Effect of Feedback Prediction on OFDMA System Throughput

Mohammad Abdul Awal
LRI, Univ. Paris-Sud, CNRS
Bat 490, 91405 Orsay, France
awal@lri.fr

Lila Boukhatem
LRI, Univ. Paris-Sud, CNRS
Bat 490, 91405 Orsay, France
lila.boukhatem@lri.fr

ABSTRACT

In wireless communication systems, adaptive modulation and coding (AMC) is used to improve the downlink (DL) spectral efficiency by exploiting the underlying channel condition. However, AMC also reduces the uplink (UL) throughput because of feedback (FB) overhead which is used to carry channel information. Opportunistic mechanisms provide a good solution to reduce the FB overhead but in each scheduling slot, they are designed to serve a very few mobile stations (MSs) only out of a large number of simultaneously active MSs. In multichannel multiuser systems, the opportunistic FB scheme will have no effect in FB reduction if all the MSs need to be scheduled in each scheduling slot. To reduce the FB overhead, we propose a base station (BS) side FB prediction strategy that reduces FB overhead for MSs experiencing any channel condition. We use recursive least square (RLS) algorithm which adapts with SNR variations over time and provides predicted SNR with minimized error. With numerical analysis and simulations, we show that this FB reduction mechanism can be used to increase the UL throughput with non-noticeable DL degradation.

Categories and Subject Descriptors

C.2 [**COMPUTER-COMMUNICATION NETWORKS**]: Local and Wide-Area Networks—*Access schemes*

General Terms

Algorithms, Design, Performance

Keywords

Feedback reduction, OFDMA, WiMAX, Prediction, Uplink Throughput

1. INTRODUCTION

With the increasing improvements in broadband wireless technologies, mobile users are desiring access to high-speed data and multimedia services from anywhere and at anytime. To meet the different users service requirements, the recent and future generations wireless communication technologies like IEEE 802.11a wireless LAN (Wi-Fi), IEEE 802.16e wireless broadband access systems (WiMAX) and 3GPP Long Term Evolutions (LTE) etc. will provide high quality of service (QoS) and bit rates to the MSs by employing a variety of techniques. In multi-user wireless environments, one of the reliable solutions is OFDMA (Orthogonal Frequency Division Multiple Access) technology which is being considered as a modulation and multiple access method for 4th generation wireless networks. The high bandwidth of OFDMA comes from thousands of orthogonal subcarriers. By grouping subcarriers, OFDMA systems can have many logical subchannels and each MS can be assigned one or more subchannels based on each MS requirements. Since the channel condition of a mobile station (MS) changes over time due to the fading, interference and path loss, the next generation wireless systems use adaptive modulation and coding (AMC) techniques to change their sending rates based on the channel quality information (CQI) in terms of measured signal-to-noise ratio (SNR). The AMC helps to increase the DL spectral efficiency with the penalty of FB overhead from each MS and reduces the UL efficiency for the MSs. A MS needs to send $M \times b$ bits of information to the base station (BS) to inform full channel information, where M is the number of subcarriers in a channel and $b = log_2(\gamma_{sc})$ bits required to represent the SNR γ_{sc} for each subcarrier. But sending FB for each subcarrier is not practical because of this large overhead.

Recently, there has been some research on reducing the channel FB information [2, 4, 7, 10]. Using multiuser diversity [10], a systems' UL (MS to BS) capacity can be maximized by picking a MS with the best channel to transmit. [7] proposed opportunistic FB to combat the FB load. In this approach, the BS sets a SNR threshold γ_{th} and announces it to the MSs. In each FB reporting slot, only the MSs with SNR γ higher than γ_{th} send FB and the BS schedules the MS reporting best γ. The performance of this algorithm increases with the number of MSs. These opportunistic approaches provide a good chance to reduce the FB overhead but in each scheduling slot, it is designed to serve only a very few number of MSs out of a large number of simultaneously active MSs. In multichannel multiuser system, the opportunistic FB scheme will have no effect in FB reduction if all the MSs need to be scheduled in each scheduling slot. Moreover, the MSs report the FB in a contention based medium, where there are possibilities of collision between FBs of MSs

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IWCNC'09, 21-24 June, 2009, Leipzig, Germany.
Copyright ©2009 ACM 978-1-60558-569-7/09/06 ...$5.00.

