Performance of Internet Services over the UMTS Enhanced Uplink

Andreas Mäder, Dirk Staehle, and Christian Gößwein
University of Wuerzburg, Department of Distributed Systems
Am Hubland, D-97074 Würzburg, Germany
{maeder, staehle, goesswein}@informatik.uni-wuerzburg.de

Abstract

The Enhanced Uplink or High Speed Uplink Packet Access (HSUPA) is designed to enable Internet applications in the uplink in UMTS networks. Its specification includes several new features which on the one hand increases throughput and decreases delay for the users, and on the other hand enables a better resource utilization on the network side. In this work, we explain the most prominent new features of the Enhanced Uplink and investigate their impact on the system and end-to-end performance with a special focus on different scheduling strategies. For the numerical results, a full-featured packet-level simulator has been implemented which enables the evaluation of UMTS networks with Rel. 99 users and Enhanced Uplink users in parallel.

1 Introduction

The UMTS Enhanced Uplink or High Speed Uplink Packet Access (HSUPA) is the counterpart to the High Speed Downlink Packet Access introduced with UMTS Rel. 5. The Enhanced Uplink is introduced with UMTS Rel. 6 and therefore completes the 3.5G HSxPA "family" of specifications in the 3GPP world. With Enhanced Uplink, 3GPP reacts to the increasing demand for broadband mobile radio access not only in the downlink, but also in the uplink direction. This demand is created by applications diffusing into the mobile realm like IP-based video conferencing (e.g. with Skype), gaming, or peer-to-peer file-sharing. This trend is amplified by the fact that hardware vendors begin to launch stationary HSxPA modems which are designed as an alternative to classical DSL connections.

In comparison to the Rel. 99 DCH radio bearers, the Enhanced Uplink has advantages from user as well as from operator perspective. The user benefits from an increased maximum throughput which is approximately 5.4 Mbps under laboratory conditions. In reality, it is expected that bit rates higher than 1 Mbps are achievable [11]. These high bit rates are realized by using up to 4 channelization codes with low spreading factors in parallel. A second improvement from user's perspective are reduced packet delays. This is achieved with short transport time intervals (TTI) of 2 ms, and additionally with the implementation of Hybrid ARQ (HARQ) on physical layer. HARQ reduces the residual frame error rate (FER) after Layer 1 processing and thus decreases the number of retransmissions on RLC layer from RNC to UE.

From operator's perspective, the benefits of the Enhanced Uplink are in the better utilization of radio resources. The relocation of the rate control/scheduling mechanism from the RNC to the NodeB enables a much faster reaction on interference fluctuations. Therefore, operators can decrease the safety margin between target cell load and maximum cell load, which ultimately means a higher network capacity. A second aspect is the flexibility of the scheduling mechanism, which allows it to implement various policies e.g. to support high-value services.

In the literature, work concentrates mostly on single aspects of the Enhanced Uplink performance. In [13], some general aspects of radio resource allocation and scheduling are investigated, supplemented with simulation results for full buffer and gaming traffic. In [12], channel-aware and channel-blind scheduling strategies are compared with help of Monte-Carlo simulations. System-level performance for for a limited set of bit rates is presented in [5], and in [10], a combination of time-division and code-division channel-aware scheduling was proposed. An analytical model has been used in [9] for the comparison of different admission control strategies.

In this work, we give an overview of the most prominent technical features of the Enhanced Uplink. Furthermore, we introduce a full-featured simulation model implemented within the OPNET environment and provide some numerical results on the performance of Internet services like video streaming and best-effort traffic. Since OPNET works on packet level, it is especially well suited for the analysis of performance measures like end-to-end-delay or TCP throughput. The model is based on the OPNET UMTS-model, which implements all relevant UMTS protocols like

PDCP, RLC, MAC and UTRAN network functionalities like mobility management and call setup. The UMTS model has been extended with Enhanced Uplink features and has been left unmodified as far as possible, such that a coexistence of DCH users and Enhanced Uplink users is possible. We furthermore propose three scheduling mechanisms: A resource-fair scheduler, a delay-aware scheduler and a channel-aware throughput-scheduler. In the numerical results we show the impact of the schedulers on the performance for video streaming and best-effort traffic.

