Multiple ARQ Processes for MIMO Systems

Haitao Zheng

Wireless Research Laboratory, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, NJ 07733, USA
Email: haitaoz@lucent.com

Angel Lozano

Wireless Research Laboratory, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, NJ 07733, USA
Email: aloz@lucent.com

Mohamed Haleem

Wireless Research Laboratory, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, NJ 07733, USA
Email: haleem@lucent.com

We propose a new automatic repeat request (ARQ) scheme for MIMO systems with multiple transmit and receive antennas. The substreams emitted from various transmit antennas encounter distinct propagation channels and thus have different error statistics. When per-antenna encoders are used, separating ARQ processes among the substreams results in a throughput improvement. Moreover, it facilitates the interference cancellation in certain MIMO techniques. Quantitative results from UMTS simulations demonstrate that the proposed multiple ARQ structure yields more than 30% gain in link throughput.

Keywords and phrases: MIMO systems, automatic repeat request, throughput, wireless communication, UMTS.

1. INTRODUCTION

Third-generation cellular systems are being designed to support high-speed packet data services. In the downlink, which has more stringent requirements in many of such services, high-speed packet access is provided through a shared channel where time-division multiplexing is used. Time slots are assigned to users at specific data rates through a scheduling algorithm based on the user data backlog and on channel quality indication (CQI) received via a feedback channel.[1] Such a transmission scheme allows multiple users to share the system resources efficiently by adapting to traffic and channel variations and it also avoids possible resource limitations that might occur if each user were allocated a dedicated code-multiplexed channel. Therefore, it has the potential to improve the capacity for delay-tolerant bursty services. Examples where this scheme will be implemented include the CDMA 1x EV-DO and 1x EV-DV and the UMTS high-speed downlink packet access (HSDPA) [1, 2]. Several advanced technologies are employed in high-speed downlink transmission to improve link throughput or reduce packet delay by adapting to the time-varying channel conditions,

traffic statistics, and quality-of-service requirements. Some of these adaptive techniques, relevant to this paper, are summarized below.

Multiple transmit and receive antennas. The use of multiple antennas at each base station sector is already part of every third-generation standard. In the downlink, specifically, these antennas can be used to provide transmit diversity and/or to direct a beam towards the intended terminal. The deployment of multiple receive antennas at data terminals is also being considered. The combination of multiple transmit and receive antennas will enable the implementation of a number of multiple-input multiple-output (MIMO) techniques that promise spectacular increases in throughput without the need for additional power or bandwidth [3, 4, 5].

Dynamic link adaptation through adaptive modulation and coding. Typically, each transmission in the downlink shared channel is at the maximum available power, with no power control. Therefore, link adaptation [6, 7], which adjusts the modulation and coding schemes (MCS), provides an efficient way of maximizing the instantaneous usage of the wireless channel. Specifically, it enables the use of aggressive MCSs when channel conditions are favorable while it reverts to MCSs that are more robust but with lower transmission rates when channel conditions degrade. The base station

---

[1]Each terminal measures its channel condition and translates it into a metric to be fed back to the serving base station.

selects the appropriate MCS based on the CQI for the user served at each time slot. We hereby refer to the MCS selection process as the mapping design.

Automatic repeat request (ARQ) or hybrid ARQ (HARQ). The performance of MCS-based link adaptation largely depends on the accuracy of the CQI, which is difficult to maintain as velocity increases. The delay tolerance of many data services enables the use of retransmission schemes to recover erroneous packets. Recently, HARQ techniques have been adopted by several wireless standardization bodies, for example, 3GPP and 3GPP2. HARQ [8, 9, 10] can improve throughput performance, compensate for link adaptation errors, and provide a finer granularity in the rates effectively pushed through the channel. Upon detecting a transmission failure, mostly by cyclic redundancy check (CRC), the terminal sends a request to the base station for retransmission. The delay due to packet acknowledgement can be significantly reduced by placing the HARQ functionality in the base station (Node B in UMTS) rather than in the radio network controller (RNC in UMTS). The packet decoder at the mobile combines the soft information of the original transmission with those of the subsequent retransmissions. The combined signal has higher probability of successful decoding. In general, there are two ways of soft combining. With *chase combining*, the base station repeatedly sends the same packet and the receiver aggregates the energy from the (re)transmissions to improve the signal-to-noise ratio (SNR) [11, 12]. A more sophisticated HARQ mechanism, named *incremental redundancy* (IR), transmits additional redundant information in each retransmission and gradually reduces the coding rate until successful decoding occurs [13, 14, 15]. Compared with chase combining, IR requires larger receiver buffers but it can achieve better performance [16]. It also provides finer granularity in the encoded rates and allows for better adaptation to channel variations.

Scheduler. In a multiuser system where user channel conditions change over time, a scheduler can exploit those channel variations by giving certain priority to the users with transitorily better channels. The scheduler critically impacts the system performance. Several scheduling algorithms have been proposed in the literature to maximize the packet data throughput, subject to various fairness conditions [17].

The above technologies are tightly coupled. However, since some of them reside in different layers, that is, HARQ in the medium access control (MAC) layer and MIMO in the physical layer, they are usually discussed and treated separately. The evaluation of each technology fails to take into account the performance improvement or degradation brought about by the other one. In particular, the link layer performance of any MIMO algorithm is usually selected according to the raw data rate at some operating point, for example, 10% packet error rate. However, when some level of channel uncertainty exists and the system supports HARQ, it may be beneficial to transmit aggressively at higher packet error rates and recover channel errors through retransmissions [18]. The throughput depends heavily on the transmission strategy. An overly aggressive transmission could produce too many unsuccessful packet transmissions that diminish the overall throughput, while an overly conservative one fails to fully utilize the channel. In this case, the overall throughput depends on the algorithms at both layers and only cross-layer design can enable the most efficient use of the channel.

In this paper, we address some of the key design issues associated with the choice of the HARQ structure to be used for MIMO physical layer transmission. We propose a new HARQ structure that matches the layered structure of the most popular MIMO architectures [19]. Simulation results show that the performance sensitivity to the choice of HARQ depends on the aggressiveness of the transmissions and on the type of CQI.

The paper is organized as follows. In Section 2, we describe the layered architectures with per-antenna encoding. Modifications to the conventional HARQ structure to fit these layered architectures are discussed in Section 3. We compare the performance of different HARQ structures in Section 4. Conclusions are drawn in Section 5.

