Provision of Guaranteed QoS with Hybrid Automatic Repeat Request in Interleave Division Multiple Access Systems

Muhammad Majid Butt
Faculty of Electrical Engineering
National University of Computer & Emerging Sciences
B-Block, Faisal Town, Lahore, Pakistan
Email: majid.butt@nu.edu.pk

***Abstract*— Provision of guaranteed quality of service (QoS) in wireless communication has always been a demanding task. QoS can be ensured by the mechanism of repeat request, however real time systems can tolerate only a finite delay. In this paper we investigate the problem of provision of guaranteed QoS by hybrid automatic repeat request (HARQ) schemes, based on soft outputs of a channel decoder. Focus is on adaptive interleave-division multiple access (IDMA) transmission. Such a system is able to provide virtually arbitrary bit error rates (BER).**

The proposed HARQ schemes use the reliability information provided by the channel decoder to decide whether a packet satisfies the quality of service (QoS) requirements. QoS can be specified in terms of a minimum required throughput/data rate or a minimum required BER. Effect of truncation of repeat requests on the bit error rate and *packet throughput* is investigated. An adaptation scheme based upon the bit error rate in the accepted packets is proposed and the effect of the adaptation on the *effective throughput* is demonstrated.

I. Introduction

In future wireless systems, adaptivity and scalability are important features for efficient provisioning of different quality of service (QoS) classes for different applications.

Traditionally, two types of channel coding schemes exist. In forward error correction (FEC) scheme redundancy is employed for encoding and decoding the data. Automatic repeat request (ARQ) schemes belong to the class of error correcting schemes where no code is used for error correction but receiver uses some error detecting code to detect the errors and provides the information about the quality of the received data on an explicit feedback channel. A combination of FEC and ARQ is termed as Hybrid ARQ (HARQ).

In this paper, we use HARQ to provide a minimum target bit error rate (BER) for a wireless transmission system based on interleave-division multiple access (IDMA) [1], [2]. IDMA has been proposed for use in next generation wireless systems [3], [4], [5]. Besides of its high spectral efficiency, IDMA is highly scalable and adaptive. Based on a soft adaptation strategy it is able to provide different data streams with different QoS parameters (such as bit error probability and data rate) to users or applications. An adaptation strategy as proposed in [3] is only reasonable if error control and recovery on the data link layer is also adaptive.

In the following, IDMA based transmission system and HARQ protocols are introduced. A reliability based retransmission criterion and the idea of packet combining for HARQ based IDMA system are discussed briefly. The impact of truncated ARQ and the adaptation of spreading factor on overall system performance is investigated and results are demonstrated.

A. Adaptive Interleave-Division Multiple Access (IDMA)

In Direct Sequence spread spectrum Code Division Multiple Access (DS-CDMA), distinct data streams $\mathbf{d}_m$ are distinguished by *different spreading sequences*. Forward error correction (FEC) coding is typically done before interleaving and spreading. Hence, interleaving is performed on a symbol-by-symbol basis. Conventionally, the same FEC encoder and the same interleaver is used for all data streams $\mathbf{d}_m$.

In IDMA, FEC encoding and spreading may be done jointly by a single low-rate encoder, subsequently denoted by ENC, followed by an interleaver. Each encoded and interleaved sequence $\mathbf{x}_m$ is referred to as a code word in the following. All data streams use the same low-rate encoder. However, for every code word *different interleavers* are used. This is often referred to as chip-interleaving. As an alternative to the described re-arrangement between interleaving and spreading, IDMA may be interpreted as DS-CDMA without spreading or as DS-CDMA with a spreading length of one.

To provide different data rates to certain users, multiple code words can be linearly superimposed in order to enhance the data rate per user, a concept which is called multi-code technique in UMTS Terristrial Radio Access (UTRA).

One main advantage of IDMA is the availability of low complexity, iterative receiver structures which deliver reliable soft outputs [1]. Based on the soft outputs, a soft adaptation strategy was proposed in [3]. This adaptation strategy adjusts the number of code words (assigned to a user or application) and the transmit power based on the soft outputs of the iterative receiver structure. The goal of the adaptation strategy is to ensure QoS in terms of a maximum bit error rate and minimum data rate while on the other hand using minimum transmission power.