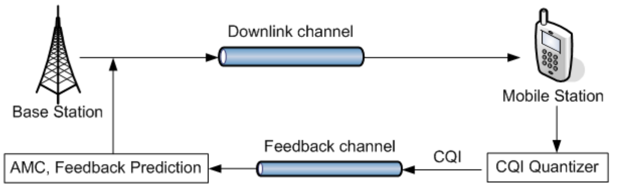


Figure 1: Multiuser SISO wireless communication system with a single base station

having same SNR opportunity. The effect of the degree of FB quantization is shown in [4] which showed that the throughput with quantization is only slightly less than the throughput obtained by using unquantized FB. The opportunistic method is applied to a multichannel system in [2] to have FBs for best-n subcarriers. Gao et al. and Devid Chen [1, 6] studied the overhead model of a multichannel OFDMA systems like WiMAX. For example, a system with 10 MSs and 2:1 DL to UL ratio, the UL overhead could be as much as 25%, out of which, more than 8% are from FB overhead [6]. With a higher number of MSs, more UL bandwidth will be consumed by overhead.

An encouraging approach to reduce these FB loads could be FB prediction. Even though applying prediction methods in wireless communication is nothing new, most of the works [8, 13] focused on estimating the channel behavior at the receiver. Schafhuber [13] proposed the novel schemes for decision directed minimum mean-square error (MMSE) and adaptive channel prediction in wireless OFDM systems. MMSE channel predictors require channel statistics knowledge which in unknown in practical applications, and thus would have to be estimated prior to the design of the MMSE channel predictor. Fortunately, adaptive channel predictors do not require any statistical prior knowledge and are able to track non-stationary channel and noise statistics. Moreover, adaptive predictors perform close to the MMSE predictors [13]. [3, 8, 13] have compared the performance of several prediction algorithms and showed that recursive least-squares (RLS) based algorithms which are adaptive channel predictors provide better performance results compared to other algorithms.

In this paper, we propose to reduce the FB overhead by BS-side FB prediction based on RLS algorithm. Upon receiving SNR from a MS, the BS may predict the probable SNR for the next frame so that the MS does not need to send the SNR back to BS in next frame. Designing a prediction algorithm is not our objective in this paper. Rather, we want to apply a well known prediction algorithm and show the feasibility to reduce the FB overhead of a multiuser system. The goal of our work is to reduce the FB load as much as possible from each MS by predicting the SNR in the BS and use these saved bandwidth to carry more user data traffic, hence, increase the overall UL throughput. We organized the paper as follows. In section 2 and 3, we introduce the system model and the system overhead model. The prediction algorithm is presented section 4 and the performance analysis is presented in section 5. Finally, we conclude our studies.

2. SYSTEM MODEL

We consider an OFDMA cellular system with bandwidth B which consists of L subcarriers. These subcarriers are equally divided and allocated to N simultaneously active MSs communicating with the single BS as shown in Fig. 1. A basic frame of an OFDMA system (Fig. 2) consists of a DL and UL subframes while there exist a TTG (Tx/Rx Transition Gap) between DL and UL subframes, and a RTG (Rx/Tx Transition Gap) between two successive frames. MSs are allocated slots for data transfer and these slots represent the smallest possible data unit. A slot is defined by a time and subchannel dimension and it varies depending on the different operating modes of the system. The baseband channel model of the system can be written as