2. LAYERED ARCHITECTURES WITH PER-ANTENNA ENCODING

In order to approach the MIMO channel capacity in rich multipath environments, the substreams radiated from the various transmit antennas should be uncorrelated [20, 21]. Nonetheless, it may in practice be advantageous to jointly encode them (Figure 1a). This has motivated a blossoming interest in the design of space-time (vector) codes [22]. Clearly, when the substreams are jointly encoded, they should share a single CRC.

The complexity of joint detection, however, explodes as the number of transmit antennas grows large. As a result, there has also been strong interest in devising alternative approaches. One such approach is that of layered architectures, which incorporate multiple scalar encoders, one per transmit antenna. In these architectures, input data is demultiplexed into multiple substreams, which are then separately encoded and radiated from the various transmit antennas (Figure 1b). At the receiver, the substreams are successively detected and cancelled [4, 5]. Specifically, the information extracted from each substream is reencoded, interleaved, and modulated to construct a replica of the transmitted substream. This replica, properly combined with the channel response, is then subtracted from the overall received signal so that—if there are no errors—the interference contribution of this substream is removed. The complexity of these architectures increases more gracefully with the number of antennas. Furthermore, they can capitalize on existing scalar coding formats.

A layered architecture can approach the MIMO channel capacity if the data rates of the different transmit antennas are appropriately adjusted [23, 24]. This adjustment requires separate CQI, one per transmit antenna, and thus the amount of feedback required increases linearly with the number of transmit antennas. We hereby refer to it as per-antenna rate and CQI. Alternatively, a common CQI—and thus the same data rate—can be used for all transmit
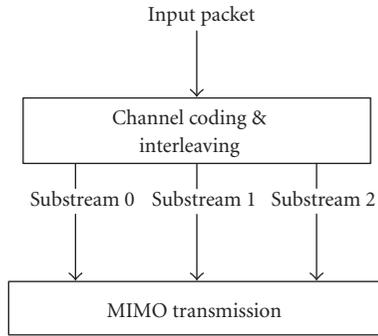
(a) MIMO with joint coding.
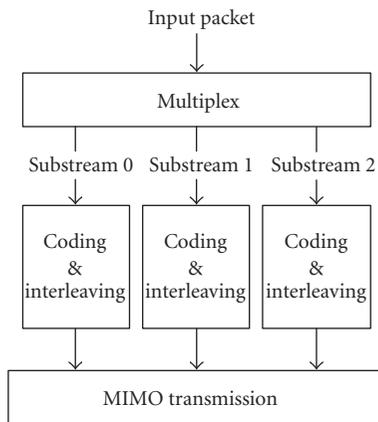
(b) MIMO with per-antenna coding.

FIGURE 1: MIMO transmitter architecture with different coding structures.

FIGURE 2: Ergodic Shannon capacity with 4 transmit and 4 receive antennas obtained via Monte Carlo simulation on a Rayleigh-faded channel with no antenna correlation.

antennas at the expense of some loss in capacity [23]. To illustrate this point, Figure 2 depicts the difference between the capacity with and without the constraint that the data rate at each of the transmit antennas be equal, for the specific case of 4 transmit and 4 receive uncorrelated antennas with Rayleigh fading. For the purpose of this paper, in any event, the most relevant feature of a layered architecture is that it does not constrain the transmit antennas to be jointly encoded and share a unique CRC.

3. HARQ MECHANISMS FOR MIMO SYSTEMS

If the MAC layer is unaware of the presence of MIMO at the physical layer, HARQ simply attaches a single CRC to the packet with such CRC encompassing the data radiated from the various transmit antennas. We refer to this scheme, depicted in Figure 3a, as MIMO single ARQ (MSARQ). Since substreams transmitted from different antennas encounter distinct propagation channels, they have different error statistics. Using a typical channel propagation model with 4 transmit and 4 receive uncorrelated antennas [21], we ob-

serve that in more than 70% of error events,[2] only the substreams from 1 or 2 transmit antennas are corrupted and thus require a retransmission (Figure 4). However, upon an error event, an MSARQ receiver has to request a retransmission of the entire packet because it relies on the single CRC over the whole packet. Retransmitting substreams that have already been correctly received wastes throughput. When multiple per-antenna encoders are used, it becomes possible to remove the constraint that the substreams radiated from multiple transmit antennas share a single ARQ process.

For per-antenna MIMO encoding architectures, we herein propose to employ multiple ARQ processes, 1 for each substream radiated from 1 transmit antenna or group of antennas. This scheme is independent of the receiver-processing algorithm and only requires that the receiver decodes substreams independently. We refer to this scheme as MIMO multiple ARQ (MMARQ). As shown in Figure 3b, a CRC symbol is appended to each substream. At the receiver, each such substream is decoded and the associated CRC is used to validate the content. Multiple acknowledgment (NACK/ACK) indications are then sent back to the transmitter. After receiving these acknowledgements, the transmitter sends fresh packets from the transmit antennas that have been successfully acknowledged and retransmits the substreams that have been negatively acknowledged through their associated transmit antennas. Hence, the HARQ operations at different transmit antennas are independent of each other. We focus on high-speed downlink data transmission so that the overhead due to multiple CRC symbols is negligible. However, we need to consider the uplink signaling overhead due to multiple acknowledgements. For each ARQ process, NACK/ACK requires an overhead of 1 bit plus error protection redundancy. Therefore, the amount of ARQ feed-

---

[2]An error event occurs when any of the substreams contains an error.

(a) MSARQ transmitter.