Such an adaptation is only reasonable if the upper layers are designed to support such an adaptivity. Traditionally, ARQ is designed for quasi error-free transmission. In our case, however, the ARQ protocol has to be adapted to the required BER. A further advantage of IDMA, the possibility of high throughput, should not be altered by the upper layers.

B. Hybrid Automatic Repeat Request (HARQ)

Analysis of pure ARQ schemes reveal the problem that an increase in retransmission requests has a severe impact on the throughput as the channel quality deteriorates [6]. The hybrid protocols counter this effect through the use of forward error correction in conjunction with error detection.

The FEC portion of the protocol is designed to correct the errors caused by noise or interference on the channel, while the error detection is used to detect the remaining errors. Hybrid protocols can thus provide throughput similar to that of FEC systems, while offering a reliability performance typical of ARQ protocols.

HARQ schemes are widely classified into two categories: Type-I HARQ and Type-II HARQ schemes. The Type-I Hybrid ARQ protocol is the simplest of the hybrid protocols. Each packet is encoded two times, for error detection and error correction. The data is first encoded using a high rate error detecting code. A cyclic redundancy check (CRC) code is a popular choice for error detection. The encoded data is then encoded again using a FEC code. On the receiver side, the transmitted packet is first decoded by employing the FEC code, and then decoded by employing the error detecting code. If an error is detected, a negative acknowledgement (NAK) is generated and sent to the transmitter. The same packet is retransmitted again and the process is repeated unless the packet is correctly received and positively acknowledged (ACK) by the receiver.

Type-II Hybrid ARQ protocols adapt to changing channel conditions by means of incremental redundancy. When the sender receives a NAK, it responds by sending additional code bits to the receiver. The receiver appends these bits to the received packet, allowing for increased error correction capacity. Type-II HARQ systems can be taken as special case of *code combining* systems in which packets are concatenated to form noise corrupted code words of increasingly lower code rates. Another class of packet combining systems is termed as *diversity combining* systems in which the individual symbols from multiple, identical copies of a packet are combined to create a single packet with more reliable constituent symbols. Code combining systems have been discussed by Chase and found superior in jamming environment [7].

C. Soft Packet Combining and Retransmission Criterion

HARQ systems require a retransmission criterion for the decision of acceptance or rejection of the received packet. In [8], reliability based retransmission criteria have been proposed for the IDMA system and average bit error probability (BEP) based criterion has been found to provide the best throughput result at a specific target bit error rate (BER).

Packet combining is a technique which stores the rejected packets and combines them with the retransmissions to improve the chances of acceptance of the new packet. In [8], the AddLLR combining technique has been found to provide the best performance by adding the log likelihood ratios (LLR) of the corresponding bits in all the transmissions of a packet,

$$Y^i[k] = L^i[k] + Y^{i-1}[k] \; . \tag{1}$$

$\mathbf{L}^i$ denotes the ith received transmission of a packet, while $\mathbf{Y}^i$ denotes the packet resulting from the combination after the ith transmission.

II. Transmission System

The transmission system assumed in this paper is depicted in Fig. 1. To enable the system for ARQ, the data sequence $\mathbf{d}_m$ of the mth user consisting of K information bits is stored in a buffer at the transmitter side. The data sequence $\mathbf{d}_m$ is also called a packet. After the buffer, the data is multiplexed to G different layers of the IDMA-based air-interface. The superimposed code words experience additive white Gaussian noise and interference by other users. At the receiver side the IDMA receiver proposed in [1] is used to provide soft estimates of the information bits. An evaluation module decides whether a packet fulfills the BER requirements. Therefore, it either accepts a packet and sends an ACK over the feedback channel or discards the packet and sends a NAK. The feedback channel itself is assumed to be perfect. While conventional ARQ schemes make decisions based on an error detecting code, we use criterions based on the soft outputs provided by the IDMA receiver described in detail. For this reason, no error detecting code, as a CRC code, is needed in these schemes.