In the next section, we describe various system aspects of the Enhanced Uplink. In Sec. 3, the schedulers are explained for different service types, which are evaluated in Sec. 4. In Sec. 5, we conclude or work and give an outlook on further topics.

2 System Description

With the Enhanced Uplink, some new transport and signaling channel have been introduced. The most prominent is the Enhanced DCH (E-DCH) and the corresponding physical channel, the E-DPDCH, which carries user data. Additionally, a new uplink control channel, the E-DPCCH has been defined. In the downlink, the E-DCH HARQ Indicator Channel (E-HICH) is used to indicate whether the actual received HARQ transmission has been decoded successfully. Scheduling information in the form of grants are sent over the E-DCH Relative Grant Channel (E-RGCH) and the E-DCH Absolute Grant Channel (E-AGCH). While the former is a dedicated channel which is co-located to a DCH, the latter is a shared channel for all E-DCH users within a sector.

2.1 Physical Layer and MAC layer

In Enhanced Uplink, high data rates are achieved on physical layer by using multi-code transmissions with low spreading factors (SF). The highest configuration uses two SF 2 codes and two SF 4 codes in parallel, which means a raw channel bit rate of 5.76 Mbps. The number of parallel codes and the Transport Block Size (TBS), which corresponds to the number of information bits [2], define together the E-DCH Transport Format Combination (E-TFC) .

On MAC-layer, the MAC-e and MAC-es entities have been introduced. One UTRAN-side, MAC-es only exists in the RNC where it is responsible for macro diversity selection and reordering. An MAC-es PDU (packet data unit) contains one or more MAC-d PDUs. MAC-e is on UTRAN-side is located in the NodeB, where it is responsible for scheduling and HARQ. MAC-e PDUs contain one or more MAC-es PDUs and constitute the final data frame which is then transmitted on physical layer to the UE. On UE-side, both entities exist. Besides the reciprocal tasks for the UTRAN-side, E-FTC selection and restriction takes place here.

Hybrid ARQ combines forward error correction with an automatic-repeat-request protocol. In case of a negative ACK, the UE saves the erroneous frame in a buffer and waits for a retransmission. The retransmission can either be an exact copy of this frame, or a version with a modified coding ratio. The first variant is known as Chase-Combining [4], while the second uses incremental redundancy, since normally the first frame has less redundancy bits than its subsequent versions. For the Enhanced Uplink, both variants are possible which is realized through redundancy flags in the rate matching entity of the coding chain.

Hybrid ARQ is realized as an N-stop-and-wait (N-SAW) protocol. This means that at a specific TTI an HARQ-process sends a layer 1 data frame and then stops and waits for a (negative or positive) acknowledgment, before it sends a new frame or initiates a retransmission. In the next TTI, another HARQ process does the same. The number of HARQ processes depends on the length of the TTI: In case of 10 ms, 4 HARQ processes are used, in case of 2 ms, 8 HARQ processes are used. After N TTIs the same HARQ-process is again selected. It should be noted that first, each HARQ-process has its own queue of MAC-e PDUs and second, a retransmission must use the same E-TFC as the initial transmission.

In the simulation model implemented for this work, an actual value interface (AVI) has been created on base of extensive Monte-Carlo simulations of the physical layer coding chain using a modified version of Valenti's ISCML toolbox for Matlab [1]. The simulation implements the E-DPDCH selection algorithm, the rate 1/3 turbo coder, and the rate matching entity which is responsible for puncturing, repetition and setting the correct redundancy version of the HARQ transmission.

If the UE is in soft handover mode, it selects one NodeB (usually the one with the best link quality) as serving NodeB (S-NodeB). In order to reduce downlink signaling load, only the S-NodeB is allowed to send NACKs. The RNC ensures that duplicates are eliminated and that the MAC-es PDUs are delivered in sequence.

