Biased Adaptive Modulation/Coding to Provide VoIP QoS over HSDPA

Mikio Iwamura, Vasilis Friderikos, Lin Wang and A. Hamid Aghvami

Centre for Telecommunications Research,
King's College London,
26-29 Drury Lane, London WC2B 5RL, United Kingdom,
mikio.iwamura@kcl.ac.uk

Abstract - The High-speed Downlink Packet Access (HSDPA) of the Universal Mobile Telecommunication System (UMTS) provides a flexible and efficient packet transmission platform, and has been attracting operators to integrate a diverging variety of services. We studied the possibilities of delivering conversational speech traffic in the form of voice over IP (VoIP), over HSDPA. To provide the VoIP QoS typified by severe delay requirements, the adaptive modulation/coding of HSDPA is biased to achieve lower error rates, thereby reducing the number of hybrid-ARQ retransmissions and the transfer delay. Each radio frame is transmitted either with a higher power or using a lower rate modulation/coding scheme, depending on the available power at NodeB. Simulation results show that the proposed scheme effectively improves VoIP QoS consuming a trifle of extra power, and substantially increases the system capacity. The results show that VoIP deliveries over HSDPA is feasible.

Keywords - HSDPA, packet transfer delay, quality of service, system capacity, voice over IP

I. Introduction

The High-speed Downlink Packet Access (HSDPA) [1] of the Universal Mobile Telecommunication System (UMTS) provides efficient packet transmissions over the cellular downlink. With features such as adaptive modulation/coding (AMC), hybrid-ARQ (HARQ), and fast scheduling, HSDPA has been designed to provide a peak user data rate of over 10 Mb/s. In HSDPA all users under the same cell share a set of code channels, i.e., High-speed Physical Downlink Shared Channels (HS-PDSCHs), in a TDM fashion. The fast scheduler at NodeB selects the user to transmit data on the next radio frame, considering the channel quality indicator (CQI) fed back from the user terminals. The exploitation of the fast fading channels produce an effect known as multiuser diversity, which significantly contributes to the spectral efficiency of the HSDPA.

Although HSDPA was originally intended for best effort services, the flexibility and spectral efficiency have been attracting operators to integrate various services into HSDPA. Previous studies [2]-[5] have shown that HSDPA is indeed capable of conveying streaming services. However, since quality of service (QoS) measures such as the user throughput and packet transfer delay are highly dependent on the traffic load and channel variations, guaranteeing a certain bandwidth or transfer delay is difficult, especially when the service is conversational and stringent on delay.

The current third generation systems carry speech traffic on circuit switched dedicated channels (DCHs), that offer robust radio links with transmission power control to resolve fast fading and soft handover to provide continuity of service. On the Internet speech traffic is increasingly carried by IP because of the scalability and low cost. The increasing volume of the packet data traffic in the mobile environment suggests an inevitable path towards packet access systems such as HSDPA. Hence, integration of the speech service into HSDPA is benefitial for affinitive interoperability between cellular systems and the Internet, and to unify the radio network management in UMTS.

Our contribution is to clarify the possibilities of accomodating conversational speech traffic, in the form of voice over IP (VoIP), in HSDPA. Assuming adaptive multi-rate (AMR) [6] 12.2 kb/s speech over HSDPA/IP/UDP/RTP suite, we evaluated the packet transfer delay using a dynamic system level simulator. Since minimizing the packet delay is crucial to provide VoIP QoS, we biased the AMC to work at lower error rates, thereby reducing the number of HARQ retransmissions and the subsequent delay. This consumes an extra fraction of the limited total power, but the reduced number of HARQ retransmissions shall mitigate the overhead to a certain extent. Our results clarify that the overall system capacity is substantially increased by imposing a larger power on each modulation/coding scheme (MCS).

II. VoIP over HSDPA

A. VoIP over UMTS

The standard codec for speech in UMTS is the adaptive multi-rate (AMR) [6] which has a frame size of 20 ms and produces a rate of 4.75 kb/s to 12.2 kb/s. As internetworking between different cellular, wireless, and fixed networks pervade, AMR deliveries over IP networks are getting indispensable. To ratify such implementations, RTP payload formats for AMR has been proposed in [7].