$$y_i(t) = h_i(t)x_i(t) + n_i(t), i = 1, 2, ..., N. \quad (1)$$

where $x_i(t)$ is the complex transmitted signal of MS i at time slot t, $y_i(t)$ is the complex received signal of MS i in time slot t, and $n_i(t)$ is zero mean complex Gaussian noise. The fading channel gain from the base station to the ith MS in time slot t is $h_i(t)$. We assume that the fading coefficients of all MSs are independent. Let $\gamma_i(t)$ denotes the instantaneous (in time slot t) received SNR of MS i. For simplicity, we omit the subscript i which is used to indicate the MS number. The system considered is time-division duplex (TDD), that is, the same carrier frequency is used for both DL and UL. We can therefore assume that a MS would most likely to experience the similar level of SNR in its both DL and UL portion of the frame. The MSs measure SNR for each DL transmission and send back these SNRs to the BS via the UL FB channel. The SNR received in a frame f is used to determine the AMC in next frame $f + j$, for $j \geq 1$. It is also assumed that the BS receives perfect SNR with zero-delay from the MS.

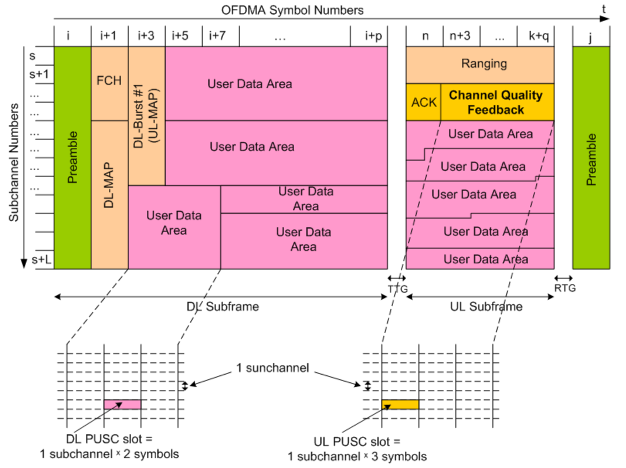


Figure 2: Frame structure of an OFDMA system (IEEE 802.16e)

3. OVERHEAD MODEL

The overhead model of an OFDMA system depends on the ratio of DL and UL, number of active MSs, slot allocation per MS and control information carried in the frame. Different OFDMA-based systems may have different allocation schemes and methods. Besides the data transmission areas, both DL and UL subframes consist of some signaling parts to ensure the optimal system operation. The DL overhead portion consists of preamble, frame control header (FCH) and resource mapping information. The UL overhead portion consists of ranging, acknowledgement, and CQI. The overall number of slots used for ACK and CQI depends on the number of simultaneously active MSs N.

To illustrate the UL overhead model, as an example, we take the parameters from mobile WiMAX (IEEE 802.16e) system with 10 MHz bandwidth at 2.3 GHz band. Fig. 2 shows the frame structure of WiMAX. In the DL (UL), there are 720 (560) data subcarriers which are divided into 30 (35) subchannels. Different schemes of permutations for data and pilot subcarriers are used to define those subchannels which are partially used subchannelization (PUSC), fully used subchannelization (FUSC), and band AMC. In WiMAX DL PUSC, one slot is a single subchannel by two OFDMA symbols and for UL PUSC, one slot is a single subchannel by three OFDMA symbols (Fig. 2). The frame is composed of 48 OFDM symbols in time dimension, corresponding to $5ms$, which includes the DL subframe, TTG and UL subframe (Fig. 2). With 2:1 DL:UL subframe ratio and 1 symbol for TTG, one sub-channel in UL subframe has 15 symbols equivalent to 5 slots (each UL slot contains one subchannel by three OFDM symbols), which makes 175 slots with 35 subchannels. In WiMAX, the ranging channel in UL is composed of one or more groups of six (or eight) adjacent subchannels. For simplicity, we assume that one group of six subchannels is enough, which requires 30 slots. For ACK and CQI channels, each ACK and CQI occupies half slot and one slot respectively [11]. For example, if the number of simultaneously active MSs is 10, ACK and CQI both occupy 15 slots. Hence, the overall control overhead is $30 + 15 = 45$ slots. So the overhead ratio is 45/175 (i.e. 25.7%). The CQI overhead ratio is itself 10/175 (i.e. 5.7%) which could increase to 11.4% for 20 MSs.

