Methods for Compression of Feedback in Adaptive Multi-carrier 4G Schemes

Víctor P. Gil Jiménez[†], Thomas Eriksson[‡], Ana García Armada[†],
M. Julia Fernández-Getino García[†], Tony Ottosson[‡] and Arne Svensson[‡]
[†]Dept. of Signal Theory and Communications
University Carlos III de Madrid (Spain)
‡Dept. Signal and Systems
Chalmers University of Technology (Göteborg. Sweden)

***Abstract*— In this paper, several algorithms for compressing the feedback of channel quality information are presented and analyzed. These algorithms are developed for a proposed adaptive modulation scheme for future multi-carrier 4G mobile systems. These strategies compress the feedback data and, used together with opportunistic scheduling, drastically reduce the feedback data rate. Thus the adaptive modulation schemes become more suitable and efficient to be implemented in future mobile systems, increasing data throughput and overall system performance as compared to wired systems ten years ago.**

I. Introduction

Wireless channels, in particular for cellular systems, have a common characteristic; they are time varying. In order to exploit the variations and approach capacity, systems are becoming adaptive attempting to adjust the transmission as much as possible to the instantaneous channel conditions. Furthermore, in a multi-user scenario, there also exists the so-called *multi-user diversity* [1]: the probability that at least one user experiences good channel propagation conditions increases as the number of users does. If there are many users, it is highly probable that not all of them will experience poor channel propagation but at least one of them will be in a good channel. Moreover, this situation is more advantageous in an OFDMA (Orthogonal Frequency Division Multiple Access) system, where the different orthogonal sub-carriers can be assigned to different users. The power and flexibility of an OFDM (Orthogonal Frequency Division Multiplexing) and OFDMA system are the reasons why OFDMA systems are one of the candidates being proposed for the downlink in the mobile 4G [2]–[4].

Regarding adaptivity, in order to maximize the throughput, at the base station or central point there is a scheduler who decides which user (or users) will transmit and in which sub-carriers. It has been shown in [5] that the throughput-maximising scheduling policy is to transmit to the user with the highest SNR (Signal-to-Noise-Ratio). Since the channel and SNR are usually estimated at the receiver, they must be fedback to the transmitter. At the transmitter side, the scheduler selects which users are going to transmit and also the adequate MCS (Modulation and Coding Scheme) based on the fedback SNR by every user. The data rate needed to convey the feedback may be very high, specially in multi-user OFDM systems [6] where terminals must feedback MCS in every sub-carrier. Moreover, it will depend on how often this information is expected. For this reason, the literature on how to efficient feedback data is increasing [7].

A very good analysis on the state-of-the-art in compression for feedback data can be found in [8] and references there in. In this paper, the authors categorize the techniques for reducing the feedback into three groups, namely *quantization*, *compression* and *SNR-limited*. In the first group the way to reduce the feedback rate is to quantize data (applied to real SNR values) and therefore the feedback is reduced [9], [10]. In the second group, some compression techniques are applied to feedback data such as LZW (Lempel-Ziv-Welch) lossless algorithm [11], Huffman codes [12] or different ones such as in [13] where entropy compression is used with an UVLC (Universal Variable Length Code). And finally in the third group, for a maximizing-throughput scheduling, the base station only allows to feedback data to those terminals with good channel conditions (above some threshold) [14], [15], sometimes denoted *opportunistic feedback*. The amount of feedback reduction using these techniques can be very high and thus they should be used jointly with the quantization or compression ones. It has been shown in [12] that opportunistic feedback may reduce more than 50% the need to feedback data. In this paper an optimized version of [12] is presented and analyzed obtaining much more compression gains. The proposal is based on the Huffman codes for compressing feedback data jointly with the opportunistic feedback.

The remaining of the paper is organized as follows: In section II the system where the algorithms are applied is described. Next, in section III the compression methods are explained and in section IV the obtained results are shown and discussed. In V some conclusions are drawn.

This work has been partly funded by the Spanish government with projects MACAWI (TEC 2005-07477-c02-02) and MAMBO2 (CCG06-UC3M-TIC-0698), and European COST Action 289.