Figure 1 (a) illustrates a basic session scenario where a mobile terminal (MT) that belongs to a non-IP UMTS terrestrial radio access network (UTRAN) communicates to a PSTN terminal over an IP backbone. The voice data is carried over IP between the VoIP gateway (GW) and the UTRAN GW. On the downlink the PSTN codec (G.711) is translated into the adaptive multi-rate (AMR) codec at the VoIP GW, and the AMR frames are packetized into RTP packets. The RTP packets are delivered through the IP backbone and decomposed at the UTRAN GW. The AMR frames are transferred over the UTRAN and decoded at the MT to playback the original speech. The uplink is vice versa.

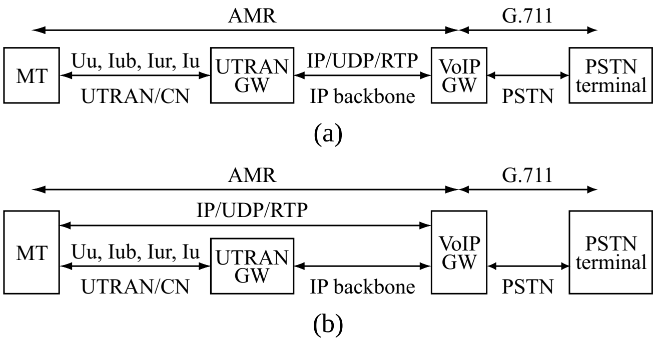

Fig. 1 *Basic session diagrams where AMR speech is carried over an IP backbone.*

An extended scenario is studied in [8] where the RTP is moved to the end MT (Fig. 1(b)), such that the VoIP packets are carried over the UMTS DCH. This allows the VoIP GW and MT to tailor the codec and RTP directly to fit the various channels crossed by the voice packet flow, hence reducing the need to translate QoS classes at a transit node. However, the erroneous radio channel and the error protection/recovery mechanism of the UMTS DCH add a considerable delay and stress the VoIP QoS. The numerical results of [8] have shown that by choosing an appropriate voice packet size and dejittering buffer, VoIP over the UMTS DCH reconciles the stringent delay and packet loss requirements.

Our interest is to further extend this scenario and consider VoIP over HSDPA, which yields a protocol stack depicted in Fig. 2. The voice data that arrive to the VoIP gateway via the PSTN are translated into a low rate codec such as AMR and delivered through the IP backbone to the radio network controller (RNC). At the RNC VoIP packets are segmented into RLC packets. These packets are carried as a single MAC-d flow by the Iub/Iur frame protocol (FP). The FP conducts a flow control over Iub/Iur. At NodeB the MAC-hs schedules the MAC-d PDUs to transmit on the shared physical channel (HS-PDSCH) and conducts AMC and HARQ. Since the residual error rate after the HARQ is very low, the ARQ functionality of the RLC is redundant and likely to prove futile. Thus, the unacknowledge mode (UM) shall be used on the RLC.

If we consider an AMR 12.2 kb/s codec with 20 ms block size, we have an octet aligned IP PDU of 31 bytes. Since each RTP (using IPv4) packet accompanies a 40 byte header (20 byte IP + 8 byte UDP + 12 byte RTP), the overhead consumes a considerable fraction of the total radio bandwidth without header compression. To reduce the overhead, the packet data convergence protocol (PDCP) is used to compress the IP header.

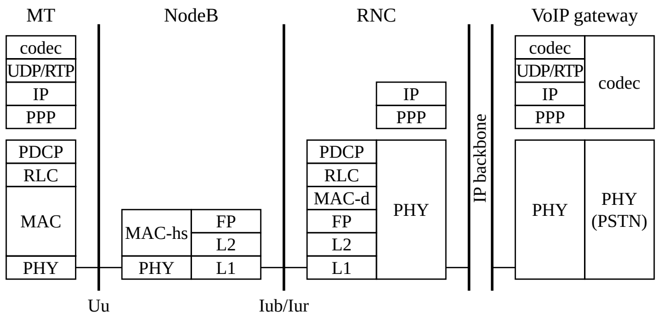

Fig. 2 *Protocol stack for VoIP over HSDPA.*

In general the corresponding terminal is not always a PSTN terminal, but may also be an IP terminal such as a PC. A more likely case would be another MT connected to a UTRAN. Nevertheless, the source does not affect the protocol stack between the RNC and MT in Fig. 2.

B. HSDPA scheduler for VoIP

The MAC-hs schedules the MAC-d PDUs as well as HARQ retransmissions to serve on the HS-PDSCH. The channel quality indicator (CQI) fed back from the mobile terminals may be utilized in the scheduling process. Various schedulers have been studied in the literature, mostly under the best effort discipline to exploit the multiuser diversity, with attemps to add a degree of fairness. A scheduler that chooses the user with the best channel quality (known as a MaxC/I scheduler) maximizes the system throughput, however, with a major deficiency on fairness. A proportionally fair (PF) scheduler [9] exhibits a comparable system throughput with much more tolerable fairness [10], [11].