(b) MMARQ transmitter.
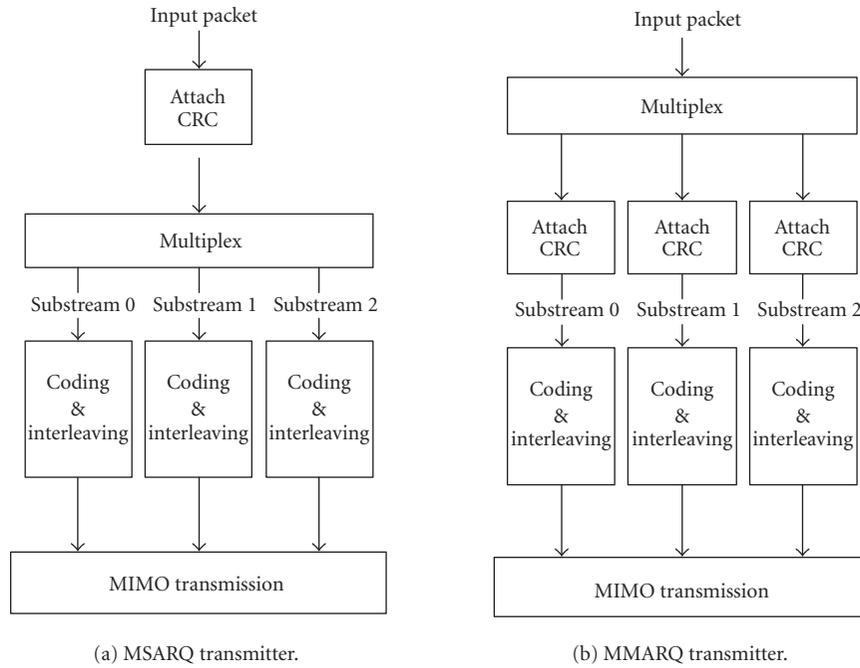
FIGURE 3: Transmitter structures of MSARQ and MMARQ.
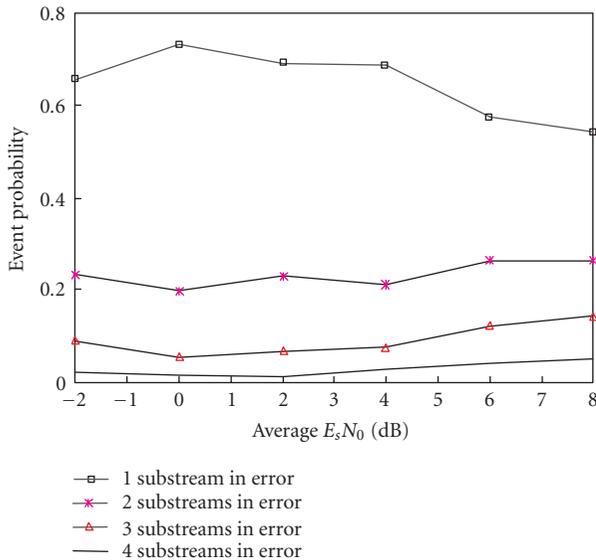
FIGURE 4: Probability distribution of the number of corrupted substreams in an error event with 4 transmit and 4 receive uncorrelated antennas and frequency-flat fading.

back overhead scales with the number of transmit antennas. When that number is large, grouping the transmit antennas and assigning a single ARQ process to each group can reduce the signaling overhead.

Next, using per-antenna encoders with successive decoding and cancellation at the receiver as an example, we describe the receiving procedures for both MMARQ and MSARQ. The receiver decodes the transmitted substreams sequentially following a certain order, which can be optimized to achieve the best throughput performance. The first substream is decoded from the overall aggregate received signal $\mathbf{Y}(t)$. The information data $S_0(t)$, extracted from substream 0, is then reencoded, interleaved, and modulated to construct a replica of the transmitted substream. This replica, combined with the channel response, that is, $F(S_0(t), \mathbf{H}(t))$, is then subtracted from $\mathbf{Y}(t)$ so that the interference contribution of this substream to the others is removed. This procedure is the so-called interference cancellation. The same process is then applied to the remaining substreams, which are thus successively extracted.

For MMARQ, the interference cancellation and HARQ packet combining procedures can be blended advantageously. In that case, the receiver would decode a substream and use its associated CRC to validate the content. If this substream carries a retransmission packet and contains uncorrectable errors, the soft symbols of the packet would be combined with those of the previous transmission(s) to extract the information data. The receiver would then perform interference cancellation to remove the interference due to this substream. Interference cancellation is performed regardless of the results of the CRC validation; therefore, all the subsequent substreams can be decoded without waiting for the retransmission of the current substream. However, the reliability of the decoded data is much higher after HARQ packet combining and, thus, using such data to reconstruct the signal replicas for interference cancellation reduces error propagation. The detailed receiver procedure is shown in **Figure 5**.
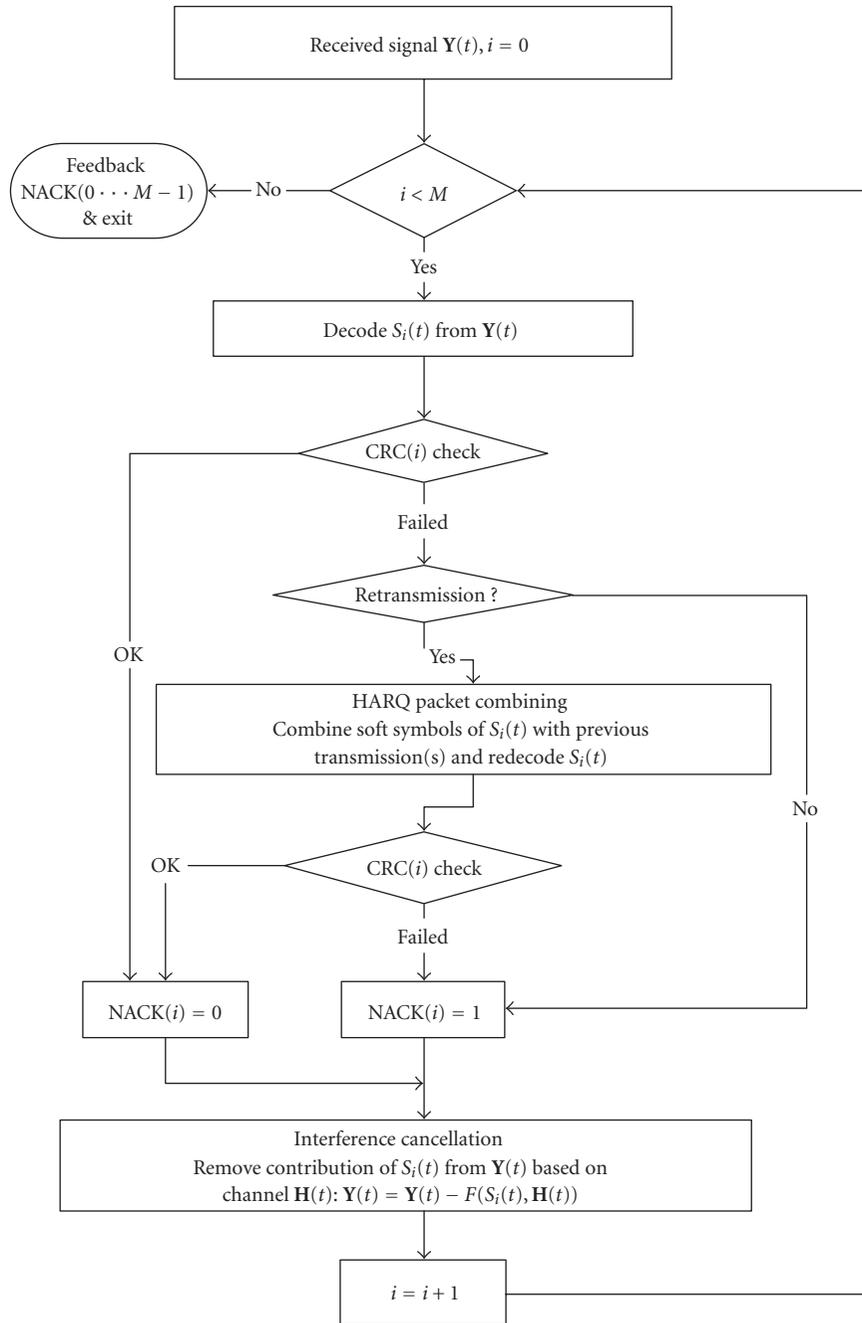
Figure 5: MMARQ receiver flow chart.