II. System Description

This paper is focussed on an Adaptive multi-carrier system that is being defined for 4G cellular transmission [4], however the proposed algorithms and results can easily be adapted to another adaptive multi-carrier system. The scheme in [4] allows two different duplexing methods, namely, FDD (Frequency Division Duplexing) and TDD (Time Division Duplexing) and, in this paper, we will focus on the FDD version. The FDD is an OFDMA allocating 512 sub-carriers where only 416 are available for user data. The basic time-frequency unit for resource allocation and partitioning is denoted a *chunk* and consists of a rectangular time-frequency grid of 8 sub-carriers along 12 OFDM symbols duration, i.e. every chunk has $8 \times 12 = 96$ sub-carriers (or data symbols). The bandwidth of the system is 20 MHz and the OFDM symbol duration is 28.8 μs, thus a chunk is a block of 312.5 kHz $\times$ 345.6 μs [4]. In Fig. 1 this structure is depicted. There are 52 chunks per frame and every chunk can be assigned to a different user or a user can be allocated into different chunks. Sub-carriers in a chunk can be modulated using from $BPSK$ to $256QAM$ (there are 8 modulation schemes available). The system has been designed so that the channel variations within the chunks are rather small and therefore the same link adaptation parameters are used for all the sub-carriers in a *chunk*. Besides, the feedback parameters will be the modulation schemes (or SNR) of every chunk at every frame. At the receiver, based on channel estimation and prediction [16], feedback data are generated for the transmitter. This means that each user must feedback data every 345.6 μs (frame duration). In order to reduce the feedback load, an opportunistic feedback strategy is also used, but the amount of data is still too high and therefore some algorithms for compressing data are proposed.

III. Lossless Compression Strategy

In this section, the proposed lossless compression strategy will be described. In this approach two main algorithms have been developed to exploit jointly the frequency and time correlation properties of the channel response, although two other methods that only exploit time or frequency correlation have also been evaluated for comparison purposes. It should be noted that we are going to feedback the MCS, i.e. according to the predicted channel conditions in the whole chunk, the optimum modulation scheme is selected for each chunk at the receiver, and these data are the feedback.

A. Compression principles

One of the best options for compressing data are Huffman codes which assign a small number of bits to commonly occurring symbols and larger number of bits to infrequent ones. Another characteristic of these codes is that they are lossless, i.e. you can send compressed data and, at the receiver, information can be restored without error (if no errors occur in the transmission channel). However, in order to obtain good performance (i.e. high compression rates) the information should be correctly conditioned in the following sense: only a few sequences have to be very frequent and the others should be very infrequent. In this situation, since Huffman codes give small number of bits to the frequent sequences the amount of data is compressed. Originally feedback data are not well conditioned, because almost all the modulation schemes have the same probability to be used in the whole range of operation. Maybe at lower SNR the lowest schemes are more frequent or at higher SNR the highest ones, but in average, in all the operation range, the probabilities are close to uniform, and therefore, the Huffman codes will not obtain good compression rates. For this reason, we propose algorithms to transform feedback data into a more useful and well conditioned form. In order to do that, time and frequency correlation have been taken into account.

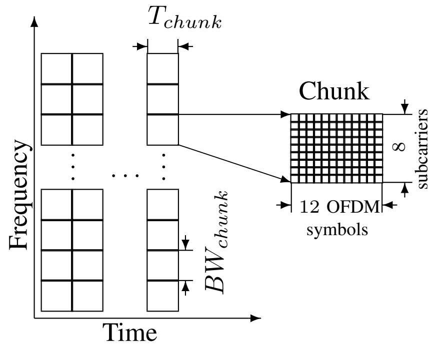


Fig. 1. Chunk description for the FDD mode

B. Time or Frequency Correlation algorithms

In a wireless channel, because of the coherence bandwidth, there exists some frequency correlation between different sub-carriers (applied to this system, different chunks); the modulation scheme used in one of the chunks may be the same or very similar to that in adjacent chunks. Furthermore, because of the coherence time, there also exists time correlation among chunks: the modulation scheme in one specific chunk may be the same or very similar to that in the previous chunk. For this reason, instead of feedback the modulation scheme, it can be advantageous to feedback the difference with respect to the adjacent chunk (frequency correlation) or previous chunk (time correlation). Once the difference has been computed the data are encoded by using Huffman codes. With such a pre-coding scheme, there will be a few symbols that occur far more frequent than the others such as 0, 1 or -1[1] as can be seen in Fig. 2 (where the probabilities of ocurrence of every

[1] 0 means that the modulation scheme is the same as previous, 1 means that the modulation scheme increases one and -1 that the modulation scheme decreases one.