A number of modifications have been proposed to fit the PF scheduler for streaming services. The main aim of these attemps are to guarantee a lower delay jitter, so that a concise dejittering buffer enables the streaming QoS. In [4] a simple but effective method to calculate the priority metric is presented, in which the head of line delay is multiplied to the PF priority metric. Simulation studies in [5] have shown that the delay sensitive PF improves the streaming performance over HSDPA. A similar discipline may be applied to schedule conversational traffic.

The HS-PDSCH, on which the user data is carried through the air, has a frame size of 2 ms. Hence without code multiplexing, only ten frames are available within a 20 ms AMR block interval. This implies that the HSDPA accomodates a maximum of mere ten flows (assuming continuous conversations without a voice activity detection). This number is further decreased by the HARQ retransmissions. We must also budget for call blocking. Since a single VoIP flow requires only a fraction of the total HSDPA bandwidth, code multiplexing is essential to serve more VoIP flows. A water filling strategy may be applied to the multiuser scheduler to exploit the available power and codes on each frame.

C. Biased AMC to reduce HARQ retransmissions

Since conversational services scarcely tolerate delay, we attempt to reduce the number of HARQ retransmissions at a cost of extra power. According to [1] the AMC is performed such that the BLER is expected to be around 0.1. Consequently, about ten percent of the HS-PDSCH frames require at least one HARQ retransmission, hence adding a considerable delay. Instead, we transmit each MCS with a larger power to lower the error rate, provided the power is available at the NodeB. If the power is unavailable, we simply lower the MCS by an equivalent amount.

Figure 3 illustrates the MCS selection under this rule. The MCS levels of HSDPA are designed such that the received symbol energy to interference power density ratio (E_s/I_0) to achieve a BLER of 0.1 is about 1 dB larger than the next lower MCS level. In Fig. 3 BLER curves for five adjacent MCS levels are shown. Assuming the MCS that corresponds to the reported CQI is level three, the point at "O" represents the original operative point. If we bias the power by 1 dB, the MCS level three is still used with, however, 1 dB higher power, provided the power is available. The working point is thus moved to "A," yielding a lower BLER. If the total power runs short, the MCS is degraded to level two, thereby operating at "B." A bias of 2 dB shifts the point to "C" if the power is available, and to either "D" or "E" if not, depending on the available power.

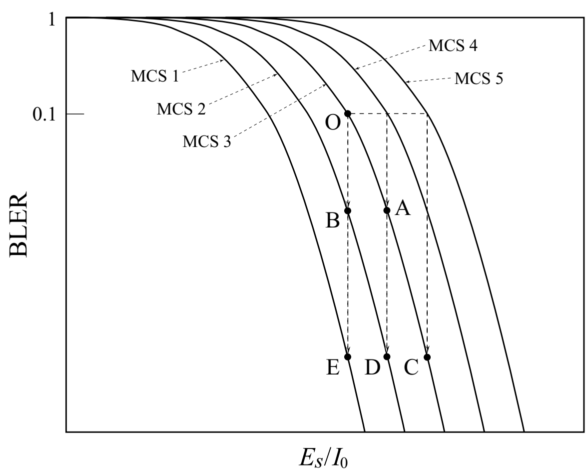

Fig. 3 *Biased MCS selection schema.*

III. System evaluation

A. System model

To evaluate VoIP over HSDPA, we carried out dynamic system level simulations. An original simulator was developed in C++. Poisson arrivals of calls were assumed, with exponentially distributed call lengths having a 90 s mean. Voice activity was also assumed with exponentially distributed spurt and gap durations having a mean of 1 s and 1.5 s, respectively. We assumed the AMR 12.2 kb/s codec, considering that the corresponding source would often be a mobile terminal transmitting via the UMTS uplink, and to impose a comparative QoS to the DCH. Although the standard AMR block size is 20 ms, we also considered 10 ms and 30 ms to investigate other possibilities. This yields octet aligned RTP packet sizes of 18, 33, and 48 bytes, corresponding to 10, 20, and 30 ms, respectively, including a 2 byte compressed IP header per packet, assuming an RTP packet carries one AMR frame. (This is equivalent to having multiple 10 ms voice codec frames per RTP packet.) We assumed a constant delay from the source to NodeB for simplicity, thereby allowing the RTP packets of a single spurt to arrive at fixed intervals given by the AMR block size. We omitted any control packets other than the actual voice packets.