In Rel. 99, the outer loop power control (OLPC) is responsible for adapting the target-SIR γ^* of the DCH in order to meet a certain residual frame error rate. For the Enhanced Uplink, it is proposed to use the number of HARQ retransmissions instead, or more specific, a certain target-RSN (retransmission sequence number). In this work, the OLPC is implemented following [3]:

$$\gamma^*_{\text{new}} = \begin{cases} \gamma^*_{\text{old}} + \Delta_{OLPC} \cdot (1 - \text{FER}^*) & \text{if RSN} > \text{target-RSN} \\ \gamma^*_{\text{old}} - \Delta_{OLPC} \cdot \text{FER}^* & \text{otherwise,} \end{cases} \tag{1}$$

where Δ_{OLPC} is the step size and FER* is the target frame error rate for the first transmission (usually 10%).

2.2 Resource Assignment

The most prominent feature of the Enhanced Uplink is rate control. Rate control is realized indirectly by transmitting scheduling grants, which denote the maximum allowed power offset the UE may use for the transmit power of E-DPDCH physical data channel over the DPCCH physical control channel. The received power S_k of a UE k at a NodeB can be expressed as

$$S_k = S_{c,k} + \Delta_{G,k} \cdot S_{c,k}, \quad (2)$$

where $S_{c,k}$ is the received power of the DPCCH and $\Delta_{G,k}$ is the scheduling grant. The possible values for $\Delta_{G,k}$ are defined in [2] in a table with 38 entries, where the entry with index 0 is the zero grant which means that the UE pauses its transmission. Grants can be either set as absolute value via the absolute grant channel or as UP/DOWN/HOLD commands on the relative grant channel. The relative grant channel exists for all NodeBs in the active set of the UE, but only the Serving NodeB is allowed to send UP commands. Grants from non-serving NodeBs (other-cell grants) are used to avoid the flooding of an adjacent cell with interference so that here only DOWN and HOLD commands are allowed. The condition for sending DOWN commands is defined by the total received power at the NodeB and by the ratio between other-cell E-DCH power to own-cell E-DCH power and can be chosen by the operator. Mobiles which are not in soft handover only have one associated RGCH, while the one at the cell border may additionally receive DOWN commands.

The selection of the E-TFC (and therefore of the instantaneous bit rate) according to the signaled power grant is task of the UE. This is done in two steps: First, the E-TFC restriction algorithm identifies the set of E-TFCs which require a power offset lower then the maximum allowed power offset, which is determined from the current power grant and the uplink power headroom (UPH) as following:

$$\Delta_{max,k} = \min\left\{\Delta_{G,k}, \frac{T_{max,k}}{T_{c,k}}\right\}, \quad (3)$$

where $T_{max,k}$ is the maximum allowed transmit power and $T_{c,k}$ is the power for the DPCCH. The power offset requirements for the E-TFC is calculated according to a reference E-TFC, gain factors and the number of codes. If the power offset of an E-TFC exceeds the maximum for a certain time, it is marked as restricted, otherwise it is in "allowed" state. The time periods of this filter mechanism are defined in the specifications, however, for a more detailed discussion see [6]. After the restriction process, the UE may then choose a E-TFC suitable for data volume in the transmit buffer.

2.3 Radio Resource Management

The task of the Radio Resource Management (RRM) is responsible for calculating the power grants for the UEs. This can be done according to different criteria. One possibility which has been chosen in [8] is to calculate the available load as the difference between a target load η^*, the load of the DCH users and the other-load. Another possibility is to choose the a certain maximum own-cell load as reference like proposed in [7]. In this work, we use a certain maximum interference or received power I^* as threshold which must not exceeded. The following condition should hold:

$$I_{D,own} + I_{E,own} + I_{oc} \le I^*, \quad (4)$$

with $I_{D,own}$ as received power from DCH users within the own cell, $I_{E,own}$ as corresponding received power from E-DCH users and I_{oc} as other-cell interference from UEs in adjacent cells. In case of active E-DCH connections, the RRM entity tries to maximize the power from E-DCH users, such that the total received power is close to I^*. We further define the received power share which is available for E-DCH connections as I_E^* and assume that the total received power meets the target interference, such that

$$I_E^* = I^* - I_{D,own} - I_{oc}. \quad (5)$$

The principle of RRM is illustrated in Fig. 1. A general goal

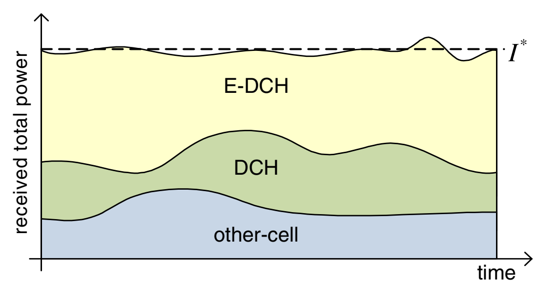

Figure 1. RRM principle: E-DCH power adapts to DCH and other-cell interference

of an operator in this context is to avoid "overshoots" due to rapid and strong fluctuations of the interference, since this can lead to outage and decreasement of link quality.

3 Scheduling

The scheduler in the NodeB is responsible for the assignment of power grants to the UEs. The scheduler can use the following feedback information from the UEs for its decision: The uplink power headroom (UPH), defined as $U = \frac{T_{max,k}}{T_{c,k}}$, the buffer status (total buffer and per priority class) and the "happy bit", which indicates whether the incoming traffic at the UE can be transmitted with the current resource assignment. Further, in [10] it is suggested to calculate an uplink CQI similar to the channel quality indicator used for HSDPA.

The assignment of scheduling grants is theoretically every TTI possible, however, it is expected [13] that the scheduling interval is larger in order to avoid signaling overhead on the downlink. For that reason we used mean values over a time window for scheduling. An additional factor which leads to some delay between the assignment and an effect is the E-TFC restriction algorithm.

In the following, we investigate three different schedulers: A resource-fair, a delay-aware and a throughput-aware scheduler, whereas the resource-fair scheduler is the reference for the two other scheduling disciplines.

Resource-fair Scheduler

The first, which we denote "resource-fair", grants each E-DCH UE $k \in \mathcal{E}$ in the cell the same amount of interference, which means that all UEs get the same power grant Δ_G. The scheduler calculates the power grant from the target E-DCH received power and Eq. 2 as following:

$$\sum_{k \in \mathcal{E}} S_{c,k} \cdot (1 + \Delta_G) = I_E^* \quad \Rightarrow \tag{6}$$

$$\Delta_G = \frac{I_E^* - \sum_{k \in \mathcal{E}} S_{c,k}}{\sum_{k \in \mathcal{E}} S_{c,k}} \tag{7}$$

Consequently, this means that all UEs in the inner area of the cell have approximately the same bit rate. In the outer region and depending on the number of concurrently transmitting UEs, the required transmit power may be higher than the maximum transmit power. In this case, some of the resources are wasted.

Delay-aware scheduler

For multimedia traffic like video conferencing the packet delay is an important factor for the user satisfaction. For such traffic, we propose to use a delay-aware or buffer-aware scheduler. Let K_E be the number of E-DCH UEs, and let us assume that the received powers of the DPCCH of all UEs is approximately the same. Then,

$$K_E \cdot S_c + S_c \cdot \sum_{k \in \mathcal{E}} \Delta_{G,k} = I_E^* \tag{8}$$

$$\Rightarrow \quad \sum_{k \in \mathcal{E}} \Delta_{G,k} = \frac{I_E^*}{S_c} - K_E. \tag{9}$$

