Linköping University Post Print

On Coding of Scheduling Information in OFDM

Jonas Eriksson, Reza Moosavi, Erik G. Larsson, Niklas Wiberg,
Pål Frenger and Fredrik Gunnarsson

N.B.: When citing this work, cite the original article.

On Coding of Scheduling Information in OFDM

Jonas Eriksson[†], Reza Moosavi[†], Erik G. Larsson[†], Niclas Wiberg[‡], Pål Frenger[‡], and Fredrik Gunnarsson[‡]

[†]Dept. of Electrical Engineering (ISY), Linköping University, Linköping, Sweden. Email: {joner,reza,egl}@isy.liu.se

[‡]Ericsson Research, Linköping, Sweden. Email: {niclas.wiberg,pal.frenger,fredrik.gunnarsson}@ericsson.com

Abstract—**Control signaling strategies for scheduling information in cellular OFDM systems are studied. A single-cell multiuser system model is formulated that provides system capacity estimates accounting for the signaling overhead. Different scheduling granularities are considered, including the one used in the specifications for the 3G Long Term Evolution (LTE). A greedy scheduling method is assumed, where each resource is assigned to the user for which it can support the highest number of bits. The simulation results indicate that the cost of control signaling does not outweigh the scheduling gain, when compared with a simple round-robin scheme that does not need signaling of scheduling information. Furthermore, in the studied scenario, joint coding and signaling of scheduling information over all selected users is found to be superior to separate coding and signaling for each user. The results also indicate that the scheduling granularity used for LTE provides better performance than the full granularity.**

I. Introduction

Many technical solutions for mobile broadband multi-user wireless access, such as 3G Long Term Evolution (LTE) and WiMAX, use orthogonal frequency-division multiplexing (OFDM) in the downlink. OFDM allows for scheduling of the different users both in different timeslots and in different frequency slots, by assigning them different subsets of OFDM sub-carriers in different OFDM symbols. Such scheduling can exploit multi-user diversity both in the time- and in the frequency-domain and thus achieve a high system throughput. In principle, users are scheduled to receive data in the specific frequency slots which are the most beneficial to them. Such opportunistic scheduling can increase the system throughput, and it is used to achieve high data rates in forthcoming mobile broadband access systems.

The gain from opportunistic scheduling comes at a cost, though: the receiving users must be told in what time- and frequency slots their data are located. Conveying this information takes a lot of control signaling. For instance in a LTE system [1] up to 20% of the channel resources are set aside for control signaling and a major portion of this is signaling related to the scheduling. The finer the granularity of the scheduling is, the better one can tune the scheduling to each channel realization. However, a finer granularity also means more control signaling, and transmitting this information will consume resources that could otherwise have been used for payload data. Hence, a tradeoff is implicit. An additional complication is that the scheduling information of different

users is correlated, since no two users can be scheduled in the same time/frequency slot. How can one efficiently communicate control information with such an inherent structure, and how can one quantify the cost of transmitting this control information? In this work, we formulate a basic framework for studying this fundamental problem, and provide some initial results.

A. Previous work

The control signaling in cellular mobile telephone systems is a problem that has been around for some time, but which has attracted relatively little attention. Some early work (e.g. [2]) focused on the optimization of call-setup channels in high-capacity mobile telephone systems. Later work [3] has focused on optimizing the ratio between the number of channels dedicated for control signaling and for traffic, respectively. When it comes to the specifics of time/frequency-scheduled OFDM systems, where the control channel dynamically competes with the payload data for resources, there are a few related results available. For instance, within the FP6-WINNER [4] project, suggestions for control channel design were presented in [5]. The solution in [5] suggests broadcasting a small amount of scheduling data to the mobile users, clustering them into groups according to channel quality, and then use a separate link adaptation mechanism for each such group.