The main system parameters are summarized in Table 1. A regular hexagonal cell layout was assumed with the so-called wrap around technique applied to avoid the boundary effect. The radio propagation was simulated as a concatenation of the Hata loss, lognormal shadowing (std. deviation = 8 dB, inter-site correlation = 0.5), and 3-path Rayleigh fading with the maximum Doppler frequency of 5 Hz. The AMC was performed using 25 MCS levels from 68.5 kb/s to 7.2 Mb/s (Category 8 in the 3GPP specification [12]), with a CQI feedback delay of 2 HS-PDSCH frames. Moreover, the Chase combining HARQ [13] was modeled with 6-channel stop-and-wait (SAW) process per user.

Table 1 *System configuration.*

Cell layout	19 cells, 3 sectors, wrap around
Site separation	1 km
Chip rate	3.84 Mc/s
Path loss exponent	3.44
Shadowing	spacially correlated lognormal (8 dB)
Multipath fading	3-path Rayleigh (5 Hz)
Handover hysteresis	3 dB
Receiver noise figure	9 dB
Total Tx power	20 W
P-CPICH Tx power	2 W
HS-PDSCH Tx power	16 W (max)
HS-PDSCH spreading factor	16
Number of HS-PDSCH codes	10
Code multiplexing	Max 4 users
AMC	25 MCS levels (68.5 kb/s to 7.2 Mb/s)
HARQ	6-channel SAW, Chase combining
CQI feedback delay	2 HS-PDSCH frames
Scheduling discipline	PF

In a VoIP session, a number of factors such as coding delay, transmission delay, and dejittering delay at various stages constitute the overall mouth-to-ear delay. Moreover, some packets may be lost on the way, due to a buffer overflow or timeout. Since our aim is to investigate the impact of HSDPA on VoIP, we focused on the transfer delay and packet losses on the HSDPA radio link. The delay was measured per RTP packet, after reordering to fix the shuffle caused by the 6-SAW HARQ in the receiver. An RTP packet that took longer than 100 ms to transfer from the NodeB to the MT was considered as a packet loss. Moreover, any RTP packet that has not been fully or partly transmitted for 100 ms after arrival (and therefore not stored in one of the SAW processes) was dropped at the NodeB. If a VoIP call experiences a packet loss rate (including dropping) of over 0.01, the call was considered as an "outage." If the HARQ is unable to recover the data after 20 retransmissions, the call was dropped and also counted as an outage. (Remind that the outage is not the probability of the total power exhaustion as often seen in the literature, but the probability of users who experience unsatisfactory QoS.)

B. Simulation results

We first examine the impact of the AMR block size on the outage probability (Fig. 4). As the traffic load is increased, the outage probability increases due to longer queues, with the block size of 10 ms being the worst. This is associated with the lack of radio frames. The block size of 10 ms generates more RTP packets per unit time. These packets arriving to the NodeB in rapid succession increases the probability of each radio frame being transmitted with few data. These lightly loaded frames occupied a considerable number of radio frames, while lengthening the queue and causing outages. Although the 30 ms block size exhibits the lowest outage, a larger coding delay required at the source (or a codec translator) may limit its use. Hence, we focus on 20 ms in the sequel. We also notice that the outage exceeds 0.02 even at a low traffic load of 400 BHCA. This is partly caused by HARQ retransmissions.

If we apply the biased AMC, the outage probability is reduced remarkably (Fig. 5). The proposed AMC provides lower error rates and reduces the number of HARQ retransmissions. This reduced the transfer delay as well as the jitter (Fig. 6) and allowed new packets to transmit with a shorter queuing time. Hence, each packet experienced shorter delays and the outage was improved. An outage less than 0.05 at a load of 2,000 BHCA, with a power offset larger than 2 dB, raises the prospect of VoIP over HSDPA. A drawback is that the outage probability still has a floor at around 0.02. Such outage users are likely to be around the cell edge, and would be in a soft handover state to guarantee QoS if DCHs are used instead. Although the outage floor (associated with handover) remains a problem, the biased AMC sets VoIP over HSDPA a feasible alternative to DCHs.