In contrast, it is not so easy to combine HARQ with interference cancellation when MSARQ is employed. As illustrated in Figure 6, MSARQ separates HARQ packet combining from interference cancellation. The receiver performs packet decoding and interference cancellation to extract the substreams and then combines those substreams into a compound packet. In this case, decoding errors at each substream could propagate to the substreams that are decoded afterwards. Such error propagation could severely degrade the performance. Another alternative would be to recancel interference on the HARQ combined signal upon a CRC failure. This procedure is shown in Figure 7, wherein interference cancellation is conducted twice. We refer to it as MSARQ IC. The resulting hardware design, however, could be problematic, as the receiver would need to quickly feedback the NACK indicator to the transmitter.
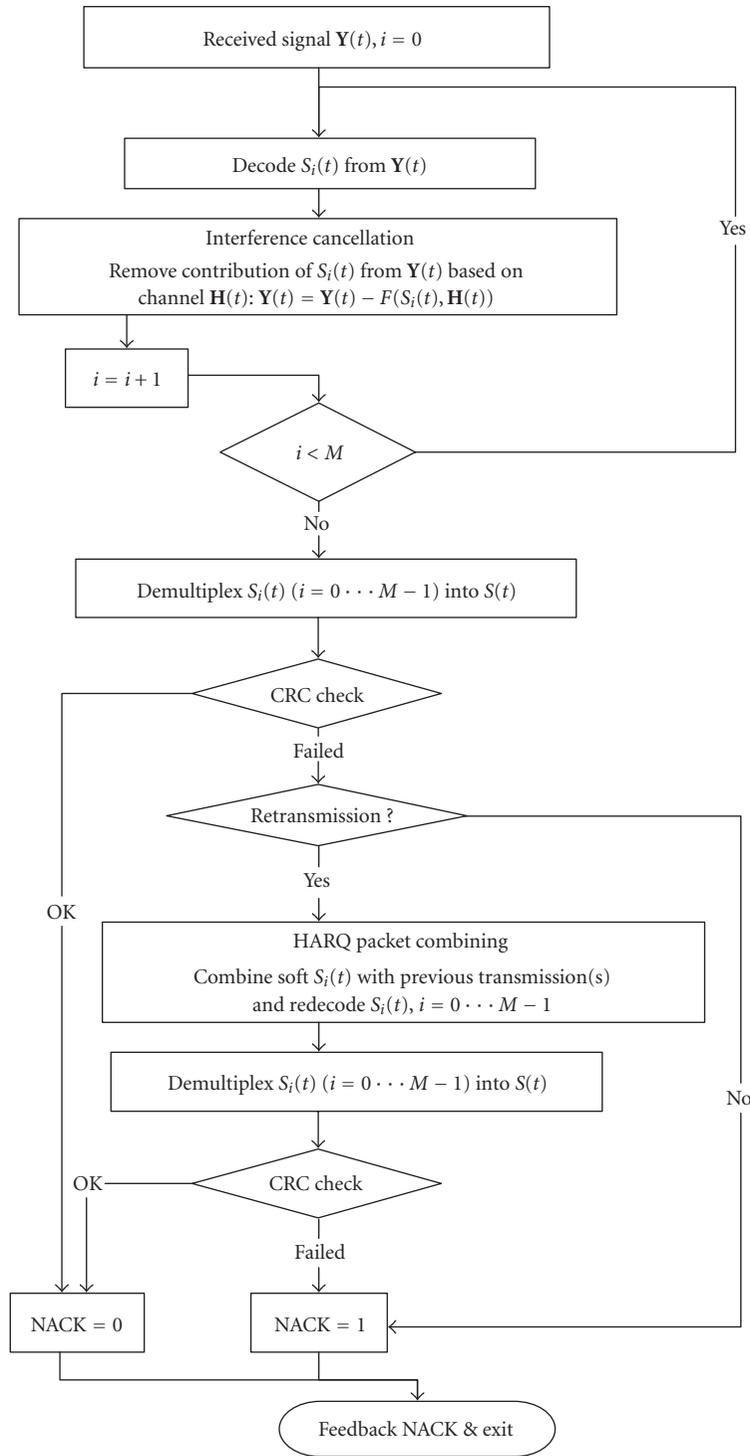
FIGURE 6: MSARQ receiver flow chart type I.

4. COMPARISON OF MSARQ AND MMARQ

In this section, we compare the performance of MSARQ and MMARQ in the context of UMTS HSDPA [25]. The most prominent features of HSDPA, which is specifically geared towards delay-tolerant data, are as follows.

(1) A fraction of the power and code space available at the base station is allocated to HSDPA while the rest is assigned to pilots, overhead channels, and voice traffic.