B. Contributions

We are interested in the fundamental limits for the coding of control signaling (scheduling information) and its associated cost. The specific contributions of our work are as follows: (i) We formulate a system model that captures all relevant physical phenomena (path loss, shadow fading, time- and frequency-selective fading) to the first order. (ii) We formulate a performance measure for the effective system capacity, that is, the system capacity taking into account the channel resources consumed by the control signaling. (iii) We evaluate empirically the relative performance of different control signaling strategies, using the proposed system model and performance measure.

II. System model

We assume that we have a base station surrounded by a random number of mobile units that want to be scheduled for reception of payload data. The users are assumed to be greedy, wanting as much throughput as they can get. That is, we are studying a maximal load case for the system. The communication method is OFDM with N_c subcarriers occupying a
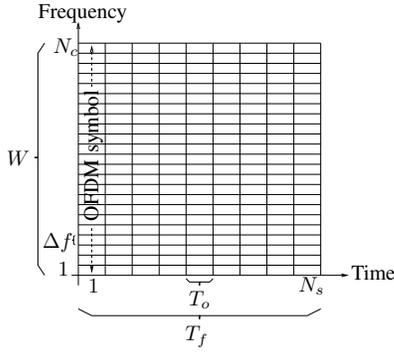
Fig. 1. OFDM time/frequency resource grid

bandwidth of $W = N_c\Delta f$ Hz, where Δf is the subcarrier spacing. Each subcarrier is narrow enough to be treated as flat fading. The time axis is divided into scheduling timeframes where the users receive new scheduling information at the start of each frame. The duration of a scheduling frame is T_f and can equivalently be expressed in terms of the OFDM symbol length as $T_f = N_s T_o$, where T_o is the duration of one OFDM symbol and N_s is the number of OFDM symbols per frame. For each frame, the channel resources can be thought of as an $(N_c \times N_s)$ grid of time/frequency *resource elements*. See Figure 1.

During a specific scheduling time interval, we denote the gain of the jth subcarrier at time instance (OFDM symbol number) i for user k by $h_{i,j}^{(k)}$. User k then experiences a fading channel summarized in the $(N_c \times N_s)$ matrix $H^{(k)} = \{h_{i,j}^{(k)}\}$. We assume that all these channel matrices are known to the base station. Based on the collection of channel matrices for a certain scheduling frame, the base station makes its scheduling decision for that interval. The scheduling decision is expressed in the terms of a *scheduling matrix* $U = \{u_{i,j}\}$ with integer entries, that indicate which of the users that are scheduled in each resource element. The scheduling procedure produces an *effective system channel* represented by the matrix S with entries $s_{i,j} = h_{i,j}^{(u_{i,j})}$.

We consider a single cell, served by one base station. The population of mobile users in the cell during any given frame is governed by a binomial distribution, and on the average a fraction p_u of a total number of potential users N_p are requesting service from the base station. The realization of the number of users in a specific time frame is denoted by N_u. The N_u active users in a specific scenario are randomly placed according to a uniform distribution in the circular ring bounded by an inner radius R_0 and an outer radius R_c, as illustrated in Figure 2. A few basic assumptions are made regarding the identification of the N_u individual users that request channel resources in a frame. We assume that all users have been given a unique identification number which is an integer in the range $0, 1, \ldots, N_u - 1$. We also assume that the users have knowledge of the total number of users requesting channel resources in a specific frame (N_u).