4. PREDICTION METHOD

We propose to implement a RLS prediction algorithm [12] in the BS-side for the specific purpose of anticipating the future SNR level of the received signal at the receiver side. BS-side prediction with RLS algorithm is also proposed in [3,9]. Jordan et al. [9] used the RLS in opportunistic beamforming to mitigate the reporting delay from the MS because of propagation delay. They calculate filter coefficients only once for every angle in the beamforming process and reuse these coefficients in further cyclic periods of the beamforming process. Duel-Hallen et al. [3] used RLS-based long range prediction to track correlated sub-carriers so that the MS sends only one CQI for all the correlated sub-carriers. In our work, we assume that each MS sends the average SNR for all the subcarriers assigned to it. We used short range prediction in every time-slot to estimate the probable SNR level that the receiver is going to experience and avoid explicit FB from the receiver. The predictability relies on error measures expressed in terms of a time average of the actual received SNR instead of a statistical average. The least square error e_{RLS} for a MS based on time average of window size w is defined as [12]

$$e_{RLS}(t) = \sum_{j=t-w}^{t} (\lambda^{(t-j)} e^*(t)e(t)) \qquad (2)$$

where λ is the scalar weighting factor with $0 < \lambda \leq 1$ that can change the performance of the prediction. λ is applied to the previous w input data. $e^*(t)$ is the complex conjugate of $e(t)$, and the error $e(t)$ according to [12] is defined as

$$e(t) = \gamma(t) - \hat{\gamma}(t), t \geq 1 \qquad (3)$$

where $\gamma(t)$ is the actual SNR received at time slot t from MS i, and $\hat{\gamma}_i(t)$ is the estimated SNR for time slot t. The SNR is estimated as [12]

$$\hat{\gamma}(t) = F(t-1)\gamma_w(t), t \geq 1 \qquad (4)$$

where $F(t)$ is the w-th order prediction filter coefficient and $\gamma_w(t)$ is the previous actual or estimated SNRs of moving window of size w up to time slot t. The prediction filter coefficient is updated as

$$F(t) = F(t-1) + K(t)e^*(t), t \geq 1 \qquad (5)$$

where the $K(t)$ is the RLS gain vector given by

$$K(t) = \frac{R^{-1}(t-1)\gamma_w(t)}{\lambda + \gamma_w^T(t)R^{-1}(t-1)\gamma_w(t)}, t \geq 1 \qquad (6)$$

Here, $\gamma_w^T(t)$ is the transpose of $w \times 1$ matrix $\gamma_w(t)$. The matrix $R^{-1}(t)$ is the inverse of the $w \times w$ sample covariance matrix, it can be calculated recursively as

$$R^{-1}(t) = \frac{1}{\lambda}[R^{-1}(t-1) - K(t)\gamma_w^T(t)R^{-1}(t-1)], t \geq 1 \qquad (7)$$

The recursion is initialized as

$$F(0) = K(0) = \gamma(0) = 0, R^{-1}(0) = dI_{ww}$$

where I_{ww} is an $w \times w$ identity matrix and d is a large positive constant. We put $d = 5000$.

