# Channel Aware Deferring Strategies to Improve Packet Scheduling in OFDMA Systems

Israel Guío #1, Ángela Hernández-Solana#2, Vanesa Montero #3, Javier Lafuente-Martínez #4, Antonio Valdovinos #5

*Aragón Institute for Engineering Research, I3A, University of Zaragoza. C\ María de Luna, 3, 50.018. Zaragoza (Spain)*

*e-mails:1 iguio@unizar.es, 2 anhersol@unizar.es, 3 vmontero@unizar.es, 4 javierlm@unizar.es, 5 toni@unizar.es*

***Abstract*—In this paper, we analyze the performance gain achieved when a channel-aware deferring (CaD) strategy is applied to defer the allocation of users with transitory bad channel state in a mobile OFDMA system. Users with good channel conditions receive more resources and the radio resource utilization improves. To reduce the impact of the CaD strategy on the quality of service (QoS) provisioning for the deferred users, a delay-dependent criterion is also applied. We show that the CaD significantly reduces both the dropping rate and the mean delay for the most restrictive traffic in a multi-service scenario, under a certain power and subcarrier allocation (PSA) algorithm that efficiently controls the inter-cell interference (ICI).**

## I. Introduction

Orthogonal frequency division multiple access (OFDMA) has been selected as the multiple access technology for future mobile broadband cellular systems, as in fixed and mobile WiMAX and in LTE, due to its high resource allocation flexibility and scalability to accommodate a great number of users with different traffic demands and channel conditions. A frequency reuse of 1 deployment, i.e. all cells use the same frequency bandwidth, has extensively been proposed to achieve high spectral efficiency. However, ICI coordination (ICIC) schemes must be applied in order to improve user performance in the outer part of the cell. ICIC schemes define a set of restrictions on resource allocation, generally frequency and power, which must be taken into account when designing power and subcarrier allocation (PSA) algorithms. The aim of the joint ICIC and PSA operation is to achieve better interference performance and resource utilization.

Fractional frequency reuse (FFR) is a widely accepted ICIC scheme in OFDMA networks with frequency reuse of 1 [1-3]. The assignment flexibility it provides for cell-center users increases the ICI fluctuations, since the strongest ICI can come from any sector of the neighboring cells. Besides, as shown in [4] for low-loaded scenarios, quick changes in frequency resource occupation can be seen as an ON-OFF effect in transmit power. Thus, ICI shows strong fluctuations which yields to high channel state estimation inaccuracy. This fact raises the block error rate (BLER) and eventually increases the packet dropping rate and transmission delay as a result of the erroneous data block retransmissions. To overcome these drawbacks, in [4] we proposed a PSA algorithm, denoted as *FFRopa*, that improves the ICI stability and reduces both the dropping rate and the mean delay by applying a coordinated ordering in the allocation of the frequency resources.

Users which suffer from temporal bad channel conditions require more resources to deliver their data packets within their QoS requirements. Users with bad channel state experience higher BLER and, consequently, they need more retransmissions with robust modulation and coding schemes (MCSs) to successfully deliver their data packets. These requirements may starve other users of resources in high-loaded conditions, increasing their packet dropping rate and mean delay. To prevent this fact, we propose to defer the allocation of users with bad channel conditions (namely those users demanding the lowest MCS in the system) in the early steps of the allocation procedure to save more resources to the other users. Users with good or reasonable channel conditions are allocated on the frequency resources where they observe the best channel conditions, so they can use efficient MCSs, and the remaining resources are finally allocated to deferred users. However, this channel-aware deferral should only be applied when the QoS requirements of the deferred user are not at risk. In this paper, we provide a delay-dependent criterion to decide whether a user could be deferred or not. The performance of such procedure is analyzed in a multi-service scenario for three different time domain (TD) packet scheduling policies: Modified Largest Delay First (M-LDF), Proportional Fair (PF) [5] and Modified Largest Weighted Delay First (MLWDF) [6].