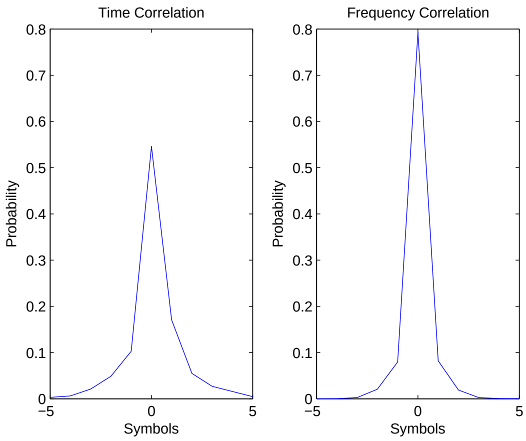
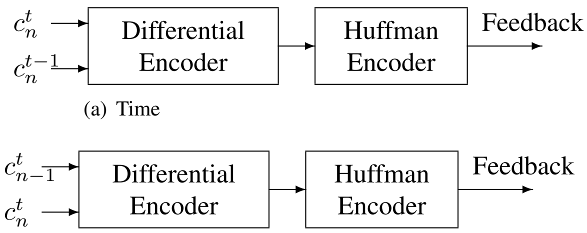


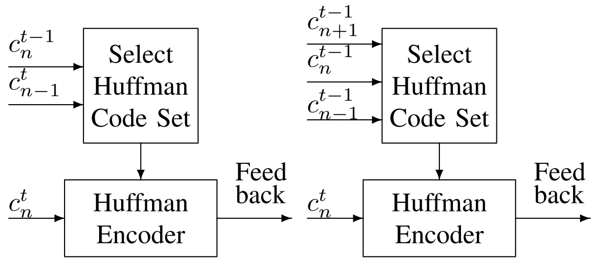
Fig. 2. Time and Frequency Correlation Probabilities for UMTS Vehicular A channel



(a) Time

(b) Frequency

Fig. 3. Time and Frequency Correlation Encoded Systems



(a) Iterative

(b) Block

Fig. 4. Time-Frequency Correlation Encoded Systems

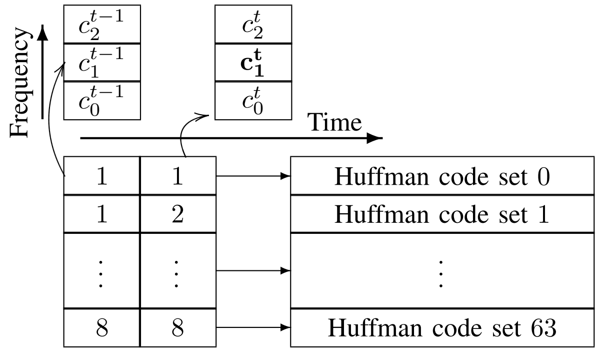


Fig. 5. Iterative Time-Frequency

symbol are shown). Thus we will obtain compression gains. It can also be observed that for a UMTS Vehicular A channel [17], the frequency correlation is higher than time correlation, and therefore it is likely that frequency correlation algorithm will obtain better compression gains than the time algorithm. Thus, these two algorithms will be denoted as time and frequency correlation respectively. These algorithms are depicted in Fig. 3, where c_n^t denotes the modulation scheme on nth chunk at tth frame-time.

C. Time-Frequency Correlation Algorithms

Since both time and frequency correlations are large, it would be interesting to be able to leverage on both at the same time in order to extract all the possible information and thus, reduce as much as possible the information needed. In this paper two algorithms that exploit time and frequency correlation properties at the same time are developed and analyzed. The first one uses the modulation scheme at the previous chunk and the actual modulation in adjacent chunk for encoding the actual chunk's modulation (Fig. 4a). The decoding process should be performed iteratively because we need adjacent chunk for decoding and therefore this scheme is denoted as *Iterative Time-Frequency (ITF)*. The second one uses the modulation schemes on the previous and adjacent chunks (Fig. 4b). The decoding process can be performed block-like and thus this algorithm is denoted as *Block Time-Frequency (BTF)* through the paper.