The channels of the different users are affected by large-scale (path-loss and shadowing) and small-scale (multipath) fading factors. The large-scale fading factors are constant over
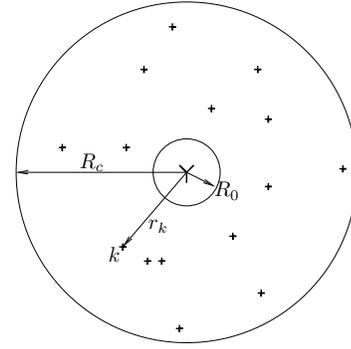

Fig. 2. Illustration of the system under consideration.

time and frequency and we model them via a multiplicative factor $(r_k/R_0)^{-\alpha}10^{\frac{\chi}{10}}$ where r_k is the distance to the base station for user k, R_0 is a reference distance, α is the path loss parameter, and χ is drawn from a normal distribution $N(0, \sigma)$ where σ is the shadow fading standard deviation in dB. The small-scale fading factors vary with time and frequency. To model this we use a tapped-delay line model for the channel impulse response. The tap coefficients are modeled as Rayleigh fading stochastic processes with a Jake's Doppler spectrum, but they are assumed to be constant over the duration of one OFDM symbol interval.

We define the *system operating point* in terms of the SNR conditions on the cell border, and we restrict the study to a noise limited system. Consider a hypothetical mobile receiver located at the reference distance R_0 from the base station. We denote the power received by this mobile, in the absence of fading, by P and we call it the *normalized transmit power*. Then the average (over the fading) signal-to-noise ratio experienced by a user at distance r from the base station is

$$\text{SNR}(r) \triangleq \frac{P}{N_c\Delta f N_0}\left(\frac{r}{R_0}\right)^\alpha$$

where N_0 is the noise spectral density at the receiver frontend, and $N_c\Delta f$ is the total system bandwidth. We define the system operating point to be the average SNR at the cell border ($\text{SNR}(R_C)$).

III. PERFORMANCE MEASURE

To compare the performance of different scheduling methods with different signaling granularity and different control signaling strategies, we need a performance measure that is easy to evaluate. For a given effective system channel S, we propose the following spectral efficiency performance measure:

$$C(S) = \max_{\sum_{i,j} P_{i,j} \leq N_s P}\left\{\frac{1}{N_c N_s}\sum_{i,j}\log\left(1 + \frac{P_{i,j}|s_{i,j}|^2}{\Delta_f N_0}\right)\right\}$$

[bits/s/Hz] (1)

where $P_{i,j}$ is the transmit power used during the (i, j)th resource element. The maximum in (1), which constitutes the optimal distribution of power over time and frequency,

is computed by classic waterfilling. The interpretation of the summation in (1) is the same as in the definition of ergodic capacity [6]: we assume that there is capacity-achieving coding across the OFDM subcarriers (and across time), and $\log\left(1+\frac{P_{i,j}|s_{i,j}|^2}{\Delta f N_0}\right)$ then measures the mutual information that flows in the (i,j)th resource element.

To take into account the cost for transmitting the scheduling information, we omit from the summation in (1) the indices corresponding to channel resources occupied by the control signaling. In this work we consider the control signaling to be error-free, but for future work our proposed measure is well suited to consider the effects of errors in the control signaling. For instance, errors in the scheduling information of a specific user would typically mean that the capacity associated with the resource elements scheduled for that user is lost. This approach also gives us the possibility of studying per-user rates if we so desire.

The main advantage of the measure in (1) is that it captures the fundamental way in which rate depends on power and bandwidth, without making assumptions on specific channel codes or modulation formats. Of course, the absolute spectral efficiencies predicted by (1) are not achievable in practice (this would require the use of a capacity-achieving code and infinitely long data blocks) but it is a useful tool for relative comparisons of different signaling strategies [6]. It should be pointed out that this measure is biased towards an overall capacity-maximizing approach, whereas not all scheduling methods have that objective. Therefore, care should be taken when comparing the performance of different schedulers using this measure. However, the measure should be useful for studying the impact of different signaling strategies for one given type of scheduler.

IV. MAIN STUDY PARAMETERS

There are three main parameters that will affect our results and conclusions: the scheduling granularity, the scheduling method, and the control signaling strategy.