The packet combiner is used for combining the information bits from the retransmitted packets. It stores copies of discarded packets. These copies are afterwards combined with the newly received packets.

III. Adaptation for Real Time Systems

AddLLR combining in conjunction with the average BEP based criterion provides the target BER in the accepted packets but this requirement can only be guaranteed if infinite retransmissions are allowed. In real time systems, this condition can not be fulfilled all the time. In this section, some adaptation schemes are proposed.

A. Deadline Dependent Coding (DDC)

Deadline dependent coding (DDC) is the concept of making a communication protocol deadline dependent [9]. The protocol should try to minimize bandwidth, transmit energy and the time required to successfully deliver the information. There are two classes of real time systems: *hard* and *soft* real time systems [10]. In a hard real time system, late delivery cannot be tolerated. In contrast, soft real time system can tolerate a specified low probability of late delivery.

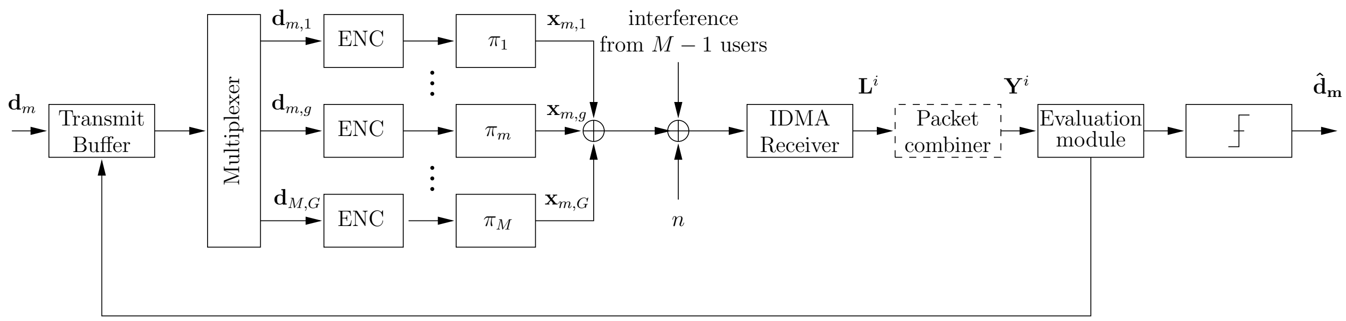

Fig. 1. IDMA-based transmission system including a hybrid ARQ scheme with optional packet combiner.

In this paper, we consider the time deadline of the data delivery as the hard constraint and try to minimize the BER, i.e. the BER is considered as a soft constraint. We limit the maximum number of transmissions to T_{max} and accept the packet after T_{max} transmissions ignoring the retransmission criterion. This type of ARQ is termed as *truncated ARQ* [11]. We set the receiver iterations to a fixed value. It is different from [11] where after each receiver iteration, stopping criterion is tested and packet is NAKed if stopping criterion is not fulfilled after the maximum number of allowed iterations. In our proposed scheme, after the fixed receiver iterations and AddLLR combining, the packet is tested against the average BEP based retransmission criterion. If the retransmission criterion is satisfied, resultant packet is accepted, otherwise it is stored in the buffer for later use. After the maximum number of allowed transmissions, if retransmission criterion is not satisfied, resultant packet is still accepted to satisfy the requirements of provision of a minimum guaranteed data rate.

B. Adaptation of Spreading Factor (SF)

In the literature [12] adaptive modulation and channel coding (AMC) is widely discussed as a form of adaptation. According to [12], (H)ARQ is essential to compensate for erroneous packet transmission in AMC caused by the false selection of modulation and channel coding scheme. We use the adaptation of the spreading factor (or code rate) as an adaptation strategy. User data varies all the time in the real scenarios and the spreading factor can be adapted to meet the requirements. We use a fixed data size of the packet in our simulations. The idea is to start a new transmission always with *small* spreading factor and *increase* it if packet is not accepted by the receiver and this information needs to be conveyed to the transmitter on the feedback channel. After a packet is accepted, spreading factor is again set to the low value. This strategy may originate unnecessary retransmissions but it has the potential of providing significant advantage when user density is low or channel conditions are good.