1) Iterative Time-Frequency Algorithm (ITF): As mentioned above, in this algorithm we use the modulation scheme in the previous chunk (c_n^{t-1}) and the actual adjacent chunk (c_{n-1}^t) for selecting the set of Huffman codes to encode the current chunk (c_n^t). In this way we have to design n_{mod}^2 Huffman code sets[2], where n_{mod} is the number of available modulation schemes. The reason to do that is the following, if for example on the previous chunk (c_n^{t-1}) the modulation was BPSK and in the current time in the adjacent chunk (c_{n-1}^t) is BPSK too, it is highly probable that the modulation in c_n^t will be BPSK or maybe QPSK, but it will be very unlikely that it will be higher than QPSK. In this way, the Huffman codes are designed with a very well conditioned data and the compression gains will be higher. A description of this algorithm is depicted in fig. 4a and 5 for a scenario with 8 available modulation schemes.

2) Block Time-Frequency Algorithm (BTF): This algorithm is similar to ITF with a small difference when selecting the Huffman code set. In this case, we use the previous chunk modulation scheme (c_n^{t-1}) as in ITF but instead of the actual adjacent chunk c_{n-1}^t we use adjacent previous chunks (c_{n-1}^{t-1} and c_{n+1}^{t-1}). Therefore we need to design n_{mod}^3 different Huffman code sets. The philosophy of this method is similar to ITF and it is depicted in fig. 4b and 6 in a scenario with 8 available modulation schemes.

[2] A Huffman code set is composed by n_{mod} Huffman codes, one for every combination.

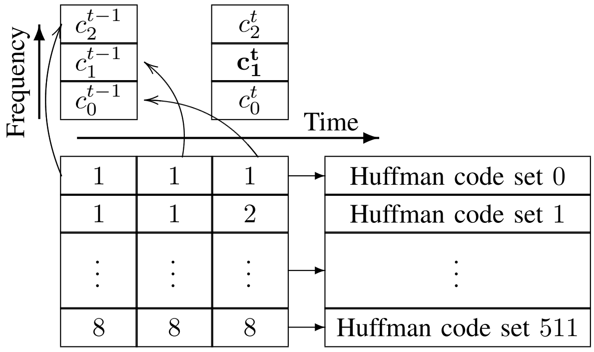


Fig. 6. Block Time-Frequency

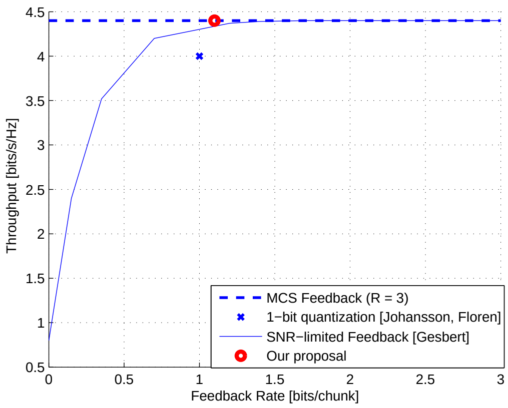


Fig. 7. Throughput as a function of the feedback rate for a multi-carrier system with 30 users for UMTS Vehicular A channel at 50 km/h@SNR = 20 dB

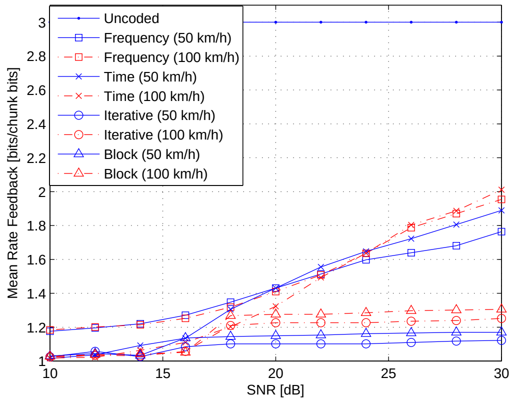


Fig. 8. Mean Feedback Rate as a function of SNR for UMTS Vehicular A channel at 50 and 100 km/h

It can be argued that the complexity of managing and storing all these codes may be high. However, once the Huffman codes have been designed the complexity is no more than to store into a lookup table, with two or three indexes (the correct one for the iterative and the block algorithm respectively) and the Huffman encoder complexity, which is very low too. Moreover, because of the symmetry of the problem, the number of codes to be stored can be significantly reduced. In order to obtain the joint probabilities and be able to design Huffman code sets, long simulations have been carried out.