So, the available resources are expressed by the right hand side of the second equation. For the delay-optimizing scheduler, we set the buffer occupancies B_k in the UEs in relation to the power grant the UEs should receive:

$$\Delta_{G,k} = \frac{B_k}{\sum_{j \in \mathcal{E}} B_j} \cdot \left(\frac{I_E^*}{S_c} - K_E \right) \tag{10}$$

Throughput-aware Scheduler

For best-effort traffic, the throughput is an important QoS measure. A throughput-maximizing scheduler would assign the complete power resources to the UEs nearest to the NodeB. However, this can lead to starvation of users further away, so we implemented a scheduler which is channel-aware, but still gives all UEs power resources. The principle is the same as for the delay-aware scheduler, but instead of using the buffer occupancy, the uplink power headroom (in dB) is used in Eq. 10:

$$\Delta_{G,k} = \frac{U_k}{\sum_{j \in \mathcal{E}} U_j} \cdot \left(\frac{I_E^*}{S_c} - K_E \right) \tag{11}$$

This gives UEs more resources proportional to their UPH, which corresponds to the link quality and also roughly to the distance to the NodeB.

4 Numerical Results

In this section we present some numerical results to show the impact of the different schedulers. The simulation parameters are listed in Tab. 1.

Table 1. Simulation parameters

Parameter	Value
target RWTP	-103 dBm
target RSN	1
target FER at first HARQ transmission	10%
max. number of transmissions	3
RLC mode	ack./unack.
OLPC step size	1 dB
UE class	4
MAC-d packet size	320 bit
RRM interval	50 ms, 100 ms
Max. UE transmit power	21 dBm
Distance between NodeBs	1.2 km
Prop. loss model	COST-231-Hata
Channel model	AWGN
Thermal noise density	-174 dBm

4.1 Interference

The first scenario gives an impression of the distribution of the received powers at the NodeB. In the center cell, 5 E-DCH users and 2 DCH users have been placed, in the outer

6 cells 2 DCH users and 1 E-DCH users each. The DCH users send with constant bit rate of 128 kbps, the E-DCH users have video traffic. The RRM interval is set to 50 ms. Figure 2 shows the received powers at the center NodeB for the different sources. The sold line with exception of the line annoted with “total” denotes DCH as source, while the dashed lines denotes E-DCH. In this scenario, the received powers from the DCH is less variant than the powers from E-DCH users, which is because the user are immobile and the DCHs are always ON. Furthermore, the target RWTP is exceeded with a probability of ca. 60%. However, the 99% quantile of the total RWTP is with -100 dBm, which corresponds to a cell load of approx. 0.85, in acceptable regions. It should be noted that this scenario is a static example scenario, not a full evaluation of the interference distribution.

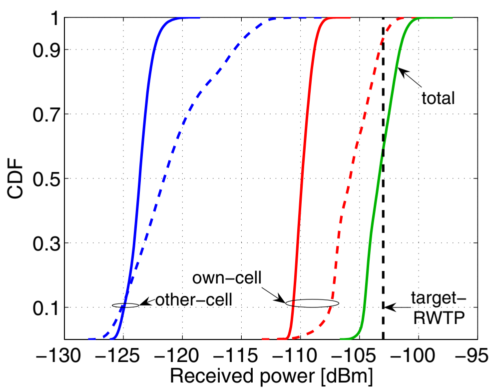

Figure 2. Received powers at the central NodeB for resource-fair scheduling

4.2 Delay

In the next scenario, we compare the resource-fair scheduler with the delay-aware scheduler. The video traffic for the E-DCH UEs is generated from a H.263 video trace with high variance, the length of the trace is 120s. The 5 UEs start transmitting with random time-offsets. Total simulation time was 1000 s. The distance to the NodeB is 300 m for all 5 UEs. In Fig. 3, the CDF of the end-to-end UDP packet delay is shown. For a RRM interval of 100 ms, the delay-aware scheduler leads to significant better results than the resource-fair scheduler, which has an 90%-quantile of 900 ms. With a smaller RRM interval, the results become overall better for both schedulers, however the delay-aware scheduler is still signficant better especially for higher quantiles.