(2) HSDPA users are time-multiplexed in short frames. A scheduler at the MAC layer determines the user to be served at each frame. Each scheduling interval or
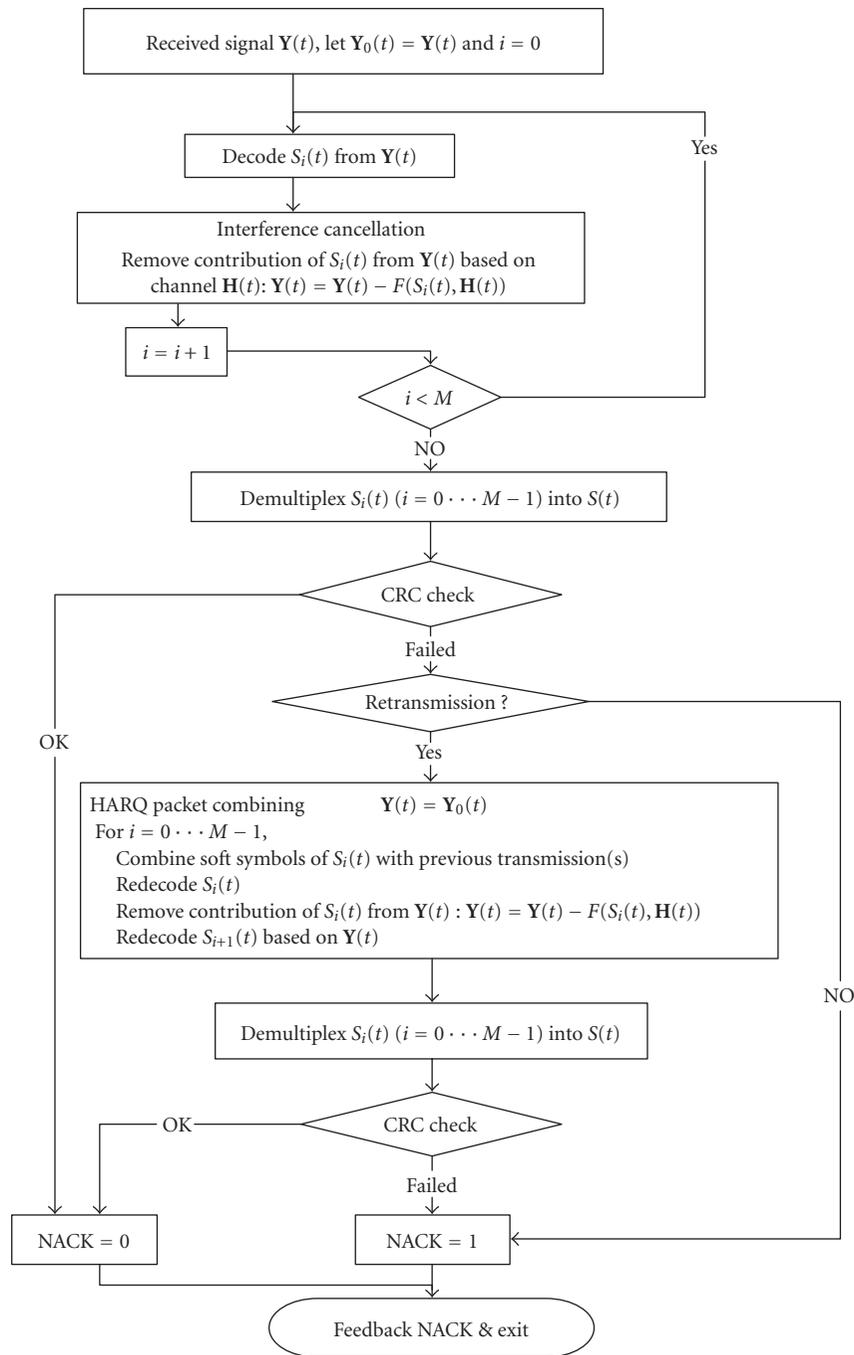
FIGURE 7: MSARQ IC receiver flow chart.

frame lasts 2 milliseconds. We assume that the entire HSDPA code space (10 codes in this paper) and transmit power are assigned to the scheduled user. That is, the base station transmits to only one user in each frame using 10 codes and full power. The transmit signal consists of a superposition of such 10 orthogonal codes.

(3) The Node B (or base station) MAC determines the

transmission rate for the user being served, based on the CQI.

(4) The HARQ functionality resides between the Node B and the mobile terminal to permit soft combining and fast NACK/ACK feedback.

We have developed a simulation tool that captures the dynamic processes in a radio network. The simulated radio network consists of a base station (Node B) and multiple user

terminals. The Node B possesses the following functionalities.

(a) MAC$_{HSDPA}$. It performs scheduling, MCS selection, and HARQ, based on the CQI feedback and the NACK/ACK signaling from each terminal.

(i) *Scheduler*. System performance depends heavily on the scheduling algorithm. For the purpose of this work, we limit ourselves to a round-robin scheduler, which exhibits maximum fairness across users. Additionally, with such scheduler, it is easy to quantify the system-level performance from the single-user performance.

(ii) *MCS selection*. The MCS at each transmit antenna is separately controlled through CQI feedback from the receiver [23, 24].

(iii) *HARQ*. The downlink HARQ operates asynchronously, that is, the retransmissions can take place anytime after the Node B receives a NACK/ACK. The scheduler determines the exact time. To compensate for the NACK/ACK feedback delay of 2 frames, each HARQ entity operates in terms of three stop-and-wait (SAW) processes. This allows HARQ to operate continuously without waiting for a NACK/ACK signal. For MSARQ, all transmit antennas use a single HARQ entity with 3 processes while, for MMARQ, each transmit antenna uses one HARQ entity with 3 processes. Chase combining is used to combine the initial transmission with the retransmissions. The maximum number of retransmissions is 30. If a corrupted packet cannot be recovered after exhausting the maximum number of retransmissions, the packet is discarded and the associated loss should be recovered by higher layer error control mechanism.

(b) *PHY*. The physical layer simulation consists of a sequence of events such as transmission and reception of signals, signal-to-interference-and-noise ratio (SINR) evaluation, and channel estimation. It employs a bandwidth of 5 MHz with 3.33-milliseconds frames. We assume that the uplink channel operates at a rate of 64 kbps. At the terminal, the substreams radiated by the various transmit antennas are decoded according to a fixed order. The MCS of each such substream is selected based on its detected SINR at the receiver and it is then fed back as a CQI message. Some additional premises are summarized below:

(i) fading is Rayleigh-distributed and frequency-flat and the channel is either perfectly known at the receiver or modeled by adding simulated estimation noise onto the actual channel;

(ii) pedestrian speed (3 Km/hr);

(iii) 70% of transmit power dedicated to HSDPA;

(iv) 10 out of 16 orthogonal codes dedicated to HSDPA;

(v) 4 uncorrelated transmit and 4 uncorrelated receive antennas;

(vi) 7 MCSs employing turbo codes with varying rates and symbol repetition [4]: QPSK rate 1/4 repeated 4 times, QPSK rate 1/4 repeated 2 times, QPSK rate 1/4, QPSK rate 1/2, QPSK rate 3/4, 16-QAM rate 1/2, and 16-QAM rate 3/4.