## II. Channel- and QoS-Aware Scheduling

Frequency and time domain packet scheduling, along with the power allocation and the adaptive MCS selection, are located at the base station (BS). Each cell independently performs its scheduling decisions taking into account the constraints imposed by the ICIC scheme. Following the LTE recommendations, the system frequency resources are divided into $N_{RB}$ resource blocks (RBs), which is the minimum allocable resource unit. An RB comprises $N_{SC}$ subcarriers for $N_T$ consecutive OFDM symbols (data symbols of one subframe/TTI), and the same transmit power and MCS are applied to all the subcarriers within an RB. Two transmit power constraints (defined by standard recommendations [7]) limit the maximum transmit power per RB ($P_{RB}^{max}$) and the maximum transmit output power in the overall cell ($P_T^{max}$). Thus, if $P_i$ is the transmit power on the RB $i$ and $P_{out}$ is the total transmit power in the cell, these constraints impose that:

$$P_{out} = \sum_{i=1}^{N_{RB}} P_i \le P_T^{max} \quad with \quad P_i \le P_{RB}^{max} \tag{1}$$

The transmit power level per RB can firstly take two values: $M^0_{med}$ in the inner subband and $M^1_{med}$ in the outer subband, which are computed so that the $P_{out}$ in a sector never exceeds $P_T^{max}$, even when it assigns all its allocable RBs: the whole inner subband plus its outer subband portion. However, a joint power and MCS adaptation according to the user's channel state will be suggested in order to improve the overall system efficiency. In order to limit the ICI variability, the allocated transmit power is suggested to be confined within a minimum mask ($M_{min}$) and a maximum mask ($M_{max}$), which can easily be computed as a decrement $\Delta_{min}$ or increment $\Delta_{max}$ from $M_{med}$.

Channel state information (CSI) is reported by each user $k$ using Channel Quality Indicator (CQI) messages, including the channel to interference and noise ratio ($CINR_i^k$) on each RB $i$. $CINR_i^k$ is computed as $CINR_i^k = h_i^k / (N^k + I_i^k)$, where $h_i^k$ is the estimated channel gain on RB $i$ for user $k$, $N^k$ the thermal noise and $I_i^k$ the ICI power on RB $i$. Multiplying this CINR value by the corresponding transmit power $P_i$, we compute the expected mean SINR on each RB.

According to the user's channel conditions, the target MCS ($MCS_k$) is defined as the lowest MCS that allows transmitting the user's required data rate $T_{req,k}$ in the minimum number of RBs $n_k$ (in this case $n_k=1$), always satisfying the SINR requirements. The resource allocation is initially performed assuming this target MCS. Once frequency resources and transmit power have been allocated to serve the required data rate, we allocate more resources to those users with remaining data in order to improve their QoS and the spectrum efficiency.

### *A. ICIC Scheme. FFRopa*

#### *1) ICIC Framework*

The FFR ICI coordination scheme provides a reuse of 1 subband for cell-center users, whereas a reuse of 1/3 is defined in the outer part of the cell to reduce the ICI power for cell-edge users. This solution provides high assignment flexibility but yields strong ICI fluctuations and an interference ON-OFF effect. To avoid these drawbacks, the *FFRopa* algorithm confines the transmissions of cell-center users on a reduced set of RBs, so the ICI power is more stable and thus, SINR estimation errors decrease. The overall frequency resources are allocated following the order depicted in Fig. 1. Inner subband is divided into 3 "sector subbands", each one intended for one sector of the cell. In this way, users are preferably allocated on the RBs intended for their serving sector although all RBs are really accessible. Neighboring cells start their allocations on different RBs, depicted as a shaded circle in Fig. 1, so ICI is avoided in very low-loaded scenario. At the beginning, a number of RBs, denoted as Stat Group (SG) and estimated from previous allocations, is not allocated following the ordering: each user gets the RB where it has the best channel conditions. The SG avoids the blockage due to the occupation ordering at early allocations as there are very few allocable RBs. The number of RBs that comprises the SG of sector s (SGS) must be carefully chosen to ensure that the transmissions are really confined, so we propose to set SGS to the number of RBs allocated in the previous TTI by this sector $s$. Out of these $SG_S$ RBs, the allocations within each sector proceed circular-wise as in Fig. 1 until all the RBs of its preferred subband are allocated.