In Fig. 4 the CDF of the buffer occupancy is shown. Note

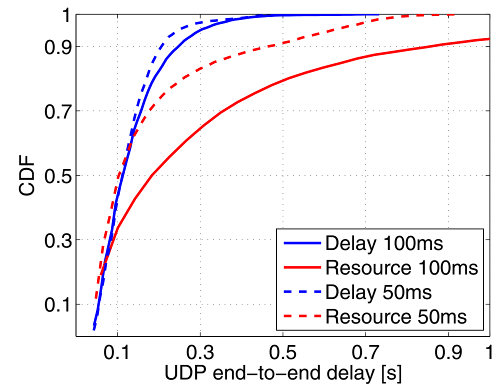

Figure 3. CDF of end-to-end packet delay

that the maximum value is according to the TEBS-table in [2] 37500 bits, which corresponds to 100%, although the real buffer size in this case is higher. The results show that if the buffer size would have been 37500 bits, the resource-fair scheduler would have lead to significant packet droppings due to buffer overflows.

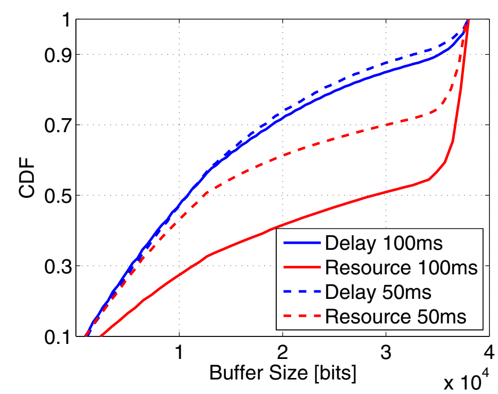

Figure 4. CDF of buffer sizes

4.3 Throughput

In the final scenario, we compare the throughput-aware scheduler with the resource-fair scheduler. In order to see the effect of the UPH on the performance of the individual UEs, they have different distances to the NodeB. Two are placed at a distance of 300 m, two at 600 m and one at 1000 m. All UEs generated Best-effort traffic over TCP (FTP) in RLC acknowledged mode. Figure 5 shows the cumulated cell throughput measured between NodeB and RLC. The throughput-aware schedule has a gain of approximately 200 kbps, which is on the one hand due to the higher grants for UEs close to the NodeB and on the other hand due to avoiding assignments to UEs which cannot fully use them.

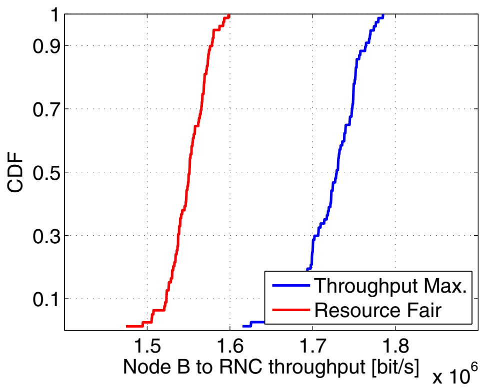

Figure 5. CDF of cell throughput

This is clarified in Fig. 6 and 7, which show fragments of the throughput traces of the individual UEs. While in Fig. 6, the throughput of the users is proportional to their distance to the NodeB, in Fig. 7 the two users close to the NodeB have the same throughput, while the UE at 1000 m has lower throughput due to the restriction of the transmit power to 21 dBm.