The probability of each substream being detected erroneously is given by a frame error rate (FER) versus instantaneous SINR curve for each MCS. For the above MCS schemes, these curves are displayed in Figure 8.

The ultimate performance measure is the single-user throughput, defined as the ratio between the number of information bits correctly received by a user and the time that the channel is allocated to that user:

$$\text{throughput} = \frac{\text{total good bits}}{(\text{total frames with transmissions}) \cdot \text{frame duration}}. \tag{1}$$

Notice that the throughput represents the peak net throughput that can be delivered to a user.

It should be pointed out that the throughput depends on the mapping between the detected SINR and the selected MCS per antenna. Such mapping is adjusted in order to maximize the throughput while maintaining some target FER measured prior to HARQ operation. When this target FER is small (less than 5%), the probability of retransmission is low and there is no large gain with any kind of ARQ. As the target FER increases, the probability of retransmissions grows and there is a considerable gain with MMARQ. Hence, we optimize the FER to maximize the throughput.

4.1. Performance with perfect channel estimation and feedback

Our initial simulations assume perfect channel estimation and error-free uplink feedback. We first examine the advantage of combining HARQ with interference cancellation by comparing the compound packet error performance of MSARQ and MSARQ IC. Separating interference cancellation from HARQ combining fails to eliminate the interference from any corrupted substream even if the substream is later fully recovered through HARQ packet combining. Such inefficiency results in a higher compound packet error rate (Figure 9). To quantify the advantage of per-antenna HARQ in MMARQ, the throughput performances of MMARQ, MSARQ, and MSARQ IC are compared in Figure 10. We observe that MMARQ achieves 10%–20% improvement over MSARQ IC and 26%–40% over MSARQ. Thus, the contributions of combined operation and multiple ARQ structures are roughly equal. The ergodic Shannon capacities for open-loop single-transmit single-receive and 4-transmit 4-receive configurations are also shown in the same figure as references.

In the above example, MSARQ, MSARQ IC, and MMARQ use the same MCS/SINR settings, which maximize the throughput for MMARQ but not necessarily for MSARQ and MSARQ IC. Through additional simulations, we find that the optimal MCS/SINR settings for MSARQ and MSARQ IC yield a compound FER of 8%–10%, while the optimal FER for MMARQ is around 15%–20%. The
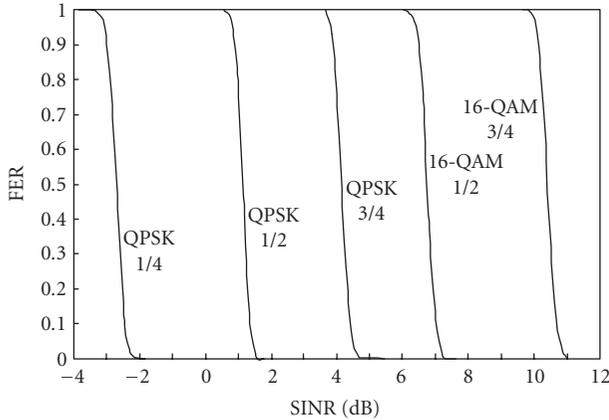
FIGURE 8: Frame error rate (FER) versus SINR for a single transmit and a single receive antenna.
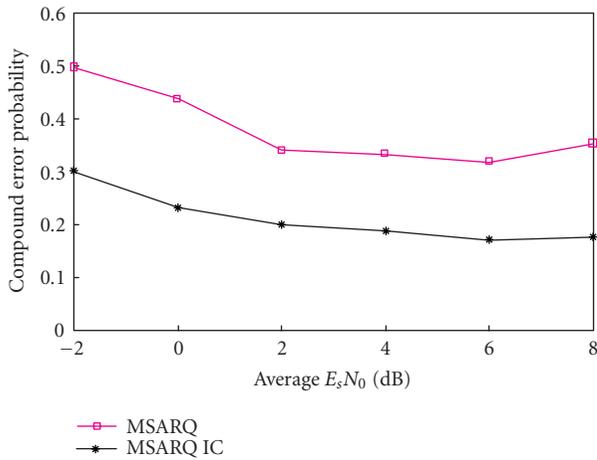
FIGURE 9: Compound packet error rate of MSARQ and MSARQ IC.
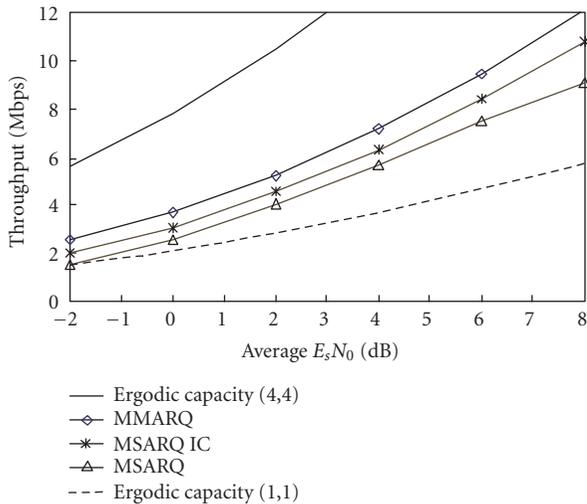
FIGURE 10: Throughputs of MMARQ, MSARQ, and MSARQ IC with interference cancellation in ideal conditions.

corresponding individual substream error rates are 2%–5% and 8%–18%, respectively. In practice, it is quite difficult to guarantee a substream error rate of 5% or less. Therefore, the optimal throughput of MMARQ would be easier to achieve in a realistic environment. Nevertheless, the optimized throughputs are shown in Figure 11, where the improvement of MMARQ drops to around 10% with respect to MSARQ IC and 20% with respect to MSARQ. By operating at a low packet error rate, channel coding and packet combining can eliminate most channel errors. As such, the throughput gap between MSARQ and MSARQ IC also diminishes.