5. PERFORMANCE ANALYSIS

5.1 Simulation Model

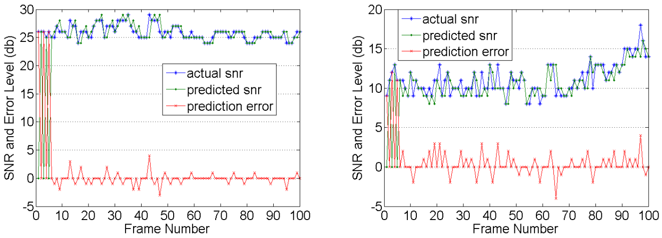
The performance evaluation of the SNR prediction method has been investigated using Matlab simulations. We consider a BS communicating with $N = 10$ to 20 active MSs. For consistency, we used 10 MSs in all the simulations except for the last set of experiment (Fig. 6), where we used up to 20 MSs. The MSs are randomly distributed in the cell with cell radius $1km$. In each analysis scenario, we chose 33% of the MSs with ITU pedestrian B channel and mobility speed 3 km/h while other MSs with ITU vehicular A channel and mobility speed 30 km/h. We assumed that all the MSs are already registered with the BS and ready to start any data session. Each MS has a session of 25 seconds equivalent to 5000 frames (200 frames/sec for $5ms$ frame duration). The BS receives the γ_i from each MS_i for $i = 1, 2, ..., N$, and maintains an instance of RLS algorithm RLS_i for each MS_i with γ_i as the input. The BS either uses the γ_i or $\hat{\gamma}_i$ to determine the AMC for next frame of MS_i. The minimum receiver SNR required to decide each level of AMC according to the WiMAX standard [11] is shown in Table 1.

5.2 Effect of Prediction Error

Fig. 3 plots γ, $\hat{\gamma}$ for an arbitrarily selected MS with period $p = 2$ and the error $e = |\gamma - \hat{\gamma}|$. The value p means the MS sends one FB out of p frames and other $p-1$ values predicted by BS. Here, for simplicity and visibility, we show the result of first 100 frames only. Even though the errors are high at the initial iterations, they converge around 0 with further

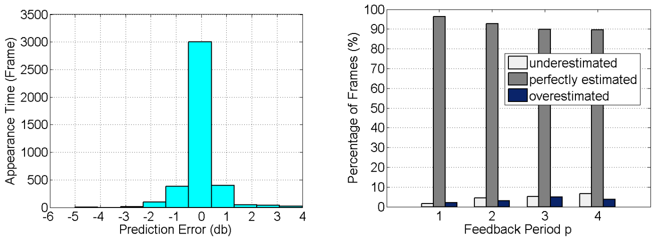
Table 1: Minimum receiver SNR required to use in AMC and their bitrates per symbol per carrier

Modulation	Level	Coding	Minimum Receiver SNR (db)	Actual data bits per slot (PUSC DL and UL)
BPSK	1	1/2	6.4	24
QPSK	2	1/2	9.4	48
	3	3/4	11.2	72
16QAM	4	1/2	16.4	96
	5	3/4	18.2	144
64QAM	6	2/3	22.7	192
	7	3/4	24.4	216



(a) (b)

Figure 3: Actual and Predicted SNR with prediction error for feedback SNR with period $p = 2$ (a) for MS with ITU pedestrian B channel and mobility speed 3 km/h (b) for MS with ITU vehicular A channel and mobility speed 30 km/h.



(a) (b)

Figure 4: (a) Predicted Error Distributions for feedback SNR with period $p = 2$ for MS with ITU vehicular A channel and mobility speed $30km/h$ (b) AMC estimation error from predicted SNR with different period p.

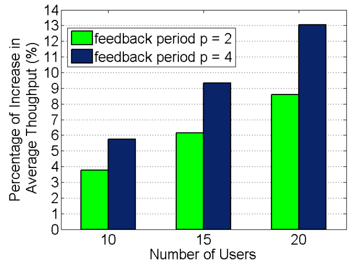


Figure 5: Percentage of increase in average uplink throughput for different number of users and different feedback periods.