IV. Numerical Results

Performance of a system employing a HARQ scheme is measured by throughput. We have defined two parameters namely *packet throughput* and *effective throughput* to compare the performance of the different strategies proposed. Packet throughput is defined as the number of packets accepted by the receiver per transmitted packet.

$$\eta_{packet} = \frac{\text{number of accepted packets}}{\text{number of transmitted packets}}. \tag{2}$$

On the other hand, effective throughput is more precise measure of the effectiveness of any scheme as it takes the performance of FEC code and idle time into account. It is defined as the number of bits accepted by the receiver per transmitted bit.

$$\eta_{packet} = \frac{\text{number of accepted bits}}{\text{number of transmitted bits}}. \tag{3}$$

Idle time or round trip time (RTT) is an other important parameter affecting system performance. It is defined as the time transmitter has to wait after sending a packet and before receiving the ACK/NAK for the packet. It is measured as the amount of data that a transmitter can send but it is not able to send due to waiting for ACK/NAK. It is waste of the resources of the system such as bandwidth and processing speed at the transmitter and the receiver. For the IDMA system, effective throughput is defined as,

$$\eta_{eff} = \frac{MK_m}{SF_1(MK_m) + SF_2(B) + T_{tr}(\Gamma)} \tag{4}$$

where M = Number of layers, K_m = Information bits per layer, SF_1= Low Spreading factor, SF_2= High Spreading factor, T_{tr} = Number of transmissions and Γ = Idle time.

$$\Gamma = D(RTT + \text{processing time}) \text{ bits} \tag{5}$$

where D is data transmission rate.

$$RTT = 2 * \frac{d}{c} \tag{6}$$

For IDMA, at bandwidth of 40 MHz with a bit load (ratio of layers to SF) equal to 2, transmission rate D is calculated as 80 Mb [5]. With the distance of 10 km between the transmitter and the receiver, at speed of light, RTT is equal to 53.2 kbps. This factor pulls the effective throughput of a *stop and wait* (SW) based HARQ protocol down for all the retransmissions. In real time systems, *sliding window* based

HARQ protocols are always preferred to minimize this effect. In this work, effective throughput has been normalized to compare the effects of various adaptation on the system. For the simulations of truncated (H) ARQ a spreading factor of 10 has been used and 20 superimposed code words of 640 bits have been allocated to a single user. The results at a single SNR point are based upon 1000 packet transmissions. The simulations show the results for a single user for whom the ARQ schemes have been applied but the idea can easily be extended to multi users.

Fig. 2 and 3 show the effect of limiting the maximum number of allowed transmissions on throughput and BER. Fig. 2 and 3 show that it is possible to ensure BER performance of the packets better than the target BER, if we do not limit the transmissions. In case of limiting the maximum number of transmissions, there is no guarantee in terms of BER performance up to a specific bit energy level and this energy level corresponds to the point where throughput curves for the limited number of transmissions converge to the curve for unlimited transmissions. This point is called *convergence point* here. After the convergence point QoS requirements of both target BER and target throughput are satisfied. At energy level less than the convergence point, target data rate can be achieved but BER is higher than the target BER. e.g. for the minimum throughput of 20 % and 33 % this point is 4 dB and 5 dB respectively as shown in Fig. 2 and 3. Therefore, a guaranteed QoS in terms of both BER and data rate is possible by the mechanism of HARQ schemes by providing specific bit energy levels.