4.2. Performance with imperfect channel estimation and feedback

Next, we examine the performance of MSARQ, MSARQ IC, and MMARQ in more realistic conditions, with imperfect channel estimation and imperfect uplink feedback. The main sources of imperfection are limited pilot power, finite channel coherence time, and feedback delay. We model these nonidealities by adding noise to the SINR, that is,

$$\hat{\gamma}_{\mathrm{dB}} = \gamma_{\mathrm{dB}} + N(0, \sigma_a^2), \tag{2}$$

where $\hat{\gamma}_{\mathrm{dB}}$ represents the SINR in dB as estimated by the receiver, γ_{dB} represents the actual SINR in dB, and $N(0, \sigma_a^2)$ represents Gaussian noise with variance σ_a^2. The estimation error not only impacts the MIMO signal detection and decoding process, but also impacts the MCS selected for each transmit antenna. In addition, the uplink feedback channel also encounters a uniformly distributed binary error rate of 6%, which could corrupt the CQI and the NACK/ACK indication(s). Figure 12 illustrates the throughput performance of MMARQ, MSARQ, and MSARQ IC for $\sigma_a^2 = 1.5$ dB. The performance degradations range from 10% to 18% for MMARQ, 17% to 32% for MSARQ, and 16% to 24% for MSARQ IC. Relatively, MMARQ is less sensitive to channel estimation noise and feedback errors. As the level of uncertainty increases, it becomes more difficult to guarantee a successful transmission without sacrificing packet throughput. In this case, it is beneficial to transmit aggressively and use HARQ to recover from channel errors. Overall, MMARQ achieves 30%–45% throughput improvement over MSARQ, while per-antenna ARQ contributes to a 15%–25% throughput improvement.

5. CONCLUSION

We have proposed a new ARQ scheme suitable for any MIMO scheme in which substreams radiated from different antennas are encoded separately. Conventionally, a single ARQ process is applied to each data packet. Upon an error event, all constituent substreams—including those that have already been correctly received—are retransmitted. In contrast, our proposed scheme separates the ARQ processes for the substreams. We have quantified the gains of the new scheme within the context of UMTS high-speed downlink data access. We first considered ideal conditions with perfect
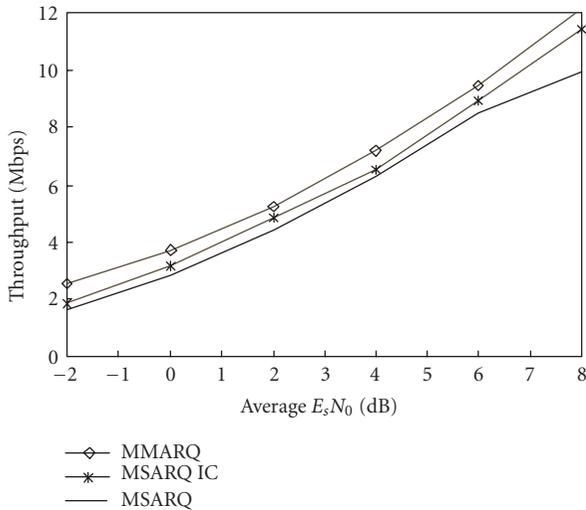
FIGURE 11: Throughputs of MMARQ, MSARQ, and MSARQ IC with interference cancellation in ideal conditions using the optimized MCS/SINR settings.
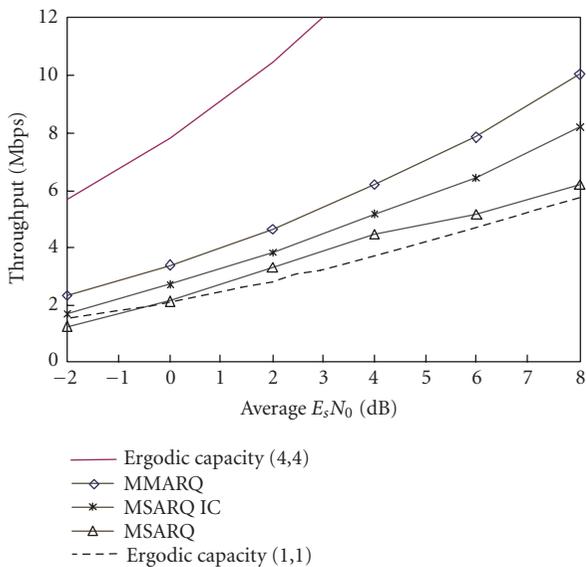
FIGURE 12: Throughput of MMARQ, MSARQ, and MSARQ IC with interference cancellation in realistic conditions (with imperfect channel estimation and imperfect uplink feedback).

channel estimation and error-free uplink feedback, where MMARQ improves the throughput by 25%–40%. We then performed the simulations in more realistic conditions, with imperfect channel estimation and possibly erroneous uplink feedback. Such uncertainty leads to a higher loss rate, and HARQ becomes a major technique for efficient error control and recovery. Hence, MMARQ is even more favorable with the performance improvement increasing to 30%–45% compared with MSARQ. It should be pointed out that the results presented here are based on the premise of frequency-

flat fading and uncorrelated antennas. Frequency-selective fading may modify this conclusion, and this problem is currently under investigation.

Traditionally, the physical layer had been considered the performance bottleneck in wireless systems due to the unpredictable nature of the radio channel. Higher layer issues, such as scheduling, link adaptation, retransmissions, and mobile routing, used to be discussed and treated separately from major physical layer issues. With the convergence of mobile communications and data services, however, there is a growing need for a cross-layer design that facilitates the interaction of multiple protocol layers. In particular, one can couple the design of link layer (i.e., MAC and RLP) with that of the physical layer. The superior performance of MMARQ confirms the benefits of such joint layer design.

REFERENCES

[1] Third Generation Partnership Project, "Physical layer aspects of UTRA high speed downlink packet access (Release 2000)," 3GPP Tech. Rep. 25.848, Technical Specification Group Radio Access Network, 2001.

[2] 3GPP2 C.S00024, "CDMA2000 high rate packet data air-interface specification," 2000.

[3] A. Lozano, F. R. Farrokhi, and R. A. Valenzuela, "Lifting the limits on high speed wireless data access using antenna arrays," *IEEE Communications Magazine*, vol. 39, no. 9, pp. 156–162, 2001.

[4] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[5] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1841–1852, 1999.

[6] K. Balachandran, S. R. Kadaba, and S. Nanda, "Channel quality estimation and rate adaptation for cellular mobile radio," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 7, pp. 1244–1256, 1999.