iterations as the algorithm have better knowledge about the SNR covariance. We discard the $\hat{\gamma}_i$ for first 1000 frames and used rest of the results for our analysis. It is possible to discard less frames as the prediction error converges as early as 7 frames only (Fig. 3). Fig. 4(a) presents the histogram of errors and shows that the average prediction error is 0 db and variance 0.57 db. The prediction for MS with higher mobility shows wider range of error than the MS with lower mobility. To see how effective our mechanism is, we use γ_i and $\hat{\gamma}_i$ to determine the AMC and compare the decision performance which is shown in Fig. 4(b) (for last 4000 frames). For $p = 2$, it shows that the algorithm decides the AMC perfectly in 96.25% frames, underestimated the AMC in 1.60% frames and overestimated in 2.15% frames. With underestimated (overestimated) AMC level, we mean a level of AMC (see Table 1) which is lower (higher) than the expected level of AMC. The underestimated (overestimated) AMC decisions would not affect the total throughput of DL data transfer if the chosen AMC is the lowest (highest) level of AMC. But the overestimation of AMC other than highest level may cause the MSs to experience higher bit-error-rate (BER) of received data. The only case that would decrease the throughput performance is if the decisions are underestimated based on predicted SNR while the actual AMC should be higher than the estimated one.

With the increase of p, the rate of perfect AMC estimation decreases but in very slow rate. With $p = 5$, the algorithm results perfect estimation in 90% frames. It is worth to notice that, in case of higher value of p, the rate of underestimated AMC decision increases compared to the rate of overestimation. This allows us not to send too many data with overestimated AMC and provides a conservative approach to maintain the BER. Fig. 4(b) shows that the performance of the prediction algorithm decreases when p increases. Other than p, the performance of RLS algorithm depends on two design parameters w and λ, where w is the moving window and λ is the forgetting factor. We obtained good results for $\lambda = 0.99$ as it gives more weight on recent SNR values than older SNR values. In a non-stationary Rayleigh fading channel, the channel varies over time and results SNR with large deviations. With a larger window size, the algorithm tends to give less accurate value for the future SNR as it tries to predict based on the data with larger deviations. In our studies, window size w equals to the FB period p gives more accurate results.

5.3 Effect on Uplink Throughput

In the overhead model section, we state that for DL:UL ratio 2:1, there are 15 symbols available for UL, which makes 5 PUSC UL slots for each subchannel. For 10 MSs, the UL overhead is 45 slots (as described in the overhead model in section IV) out of total 175 UL slots which means there are 130 slots available to carry UL data. Now, in the frame, where the CQI is predicted, 140 slots will be available to carry data as the MSs does not send the FB. We calculated the total UL throughput for both cases. One case is, every MS sends FB in every frame and another case is every MS sends 1 FB in every p frames. Then, we determined how much increase in throughput we have with prediction mechanism. Fig. 5 shows that for 10 MSs, the throughput incre-

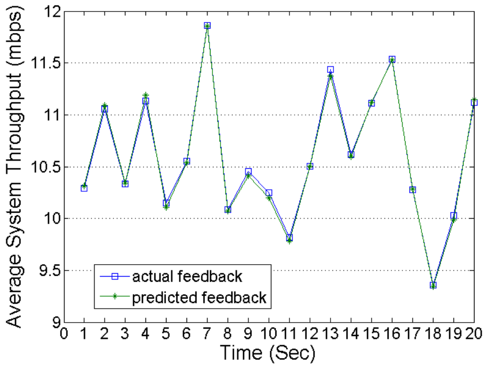


Figure 6: System downlink throughput with actual and predicted CQI.

ment is 3.76% and 5.76% with $p = 2$ and $p = 4$ respectively compared to $p = 1$. This percentage of throughput increases with the number of MSs. The throughput enhancement is 8.60% and 13.03% with $p = 2$ and $p = 4$ respectively for 20 MSs. As the UL overhead increases with the increase of number of active MSs in the network, the prediction algorithm would allow saving more FB overhead and increase the UL throughput.