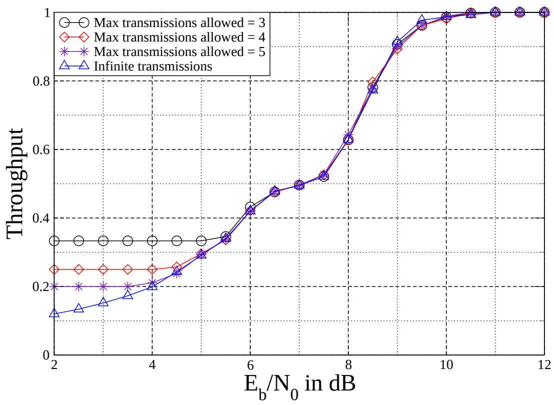

Fig. 2. Comparison of Packet throughput at target BER of 10^{-4} when maximum number of retransmissions are limited.

Fig. 4 and 5 show the effect of adaptation of spreading factor on the throughput and effective throughput respectively. From Fig. 4, it is clear that an increase in spreading factor results in an increase in packet throughput, which is obvious. The real advantage of this adaptation can be seen in Fig. 5 where effective throughput (defined in Eq(3)) is plotted. The figure shows that at low bit energy levels, higher spreading factor is advantageous but if more power is available (or users density is low), use of small spreading factor can provide better performance and capacity.

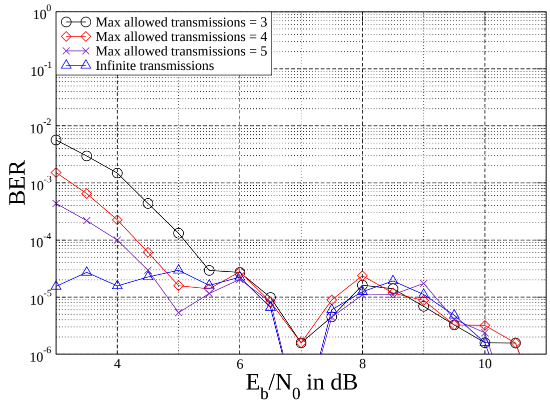

Fig. 3. BER in truncated (H)ARQ systems at target BER of 10^{-4}.

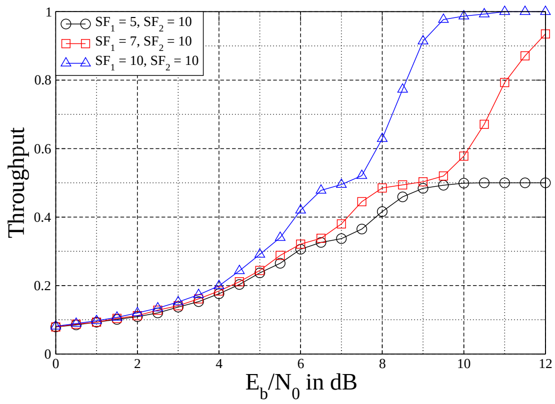

Fig. 4. Packet throughput for the adaptation of spreading factor in case of retransmissions.

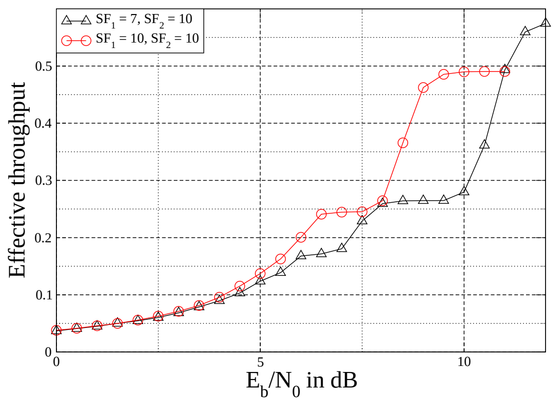

Fig. 5. Effective throughput for the adaptation of spreading factor in case of retransmissions.