The proposed AMC achieves lower error rates by transmitting each MCS with a higher power. This imposes an

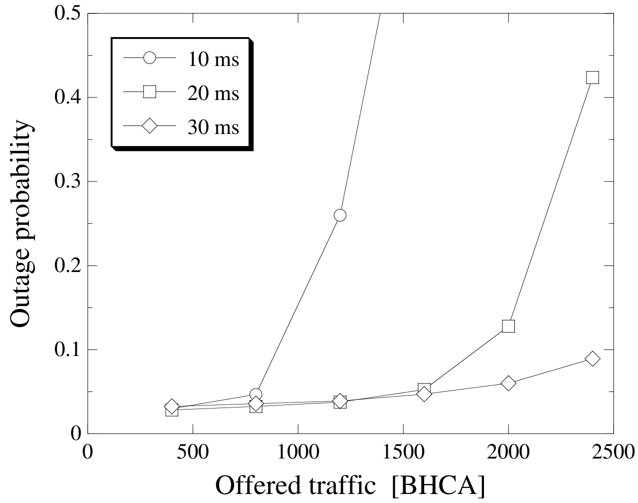

Fig. 4 *Outage probability for various AMR block sizes.*

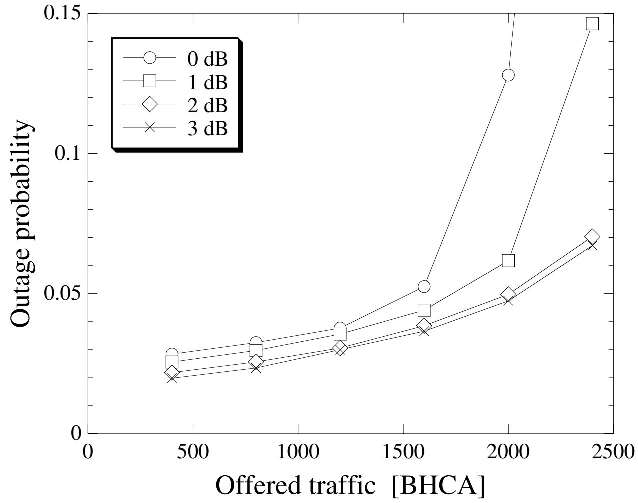

Fig. 5 *Outage probability for various power offsets.*

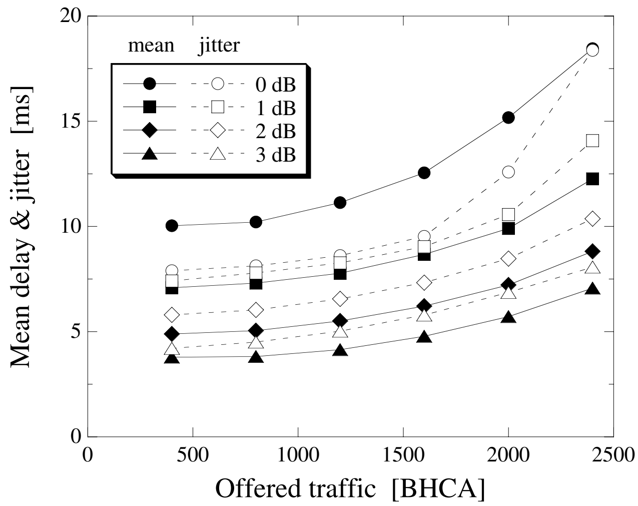

Fig. 6 *Delay statistics comparison.*

extra power consumption on aggregate (Fig. 7). However, an offset of 1 dB does not increase the average total power by 1 dB, but only about 0.25 dB even at a load of 2,000 BHCA. The total power merely increases by 1 dB at 2,000 BHCA when the offset is 3 dB. This is because the reduced number of HARQ retransmissions relieved the aggregate power increase. Therefore, the proposed scheme effectively improves the VoIP QoS at a cost of slightly increased power and interference.