A. Scheduling granularity

The scheduling granularity determines how small part of the time/frequency grid that can be allocated to a specific user. At one extreme, a user either gets the entire resource grid (all i,j). That is, the entire frame is allocated to only one user (*Min* granularity). At the other extreme, each user can get an arbitrary set of individual resource elements (i.e., one or many pairs (i,j)) (*Full* granularity). Between these two extremes there are many other possibilities. To take a relevant example, in the 3G-LTE standard [1] the smallest possible scheduling granularity is 12 OFDM subcarriers in frequency and 14 OFDM symbols in time, spanning the whole scheduling frame (*LTE* granularity). In this study we will focus on these three different scheduling granularities.

B. Scheduling method

The scheduling method determines how the base station decides to which user it should give each resource element. We will consider two methods:

1) Round-robin: The channel resources are allocated to the different users cyclically, according to a fixed predetermined order, and without regard to their channel qualities. Under the assumptions made and if the users know their identity numbers, round-robin does not need any control signaling to convey scheduling decisions. (From a performance point of view, resources may as well be assigned at random, but this would require more signaling.)

2) System maximizing: This scheduling strategy creates the system channel S which maximizes the total system throughput. For the case of by-resource-element scheduling, we use the following approximation for this scheduler to produce the scheduling matrix U. For each resource element (i,j) we set

$$u_{i,j} = \arg \max_{1 \le k \le N_u} |h_{i,j}^{(k)}|.$$

For the other two granularities we create the matrix U as above and then simply let a user decide who gets a specific scheduling block by a majority vote.

For the by-resource-element-scheduling, we estimate the number of bits that a compressed version of U will require by empirically estimating the conditional entropy of a single position in the scheduling map, given its local surrounding when all symbols in the map are mapped onto the integers $0, 1, \ldots, N_{sched}$, and where N_{sched} is the number of users being granted channel resources by U. We do the following

$$H(U) \approx H\left(\square \middle| \begin{smallmatrix}\square\square\square\\\square\square\end{smallmatrix}\right) = H\left(\begin{smallmatrix}\square\square\square\\\square\square\square\end{smallmatrix}\right) - H\left(\begin{smallmatrix}\square\square\square\\\square\square\end{smallmatrix}\right)$$

where the two terms on the right are estimated from the histogram over the respective patterns in U. Our estimate for the number of bits needed to transmit U is $\lceil H(U)N_c N_s \rceil + \lceil \log_2(N_u - 1) \rceil + N_{sched}\lceil \log_2(N_u - 1) \rceil$, where the last two terms stem from the need to supply a translation table for the integers $0, 1, \ldots, N_{sched}$ to the correct subset of the integers $0, 1, \ldots, N_u$.

For the minimum granularity case we have that since one user gets all the channel resources we only need to transmit the identification number of that user. For this we allocate $\lceil \log_2(N_u - 1) \rceil$ bits.

When using LTE granularity, the scheduling decision is a $N_c/12$ long vector. Estimating the entropy of this vector is not a good idea since we typically deal with very little data to build our statistics from, and the results will be unrealistically small. Instead we adopt a simple run-length encoding scheme. We use $\lceil \log_2(N_u - 1) \rceil$ bits per symbol s_i in the run-length code to describe the symbol value (user identity) and for each instance i we indicate the length of the run l_i with $\lceil \log_2(N_c/12 - l_{i-1} - 1) \rceil$ bits, with $l_0 = 0$. If the last symbol has a run-length of one we omit stating this run-length in the codeword.

Finally, we let M be the total number of bits required for the scheduling information. Note that in the case that there is only one user in the cell ($N_u = 1$), $M = 0$. Such a user will under our assumptions know that it is alone in the cell, and it will therefore greedily seize the whole scheduling frame.

3

C. Control signaling strategy

The control signaling strategy determines the way in which we convey the control information to the different users. This mechanism is the primary target of our study. A few assumptions are made on the basic structure of the control signaling. We assume that the control signaling region is placed in the first consecutive OFDM symbol(s) of the scheduling frame and that consecutive sub-carriers are allocated for the control signaling starting from sub-carrier 0.