IV. Simulation Results

Once the algorithms have been described, in this section some results of their performance will be shown. Monte Carlo simulations have been carried out with 10.000 frames over a UMTS Vehicular A channel [17]. The system summarized in section II [4] has been simulated with 30 active users at speed of 5, 50 and 100 km/h. The channel has been assumed to be perfectly estimated and predicted at the receiver. Two different scenarios have been analyzed: In one of them, the feedback channel has no errors, and in the other one, an AWGN (Additive White Gaussian Noise) channel has been implemented for evaluating the systems's robustness.

First in Fig. 7 a performance comparison between different algorithms in the literature is shown. In this figure, the throughput in terms of bits per second per hertz depending on the mean number of bits per chunk is ploted for the algorithms in [10], [14] and [9], and they are compared to our proposal. It can be seen that our proposal offers good trade-off between performance and compression rates. Besides, as it has been discussed in the previous section, the complexity is almost negligible (once the codes have been designed).

The compression performance in these algorithms depends on the SNR, the terminal speed and the delay spread of the channel. Thus, in order to evaluate it, in figures 8 and 9 the comparison for the four compression algorithms in this paper are drawn. The free parameters were a terminal speed of 5, 50 and 100 km/h, SNR ranging from 10 to 30 dB, and two different channel models: Pedestrian A and Vehicular A. It can be seen that algorithms exploiting both correlations (time and frequency) at the same time obtain much better results than the ones that only explore time or frequency correlations. It should be noted that the results in both figures have been obtained designing different Huffman code sets for every scenario, i.e depending on the estimated SNR and the Doppler spread, the adequate codes set is selected. Moreover, utilizing these techniques jointly with opportunistic feedback, we can reduce the feedback rate by a factor of 6 at best.

Up to now, we had assumed that there were no errors in the feedback channel but this is not the common situation. In order to evaluate the algorithm's performance in a more realistic scenario we have simulated them utilizing the same parameters as before but including AWGN in the feedback link. Since errors will cause that feedback data degradation (and therefore the throughput reduction,) we have introduced a refresh burst every Nth frame, i.e. we send uncoded information every Nth frames for refreshing data at the transmitter. In Fig. 10 these results are shown for the Block and the Iterative

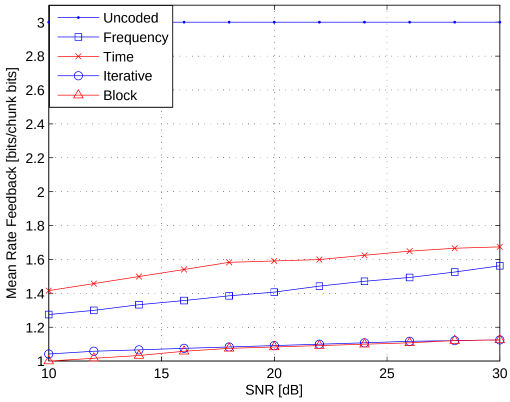


Fig. 9. Mean Feedback Rate as a function of SNR for UMTS Pedestrian A channel at 5 km/h

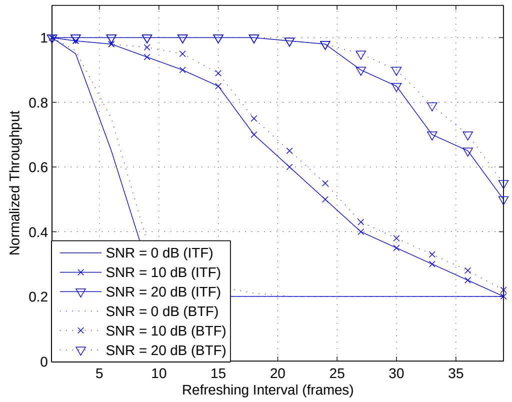


Fig. 10. Mean Throughput Loss due to errors in feedback channel at 100 km/h

algorithm and for different SNRs in the feedback channel. Throughput in the figure has been normalized to the maximum if no errors would have occurred. It can be seen that both algorithms perform similar in presence of errors, although BTF is slightly more robust against errors in feedback channel. The reason may be because it uses three chunks (instead of two as ITF does) for selecting the adequate Huffman code set. It can also be observed that for a reasonable signal to noise ratio the refresh rate may be $15 - 20$ frames. It should be noted that the lower the frame refresh rate the better performance but also the lower compression rates. It has been shown in [12] that for a refresh rate larger than 15 the loss in compression gain is rather small.