[7] B. Vucetic, "An adaptive coding scheme for time-varying channels," *IEEE Trans. Communications*, vol. 39, no. 5, pp. 653–663, 1991.

[8] S. Lin, D. Costello, and M. Miller, "Automatic-repeat-request error-control schemes," *IEEE Communications Magazine*, vol. 22, no. 12, pp. 5–17, 1984.

[9] D. J. Costello Jr., J. Hagenauer, H. Imai, and S. B. Wicker, "Applications of error-control coding," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2531–2560, 1998.

[10] Third Generation Partnership Project, "Hybrid ARQ methods for FDD in Release 2000," TSG-R WG1 document, TS-GRl#13(00)0637, May 2000.

[11] G. Benelli, "An ARQ scheme with memory and soft error detectors," *IEEE Trans. Communications*, vol. 33, no. 3, pp. 285–288, 1985.

[12] D. Chase, "Code combining–a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," *IEEE Trans. Communications*, vol. 33, no. 5, pp. 385–393, 1985.

[13] D. N. Rowitch and L. B. Milstein, "On the performance of hybrid FEC/ARQ systems using rate compatible punctured turbo (RCPT) codes," *IEEE Trans. Communications*, vol. 48, no. 6, pp. 948–959, 2000.

[14] R. Love, B. Classon, A. Ghosh, and M. Cudak, "Incremental redundancy for evolutions of 3G CDMA systems," in *Proc. IEEE 55th Vehicular Technology Conference, 2002. VTC Spring 2002*, vol. 1, pp. 454–458, Birmingham, Ala, USA, May 2002.

[15] A. Das, F. Khan, A. Sampath, and H. J. Su, "Performance of hybrid ARQ for high speed downlink packet access in UMTS," in *Proc. IEEE VTS 54th Vehicular Technology Conference, 2001. VTC 2001 Fall*, vol. 4, pp. 2133–2137, Atlantic City, NJ, USA, 2001.

[16] P. Frenger, S. Parkvall, and E. Dahlman, "Performance comparison of HARQ with chase combining and incremental redundancy for HSDPA," in *Proc. IEEE VTS 54th Vehicular Technology Conference, 2001. VTC 2001 Fall*, vol. 3, pp. 1829–1833, Atlantic City, NJ, USA, 2001.

[17] M. Andrews, K. Kumaran, A. Ramanan, K. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150–154, 2001.

[18] H. Zheng and H. Viswanathan, "Optimizing the ARQ performance in downlink packet data systems with scheduling," in *Proc. IEEE VTS Vehicular Technology Conference, 2003. VTC 2003 Fall*, Orlando, Fla, USA, October 2003.

[19] H. Zheng, A. Lozano, and M. Haleem, "Multiple ARQ processes for MIMO systems," in *Proc. of IEEE 13th Symposium on Personal, Indoor and Mobile Radio Communication (PIMRC '02)*, Lisbon, Portugal, September 2002.

[20] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[21] D. Chizhik, F. R. Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Communications Letters*, vol. 4, no. 11, pp. 337–339, 2000.

[22] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[23] A. Lozano, "Capacity-approaching rate function for layered multiantenna architectures," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 616–620, 2003.

[24] S. T. Chung, A. Lozano, and H. C. Huang, "Approaching eigenmode BLAST channel capacity using V-BLAST with rate and power feedback," in *Proc. IEEE VTS 54th Vehicular Technology Conference, 2001. VTC 2001 Fall*, vol. 2, pp. 915–919, Atlantic City, NJ, USA, October 2001.

[25] Third Generation Partnership Project, "UTRA high-speed downlink packet access (Release 4)," 3GPP Tech. Rep. 25.950, Technical Specification Group Radio Access Network, March 2001.

Haitao Zheng received her B.S. degree with highest honor in electrical engineering from Xian Jiaotong University, China, in 1995, the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD, in 1998 and 1999, respectively. From 1995 to 1998, she was an Institute for System Research Fellow at University of Maryland. She received the 1998-1999 George Harhalakis Outstanding Systems Engineering Graduate Student Award in recognition of outstanding contributions in cross-disciplinary research from the University of Maryland. Since August 1999, she is with Wireless Research Laboratory, Bell Labs, Lucent Technologies, Holmdel, NJ. Her research interests include wireless communications and networking, multimedia communications, and signal processing. Recently, she received the Bell Laboratories 2002 Presidents Gold Award in recognition of outstanding level of innovation, technical excellence, and business impact. She currently serves as the TPC Member of IEEE Multimedia Signal Processing Technical Committee, TPC of ICME 2003, TPC of Globecom 2003, Guest Editor of the IEEE JSAC Special Issue on Advanced Mobility Management and QoS Protocols for Wireless Internet, and Guest Editor of the EURASIP Journal on Applied Signal Processing Special Issue on Cross Layer Design for Communications and Signal Processing. She has served as the TPC Member of ICME 2002, ICC 2003, ICASSP 2002, and so forth.

Angel Lozano was born in Manresa, Spain, in 1968. He received his Engineer degree in telecommunications (with honors) from the Polytechnical University of Catalonia, Barcelona, Spain, in 1992 and his M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, Calif, in 1994 and 1998, respectively. Between 1996 and 1998, he worked for Pacific Communication Sciences Inc. and for Conexant Systems in San Diego, Calif. Since January 1999, he has been with Bell Laboratories, Lucent Technologies, Holmdel, NJ. Since October 1999, Dr. Lozano has served as Associate Editor for IEEE Transactions on Communications. He holds 6 patents.

Mohamed Haleem received his B.S. Eng. degree from the Department of Electrical and Electronic Engineering, University of Peradeniya, Sri Lanka in 1990, the M.Phil. degree from the Department of Electrical & Electronic Engineering, Hong Kong University of Science & Technology, in 1995, and is currently working toward his Ph.D. degree in electrical engineering at Stevens Institute of Technology, Hoboken, NJ. He was with the academic staff of the Department of Electrical & Electronic Engineering, University of Peradeniya, from 1990 to 1993, and has been with the Wireless Communications Research Department, Bell Laboratories, Lucent Technologies at Crawford Hill, Holmdel, NJ, from 1996 to 2002. His research interests include dynamic resource assignment, channel adaptive transmission techniques for wireless communication systems, multiple antenna communication systems, and application of stochastic dynamic programming and optimization techniques to wireless communication systems.