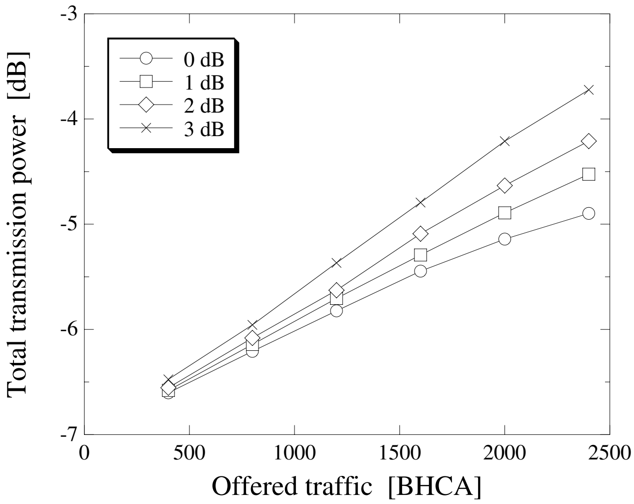

Fig. 7 *Total power consumption of the biased AMC.*

If we define the system capacity (Erlang capacity) as the maximum offered traffic load that guarantees an outage probability of 0.05, we obtain the capacity values in Table 2. Note that the capacity is presented as a relative gain from the 0 dB case. The capacity is substantially increased by setting a power offset in AMC. An offset of 3 dB increases the capacity by more than 30 percent. Therefore, the proposed AMC adjustment is essential to provide VoIP services over HSDPA. However, an offset larger than 3 dB diminishes the gain due to excessive use of power, which limits the number of code multiplexed users and increases interference.

Table 2 *Capacity gain by the biased AMC.*

Power offset [dB]	Relative capacity
0 dB	1.00
1 dB	1.15
2 dB	1.29
3 dB	1.32
4 dB	1.21

IV. CONCLUSIONS

The rapidly increasing volume of the VoIP traffic on the Internet and the high spectral efficiency of the HSDPA suggest a path towards enabling VoIP over HSDPA. We investigated the possibilities of accomodating VoIP traffic in HSDPA. To enable the VoIP QoS typified by the severe delay requirements, we introduced a biased AMC scheme, in which the transmission power is raised on each MCS to achieve lower error rates. This reduces the number of HARQ retransmissions and the transfer delay. Our simulation results showed that the proposed scheme is essential to improve the VoIP QoS over HSDPA, and to increase the system capacity. A power offset of 3 dB increased the system capacity by over 30 percent, consuming about 1 dB of extra power on aggregate. Although a problem remains that about two percent of users experience inferior QoS even at low traffic loads, our results showed that VoIP deliveries over HSDPA is feasible.

ACKNOWLEDGEMENT

This work was performed under partial supports of NTT DoCoMo, Inc. and the project IST-EVEREST (www.everest-ist.upc.es), which is partly funded by the European Community.

REFERENCES

[1] 3GPP, TR 25.848, "Physical layer aspects of UTRA high speed downlink packet access," v4.0.0, March 2001.
[2] T. Moulsley, "Performance of UMTS high speed downlink packet access for data streaming applications," *IEE Proc. 3G2002*, pp. 302-307, London, May 2002.
[3] G. Barriac and J. Holtzman, "Introducing delay sensitivity into the proportional fair algorithm for CDMA downlink scheduling," *IEEE Proc. ISSSTA 2002*, vol. 3, pp. 652-656, Prague, Sept. 2002.
[4] M. Andrews, et al., "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, vol. 39, issue 2, pp. 150-154, Feb. 2001.
[5] P. Ameigeiras, *Packet scheduling and quality of service in HSDPA*, Ph.D. Thesis, Aalborg Univ., Oct. 2003.
[6] 3GPP, TS 26.090, "Adaptive multi-rate (AMR) speech codec, transcoding functions," v6.0.0, Dec. 2004.
[7] IETF, RFC 3267, "Real-time transport protocol (RTP) payload format and file storage format for the adaptive multi-rate (AMR) and adaptive multi-rate wideband (AMR-WB) audio codecs," June 2002.
[8] F. Poppe, D. De Vleeschauwer, and G. H. Petit, "Choosing the UMTS air interface parameters, the voice packet size and the dejittering delay for a voice-over-IP call between a UMTS and a PSTN party," *IEEE Proc. INFOCOM 2001*, pp. 805-814, Anchorage, April 2001.
[9] J. H. Rhee, T. H. Kim, and D. K. Kim, "A wireless fair scheduling algorithm for 1xEV-DO system," *IEEE Proc. VTC Fall 2001*, pp. 743-746, Atlantic City, Oct. 2001.
[10] R. Love, et al., "High speed downlink packet access performance," *IEEE Proc. VTC Spring 2001*, vol. 3, pp. 2234-2238, Rhodes, May 2001.
[11] Y. Ofuji, S. Abeta, and M. Sawahashi, "Comparison of packet scheduling algorithms focusing on user throughput in high speed downlink packet access," *IEICE Trans. Commun.*, vol. E86-B, no. 1, pp. 132-141, Jan. 2003.
[12] 3GPP, TS 25.214 "Physical layer procedures (FDD)," v6.0.0, Dec. 2003.
[13] D. Chase, "Code combining - a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," *IEEE Trans. Commun.*, vol. 33, no. 5, pp. 385-393, May 1985.