As a theoretical baseline, we consider a *magic genie (MG)*, which conveys the scheduling decisions to the different users without spending any channel resources at all. This is clearly not realizable but it provides an upper bound on performance.

As a starting point for cost setting the control signaling of scheduling information we use the following strategy. The number of resource elements occupied by the control signaling is determined based on the number of bits that a specific resource element can carry. This is evaluated in the following way. For a specific resource element with instantaneous signal-to-noise ratio $\gamma_{i,j}$, we set the number of carried bits to $\log_2(1 + \delta\gamma_{i,j})$ so that the control region in total carries

$$N_{bits} = \left\lfloor \sum_{i,j \in \text{control region}} \log_2(1 + \delta\gamma_{i,j}) \right\rfloor.$$

bits. We recognize this as the conventional capacity formula for an ergodic Rayleigh fading channel, but where the SNR has been scaled down with a penalty factor δ. This penalty factor models how far from the capacity limit the code is able to operate. We will use $\delta = -3$ dB (corresponding to a very good error correcting code), and at the other extreme $\delta = -10$ dB (corresponding to poor or no coding).

The control region is expanded dynamically until it accommodates the number of bits, M, needed for the specific scheduling decision to be properly propagated over the control channel.

The basic assumptions for the control signaling structure in this study is that only resource element scheduling information needs to be transmitted over the control channel, and that this is done via error-free transmission and near entropy-reaching compression of the scheduling information. This strategy is not realistic, but it does acknowledge the fact that conveying scheduling information costs channel resources and gives a lower bound on how cheap this can get. In a practical system other information apart from the resource element scheduling needs to be conveyed to the mobile users, for example transport formats and HARQ information. However, the resource element scheduling information constitutes a substantial part of the total, and is therefore a relevant study object in itself.

The problem splits into two contrasting ways of conveying the scheduling information to the scheduled users. Either broadcast all scheduling information jointly to all users utilizing fully the correlation in the scheduling information between users in the compression stage, obtaining a small amount of data to be transmitted. We denote this strategy *compress and broadcast (CBC)*. The in principle diametrically

Simulation parameters		Simulation parameters, cont.	
N_{MC}	100000	σ	6 dB
N_c	300	N_t	3
N_s	14	P_{tapp}	$[0, -3, -6]$ dB
Δ_f	15 kHz	f_d	200 Hz
T_s	71.429 μs	D_{tndp}	$[0, 0.5, 1]$ μs
R_0	50 m	d_j	0.3 μs
R_c	1500 m	N_p	100
α	4	p_u	0.1

TABLE I
SIMULATION PARAMETERS.

opposed method of conveying the scheduling information is to treat the scheduling information of one user separately from the scheduling information of all the other users in the compression stage and transmit it separately to each user. We denote it *compress and transmit separately (CTS)*. In the CBC case we utilize the full correlation in the scheduling information but we broadcast the information over a multi-user broadcast channel which in general displays poor channel quality. We construct the control channel for this case for a specific set A of scheduled users by the rule

$$k^* = \arg\min_{k \in A} |h_{i,j}^{(k)}| \implies h_{i,j}^{(control)} = h_{i,j}^{(k^*)}$$

for each resource element in the control region.

In the CTS case we make a poor job at utilizing the correlation in the scheduling information but we transmit separately over individual channels to the users. Simulations were performed to compare the two strategies.