V. Conclusions

In this paper two algorithms for compressing feedback data exploiting the inherent frequency and time correlation properties of wireless mobile channels have been developed and analyzed. By using them jointly with the opportunistic feedback the amount of feedback data is significantly reduced, and thus adaptive modulation in multi-carrier systems can be implemented. It has also been shown that since the feedback information is an important information, some refresh frames must be sent every certain time in order to avoid error propagation in the decoding process. A reasonable value for the refresh interval may be 15 - 20 frames.

References

[1] E. G. Larsson, "On the combination of spatial diversity and multiuser diversity," *IEEE Comunications Letters*, vol. 8, no. 8, pp. 517–519, August 2004.

[2] X. Qui and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE trans. on Communications*, vol. 47, no. 6, pp. 884–894, June 1999.

[3] E. Dahlman, P. Frenger, J.-C. Guey, G. Klang, R. Ludwig, M. Meyer, N. Wiberg, and K. Zangi, "A framework for future radio access," in *Proc. IEEE 61st Vehicular Technology Conference. VTC Spring'05. Paving the path for Wireless Future.*, vol. 1, 30-31 May. 1 June 2005.

[4] M. Stenard, T. Svensson, and G. Klang, "WINNER MAC for Cellular Transmission," in *Proc. of IST Mobile Summit*, Mikonos, Grece, June 2006.

[5] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE International Conference on Communications (ICC)*, vol. 1, Seatle, June 1995, pp. 331 – 335.

[6] W. Wang, T. Ottosson, M. Sternad, anders Ahlen, and A. Svensson, "Impact of multiuser diversity and channel variability on adaptive OFDM," in *Proc. of IEEE Vehicular Technology Conference (VTC Fall)*, vol. 1, September 2003, pp. 547 – 551.

[7] D. Gesbert and M.-S. Alouini, "How much feedback is multiuser diversity really worth?" in *Proc. IEEE International Conference on Communications (ICC)*, vol. 1, June 2004, pp. 234 – 238.

[8] T. Eriksson and T. Ottosson, "Compression of Feedback for adaptive modulation and scheduling," 2007, to be published in Proceeding of the IEEE.

[9] M. Johansson, "Diversity-enhanced equal access - considerable throughput gains with 1-bit feedback," in *Proc. of IEEE Workshop on Signal Processing Advances in Wireless Communications (WSPAWC)*, Lisboa. Portugal, 2004.

[10] F. Floren, O. Edfors, and B. A. Molin, "The effect of feedback quantization on the throughput of a multiuser diversity scheme," in *Proc. of IEEE Global Conference on Communications (GLOBECOM)*, 2003.

[11] T. Lestable and M. Bartelli, "LZW adaptive bit loading," in *Proc. of International Symposium on Advances Wireless Communications (ISAWC)*, 2002.

[12] V. P. G. Jiménez and A. G. Armada, "An Adaptive MIMO OFDM system: Design and Performance evaluation," in *Proc. IEEE International Symposium on Wireless Communication Systems*, September 2006.

[13] H. Nguyen and T. Lestable, "Compression of bit loading power vectors for adaptive multi-carrier systems," in *Proc. of 47th IEEE International Midwest Symposium on Circuits and Systems (IMSCS)*, vol. 3, 2004, pp. 243 – 246.

[14] V. Hassel, M.-S. Alouini, G. E. Øien, and D. Gesbert, "Rate-optimal multiuser scheduling with reduced feedback load and analysis of delay effects," *EURASIP Journal on Wireless Communications and Networking*, pp. Article ID 36424, 7 pages, 2006, doi:10.1155/WCN/2006/36424.

[15] P. Svedman, S. K. Wilson, L. J. C. Jr, and B. Ottersen, "A simplified oportunistic feedback and scheduling scheme for OFDM," in *Proc. of IEEE Vehicular Technology Conference (VTC Spring)*, vol. 1, 2004, pp. 1878 – 1882.

[16] A. Duel-Hallen, H. Hallen, and T.-S. Yang, "Long range prediction and reduced feedback for mobile radio adaptive ofdm systems," *IEEE trans. on Wireless Communications*, vol. 5, no. 10, pp. 2723 – 2733, October 2006.

[17] "Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000," Recommendation ITU-R M.1225, Tech. Rep., 2000.