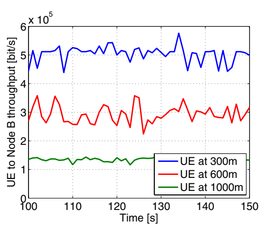

Figure 6. Throughput-aware trace

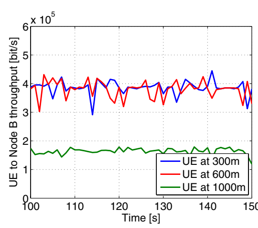

Figure 7. Resource-fair trace

5 Conclusion

In this paper we gave an overview of the UMTS Enhanced Uplink system and explained the most important new technical features and how they have been implemented in the simulation model we used for numerical results. A special focus was on the scheduling or rate assignment mechanism. As reference scheduler, a resource-fair scheduler has been proposed which assigns the available power resources equal to all users. For delay-sensitive traffic like video streaming, a delay-aware scheduler has been developed which assigns the resources proportional to the buffer occupancies of the UEs. Finally, the throughput-aware scheduler increases the overall system throughput due to channel-aware resource assignments proportional to the uplink power headroom of the UEs. The results show that both schedulers lead to significant performance gains for the traffic classes they were designed for. In this context, an area of further research is the design of a scheduler for mixed traffic situations either in the UE itself or for different UEs within the same cell. On system level, a further topic is the impact of features like other-cell grants on the system performance.

References

[1] ISCML website. http://www.iterativesolutions.com/Matlab.htm.

[2] 3GPP. TS 25.321 V6.6.0 Medium Access Control (MAC) protocol specification. Technical report, Sept. 2005.

[3] M. Boussif, J. Wigard, T. E. Kolding, and N. A. Madsen. Errors on the HSUPA E-HICH Channel and Their Effect on System Performance. In *Proc. of IEEE VTC Spring '07*, Dublin, Ireland, Apr 2007.

[4] D. Chase. Code Combining – A Maximum-Likelihood Decoding Approach for Combining an Arbitrary Number of Noisy Packets. *IEEE Transactions on Communications*, 33(5), May 1985.

[5] K. W. Helmersson, E. Englund, M. Edvardsson, C. Edholm, S. Parkvall, M. Samuelsson, Y.-P. E. Wang, and J.-F. Cheng. System Performance of WCDMA Enhanced Uplink. In *Proc. of IEEE VTC Spring '05*, Stockholm, Sweden, May 2005.

[6] T. Henttonen and T. Chen. Power Headroom Measurements for E-TFC Elimination/Selection in HSUPA. In *Proc. of IEEE VTC Spring '06*, Melbourne, Australia, May 2006.

[7] T. Liu, A. Mäder, and D. Staehle. Analytical Other-Cell Interference Characterization over HSUPA-Enabled Multi-cell UMTS Networks. In *Proc. of IEEE VTC Fall '07*, Baltimore, MD, USA, Oct. 2007.

[8] A. Mäder and D. Staehle. An Analytical Model for Best-Effort Traffic over the UMTS Enhanced Uplink. In *In Proc. of IEEE VTC Fall '06*, Montréal, Canada, Sept. 2006.

[9] A. Mäder and D. Staehle. Comparison of Preemptive and Preserving Admission Control for the UMTS Enhanced Uplink. In *Proc. of 15th KiVS*, Bern, Switzerland, Feb. 2007.

[10] C. Rosa, J. Outes, T. B. Sørensen, J. Wigard, and P. E. Mogensen. Combined Time and Code Division Scheduling for Enhanced Uplink Packet Access in WCDMA. In *Proc. of IEEE VTC Fall '04*, Los Angeles, USA, Sept. 2004.

[11] M. Sundelin, C. Milard, M. Ringström, J. Skördeman, E. Sparrman, T. Sundin, T. Tynderfeldt, and H. Schmekel. WCDMA Enhanced uplink - Test bed and Measurements. In *Proc. of IEEE VTC Spring '05*, volume 3, Stockholm, Sweden, May 2005.

[12] J. Voigt and K. Pannhorst. Optimizations on Scheduling Strategies for Enhanced Uplink on WCDMA. In *Proc. of IEEE VTC Spring '07*, Dublin, Ireland, Apr. 2007.

[13] W. Xiao, R. Ratasuk, A. Ghosh, and R. Love. Scheduling and Resource Allocation of Enhanced Uplink for 3GPP W-CDMA. In *Proc. of IEEE PIMRC '05*, Berlin, Germany, Sept. 2005.