V. SIMULATION RESULTS

Monte Carlo simulations for the round-robin scheduler and the system-maximizing scheduler for different scheduling granularities and different system operation points have been performed. The basic simulation setup is summarized in Table I, where parameters of the tapped-delay line model are included as well. N_t is the number of taps, P_{tapp} is the tap average power profile, f_d is the Doppler spread, D_{tndp} is tap nominal delay profile and d_j is the per-user tap delay jitter. As a system operation point comparison, the LTE system simulation evaluation description for macro deployments with three sector sites [7] corresponds to SNR(1500) = 20 dB in the main antenna direction and SNR(1500) = 0 dB in the direction of minimum antenna gain. Note that with the parameters of Table I, the average SNR difference between a point on the cell border and a point at the reference distance R_0 is about 60 dB. Also note that the system operation point will be even lower in a multi-cell deployment.

Figures 3 and 4 show the normalized empirical mean of the capacity measure (1) for the system-maximizing scheduler with full granularity as a function of SNR(R_C). The results for the CBC and CTS strategies for conveying scheduling decisions are depicted, normalized on the MG result. For comparison the results for the round-robin scheduler are also included. The graph reveals that in this case, CBC outperforms CTS. In Figure 5 the situation for the system-maximizing scheduler operating on LTE granularity is depicted in a similar way, including the MG results for LTE granularity, and the
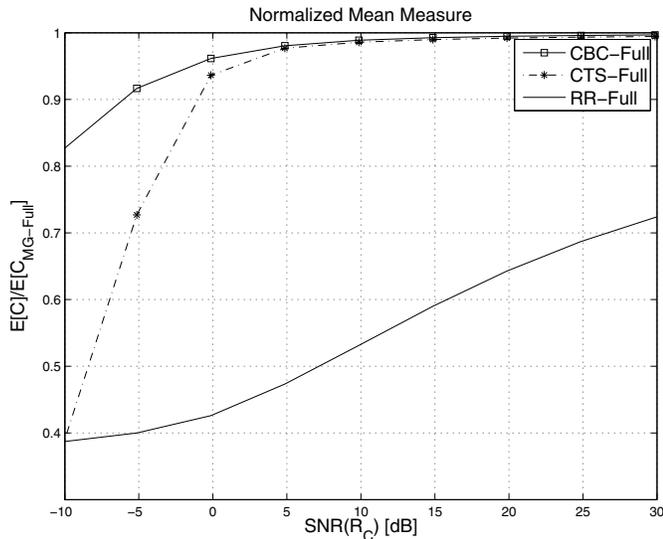
4

Fig. 3. Simulation results for a system-maximizing scheduler with full granularity. Mean of the capacity measure (1) with $\delta = 0.5$ for CBC and CTS. The cost of the scheduling control signaling is highlighted by normalizing the capacity results to the magic-genie case with no cost assumed for the signaling. The diagram shows that the relative cost of the signaling decreases with increasing channel quality, since the number of bits required to represent the scheduling information is independent of the channel quality. For comparison the results for the round-robin scheduler are also included.
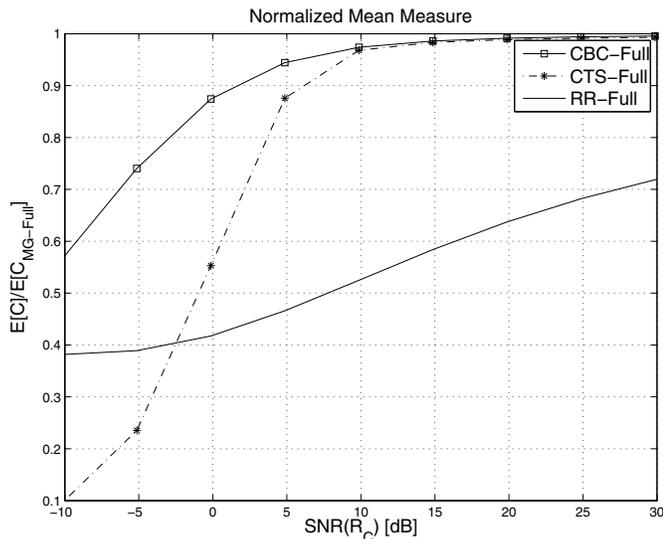

