3)

We now derive the average number of transmitted bits for the $i$th retransmission, $E[T_i]$. The random variable, $T_i$, takes the value $Ln + h$ if $L$ out of $M$ sub-blocks are retransmitted at the $i$th retransmission. The event that a particular sub-block is transmitted at the $i$th retransmission, occurs if the sub-block is detected in error $i$ times. The probability of this event is $P_{i}$. Then the probability that $L$ out of $M$ sub-blocks are retransmitted at the $i$th retransmission, is given by

$$\Pr(T_i = Ln + h) = \binom{M}{L} p_{i}^L (1 - P_{i})^{M-L}$$  \hspace{1cm} (A.4)
Then $E[T_r]$ is given by

$$E[T_r] = (Mn + h) + n \sum_{i=1}^{\infty} \sum_{L=1}^{M} (Ln + h) \binom{M}{N} P_r^L (1 - P_r)^{M - L}.$$  

$$= (Mn + h) + n \sum_{i=1}^{\infty} \sum_{L=1}^{M} L \binom{M}{L} P_r^L (1 - P_r)^{M - L}.$$  

$$+ h \sum_{L=1}^{M} \binom{M}{L} \sum_{i=1}^{\infty} P_r^L (1 - P_r)^{M - L}.$$  

(A.5)

Applying the following identity (the mean of a binomial random variable)

$$\sum_{k=1}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = np$$  

(A.6)

to the second term of Eq. (A.5) and applying the following

$$(1 + x)^n = \sum_{i=0}^{n} \binom{n}{k} x^k$$  

(A.7)

to the third term of Eq. (A.5), we can obtain

$$T = (Mn + h) + n \sum_{L=1}^{M} MP_r^L + h \sum_{L=1}^{M} \binom{M}{L} \sum_{i=1}^{\infty} P_r^L (1 - P_r)^{M - L}.$$  

$$\times \binom{M - L}{r} (1 - P_r)^{r} = Mn + h + nM \sum_{i=1}^{\infty} P_r^L + h \sum_{L=1}^{M} \binom{M}{L} \sum_{i=1}^{\infty} P_r^L (1 - P_r)^{M - L}.$$  

$$\times \left( \binom{M - L}{r} (1 - P_r)^{r} \right) = Mn + h + nMP_r^L + h \sum_{L=1}^{M} \binom{M}{L} \sum_{i=1}^{\infty} P_r^L (1 - P_r)^{M - L}.$$  

$$\times \left( \binom{M - L}{r} (1 - P_r)^{r} \right) = Mn + h + nMP_r^L \frac{1}{1 - P_r^{L+r}}.$$  

(A.8)

References


Youngsup Kim was born in Seoul, Korea in 1970. He received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, Korea, in 1994 and 1996, respectively. He has been with SK Telecom since 1996 as a Research Engineer. His research interests are in mobile telecommunications.

Saewoong Balik received the B.S. and M.S. degrees in electrical engineering from Seoul National University in 1984 and 1986 respectively, and the Ph.D. Degree from the University of Pennsylvania in 1991. From 1991 through 1994 he was with the Department of Network Operations Systems at AT&T Bell Laboratories, as a Member of Technical Staff where he worked on the projects for the development of network operation systems for AT&T networks. In 1994 he joined the faculty of the School of Electrical Engineering at Seoul National University, where he is currently an Assistant Professor. His areas of interests include performance analysis and design of communication networks, interworking of ATM and non-ATM networks, wireless ATM networks, bandwidth allocation and traffic control at high speed networks, and high-speed network protocols.