V. Conclusions

In [8], HARQ scheme based upon AddLLR packet combining and average BEP based retransmission criterion is proposed by allowing unlimited transmissions. In this paper this idea is extended to the limited number of transmissions. Aim is to guarantee a certain *minimum throughput* by allowing bit errors more than target BER below a certain threshold energy level. Effect of *deadlines* on the ideal system is examined. HARQ scheme provides the guaranteed QoS in terms of BER only if retransmissions are not limited. Contrarily, if a minimum throughput is the more desirable feature, then retransmissions are limited and no guarantee is provided about the BER at the bit energy less than the *convergence point*. After this point both *deadlines* of BER and data rate can be met simultaneously. The idea of *convergence point* (discussed in Sec:IV) has been proposed here for the provision of guaranteed QoS in terms of both minimum BER and minimum throughput of the multiuser wireless communication systems.

An adaptation scheme based upon alternating spreading factor is proposed. In case of low user density or high power, this adaptation provides better capacity than the fixed spreading factor scheme. This adaptation has the potential of even better performance in the presence of selected retransmission schemes where in stead of the whole coded packet, only a part of the redundancy or some puncturing pattern is retransmitted. In this case, first transmission is done at low spreading factor and high spreading factor is applied only on the redundancy information which is much smaller in size than the original data. Therefore, overall code rate is expected to improve as compared to fixed spreading factor schemes.

REFERENCES

[1] Lihai Liu, W. K. Leung, and Li Ping, "Simple iterative chip-by-chip multiuser detection for CDMA systems," in *Proc. 57th IEEE Semiannual Vehicular Technology Conference (VTC 2003-Spring)*, vol. 3, April 2003, pp. 2157–2161.

[2] Li Ping, "Interleave-division multiple access and chip-by-chip iterative multi-user detection," *IEEE Commun. Mag.*, vol. 43, pp. S19–S23, June 2005.

[3] H. Schoeneich and P. A. Hoeher, "Adaptive interleave-division multiple access - a potential air interface for 4G bearer services and wireless LANs," in *Proc. 1st IEEE and IFIP Int. Conf. on Wireless and Optical Communications and Networks (WOCN 2004)*, Muscat, Oman, June 2004, pp. 179–182.

[4] Shidong Zhou, Yunzhou Li, Ming Zhao, Xibin Xu, Jing Wang, and Yan Yao, "Novel techniques to improve downlink multiple access capacity for beyond 3G," *IEEE Commun. Mag.*, vol. 43, pp. 61–69, January 2005.

[5] J. Ch. Fricke, P. A. Hoeher, and H. Schoeneich, "An uplink proposal based on interleave-division multiple access," in *Proc. Wireless World Research Forum 14th Meeting (WWRF14)*, San Diego, California, USA, July 2005, paper no. WG4-16.

[6] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Prentice Hall, 1995.

[7] D. Chase, "Code combining–a maximum likelihood decoding approach for combining an arbitrary number of noisy packets," *IEEE Trans. Commun.*, vol. 33, pp. 385–393, 1985.

[8] M. M. Butt, J. Ch. Fricke, and P. A. Hoeher, "Reliability-based packet combining with applications to interleave- division multiple access," in *Proc. Int. Symp. On Turbo Codes and Related Topics in conjunction with Int. ITG Conf. on Source and Channel Coding, Munich, Germany,*, April 2006.

[9] E. Uhlemann, T. M. Aulin, L. K. Rasmussen, and P.-A. Wiberg, "Deadline dependent coding with concatenated hybrid ARQ schemes for wireless real-time communication," in *Proc ARTES Graduate Student Conference, Uppsala*, April 2002.

[10] E. Uhlemann, "Adaptive concatenated coding for wireless real-time communications," Ph.D. dissertation, Chalmers University of Technology, September 2004.

[11] E. Uhlemann, T. M. Aulin, L. K. Rasmussen, and P.-A. Wiberg, "Packet combining and doping in concatenated hybrid ARQ schemes using iterative decoding," in *Proc. 2003 IEEE Wireless Communications and Networking Conference (WCNC 2003)*, New Orleans, LA, USA, March 2003, pp. 849–854.

[12] N. Miki, H. Atarashi, S. Abet, and M. Sawahashi, "Comparison of throughput employing hybrid ARQ packet combining in forward link OFDM broadband packet wireless access," *IEEE Trans. Commun*, vol. 88, no. 2, pp. 594–603, February 2005.