Fig. 4. Same as Figure 3 but with $\delta = 0.1$

conclusions are the same. Note that the results for LTE granularity are in Figure 5 normalized to the magic-genie results for full granularity. By comparing with Figure 3, it can be concluded that the LTE scheduling granularity outperforms the by-resource-element scheduling granularity. The reason is that LTE granularity is fine enough to tune to the channel in an acceptable manner but conveying the scheduling decisions are much less costly in terms of channel resources than for the by-resource-element granularity.
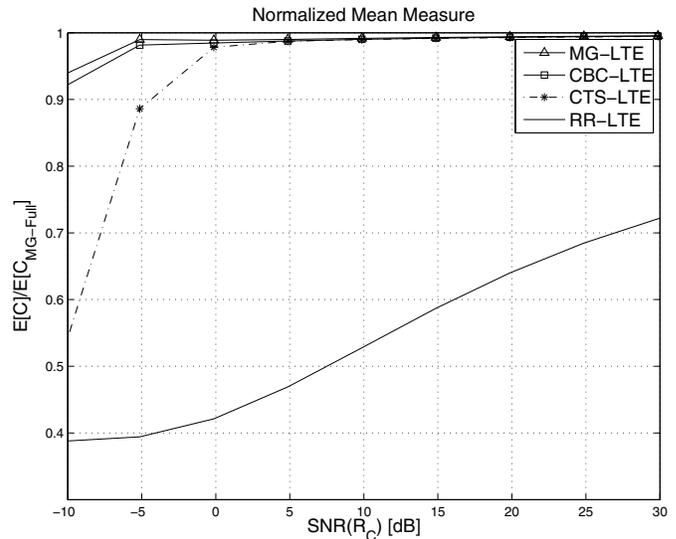

Fig. 5. Same as Figure 3 but for LTE granularity. The results are normalized against the magic-genie results for full granularity.

VI. CONCLUSION

We have proposed a framework for studying the fundamental limits associated with the signaling of scheduling information in time/frequency-scheduled OFDM systems. The methodology is exemplified by the study of joint versus separate coding and magic genie control signaling strategy for a system-maximizing scheduler. For comparison a round-robin scheduler was also included in the study. Our simulation results indicate that a system-maximizing scheduler negates much of the poor order statistics normally apparent in the multiuser broadcast situation, making joint coding of the scheduling information a plausible solution for this case. It should be duly noted that this result is due to the severe unfairness imposed by the system maximizing scheduling method. Starvation of users far from the base station is obvious and this of course makes the scheduling method highly unattractive for practical systems. We expect to refine the methodology presented here and study more attractive schedulers (e.g. proportional fair schedulers), for which we foresee severe problems with the broadcast strategy.

REFERENCES

[1] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution - HSPA and LTE for Mobile Broadband*, 2nd ed. Academic Press, 2008.
[2] S. Okasaka, "Control channel traffic design in a high-capacity land mobile telephone system," *IEEE Transactions on Vehicular Technology*, vol. 27, no. 4, pp. 224–231, Nov 1978.
[3] A. Srivastava, S. Srivastava, and K. Mitchell, "Determining traffic and control channels for a packet based cellular system having time-varying traffic," in *Proc. of IEEE VTC*, vol. 4, pp. 2167–2171, May–June 2005.
[4] G. Auer, M. Döttling, K. Doppler, J. von Häfen, N. Johansson, K. Kalliojärvi, T. Lestable, J. Luo, N. J., M. Olsson, A. Osserian, M. Sternad, T. Svensson, and C. Wijting, "WINNER II System concept Description," FP6 research program, Tech. Rep., 2008.
[5] M. Sternad, T. Svensson, and M. Dottling, "Resource allocation and control signaling in the WINNER flexible MAC concept," in *Proc. of IEEE VTC*, pp. 1–5, Sept. 2008.
[6] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
[7] 3GPP TS 25.814 "Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)".

5