5.4 Effect on Downlink Throughput

We also analyzed the DL data throughput to show the effect of the non-perfect (underestimated and overestimated) AMC decisions. For DL:UL ratio 2:1, 32 symbols are available for DL. From [1, 6, 11], the DL overhead is 7 symbols, which is a conservative estimate [1, 5]. For the sake of simplicity, in our study, we take this overhead value. After using 7 symbols for DL overhead, 25 symbols are available to carry DL data, which makes 12 PUSC DL slots for each subchannel. For whole 10 MHz DL bandwidth, it includes $12 \times 30 = 360$ slots which are used to carry data for 10 MSs. According to Table 1, the BS decides the AMC level for actual or predicted CQI. Fig. 6 shows the system throughput with actual CQI ($p = 1$) and predicted CQI ($p = 2$). The system throughput varies over time as different MSs experience different channel conditions. The throughput with $p = 2$ shows unnoticeable degradation (maximum 0.54% during the $13th$ sec in Fig. 6) than the throughput with $p = 1$.

6. CONCLUSION

In this paper, we proposed a novel mechanism to reduce the FB overhead by allowing the BS to predict the MSs FBs based on RLS algorithm. As our mechanism predicts SNR from each MS in any channel condition, it would reduce the FB load for every MS. Through simulation and numerical analysis of the chosen scenario we showed that our mechanism achieves at least 3.76% increase in UL throughput for 10 MSs without noticeable degradation (at most 0.54%) in DL throughput. The UL throughput could increase up to 8.6% for 20 users with feedback period 2 and at least 13% with feedback period 4. Although our analysis is based on a basic prediction channel model, we believe that our work would give better performance with more improvement of the prediction algorithm and with more knowledge about the channel. Our mechanism can be applied to more general partial FB mechanisms that have the advantage of saving bandwidth and transmission power. Our work can be applied for designing next-generation multimedia communication systems that deal with multichannel resources such as in CDMA, OFDMA, and multi-antenna systems. As future work, we are investigating the rate of performance degradation with the increase of FB period p and its effect in overall system throughput including a complete analysis on the BER.

7. REFERENCES

[1] D. Chen. On the analysis of using 802.16e wimax for point-to-point wireless backhaul. In *IEEE Radio and Wireless Symposium*. IEEE, January 2007.

[2] Y.-J. Choi and S. Bahk. Partial channel feedback schemes maximizing overall efficiency in wireless networks. *IEEE Transactions on Wireless Communications*, 7(4):1306–1314, April 2008.

[3] A. Duel-Hallen, H. Hallen, and T.-S. Yang. Long range prediction and reduced feedback for mobile radio adaptive ofdm systems. *IEEE Transaction on Wireless Communication*, 5(10):2723–2733, October 2006.

[4] F. Floren, O. Edfors, and B.-A. Molin. The effect of feedback quantization on the throughput of a multiuser diversity scheme. In *GLOBECOM*, volume 1. IEEE, December 2003.

[5] W. Forum. Mobile wimax - part i: A technical overview and performance evaluation, August 2006.

[6] Y. Gao, L. Chen, X. Zhang, and Y. Jiang. Performance evaluation of mobile wimax with dynamic overhead. In *IEEE Vehicular Technology Conference, VTC2008-Fall*. IEEE, September 2008.

[7] D. Gesbert and S. Alouini. How much feedback is multi-user diversity really worth? In *International Conference on Communications (ICC)*. IEEE, June 2004.

[8] P. Jain. On the impact of channel and channel quality estimation on adaptive modulation. Master's thesis, Virginia Polytechnic Institute and State University, December 2002.

[9] M. Jordan, L. Schmitt, G. Ascheid, and H. Meyr. Prediction of downlink snr for opportunistic beamforming. In *IEEE Wireless Communications and Networking Conference, WCNC2007*. IEEE, March 2007.

[10] R. Knopp and P. Humblet. Information capacity and power control in single-cell multiuser communications. In *International Conference on Communications (ICC)*. IEEE, June 1995.

[11] I. P802.16e-2005. Part 16: air interface for fixed and mobile broadband wireless access systems: amendment for physical and medium access control layers for combined fixed and mobile operation in licensed bands, December 2005.

[12] T. Rappaport. *Wireless communications: principles and practice*. Prentice Hall publications, 1996.

[13] D. Schafhuber. *Wireless OFDM systems: channel prediction and system capacity*. PhD thesis, Vienna Technical University, Vienaa, March 2003.