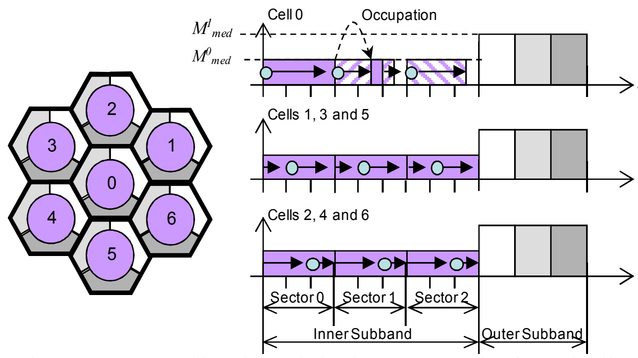


Figure 1. Resource allocation ordering in FFRopa (cell 0 shows details on resource occupation)

When a user must be scheduled but the preferred subband of its serving sector is completely assigned, the scheduler starts occupying RBs intended for other sectors of the cell. We propose to occupy RBs from the sector with the minimum number of active users in this TTI, as it will likely require fewer RBs to serve them. The frequency occupation ordering must be followed in this case as well, so only the first free RB of the occupied sector is allocated. An example of an RB occupation is depicted in Fig. 1 for cell 0, where sector 0 occupies the first free RB from sector 1 (solid-shaded). After this occupation, sector 1 proceeds allocating the following RBs to its own users according to the occupation ordering.

FFRopa also performs a power adaptation procedure in the inner subband to allocate more users at the expense of a small increase in ICI power variability. An example of this power adaptation is depicted in Fig. 2. The required power in the RB $i$ for user $k$ ($P_i^k$) is computed as $P_i^k = SINR_{req} / CINR_i^k$, where $SINR_{req}$ is the desired SINR in reception and $CINR_i^k$ is the channel to interference and noise ratio. Users demanding less power than $M_{med}$ (users 0, 2 and 3 in Fig. 2) create a power margin ($M_{med} - P_i^k$) that can be used to boost the transmissions for users with bad channel conditions (users 1 and 4 in Fig. 2), never exceeding $M_{max}$. However, in order to reduce the transmit power and hence ICI fluctuations, the allocated power is low-bounded by a $M_{min}$ value. Thus, when the required power falls below $M_{min}$, the scheduler actually considers that the required power is $M_{min}$ (user 2 in Fig. 2). On the contrary, when the required power exceeds $M_{max}$ in a certain RB, the user can not be allocated in that RB. The power adaptation is only performed in the inner subband, since the outer subband offers a very low power margin and we prefer to stabilize the ICI power transmitting always with constant power.

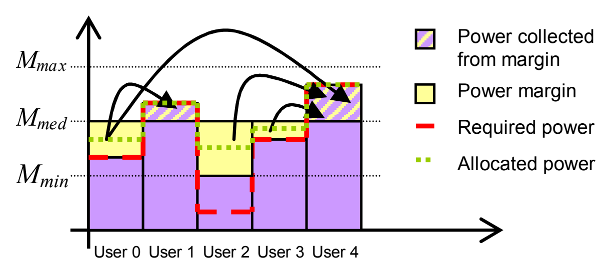


Figure 2. Power adaptation in FFRopa

Although the power margin can be immediately taken from users with the best channel state, we think that a better solution is obtained when the power margin is proportionally collected from power margins of all the allocations with good channel conditions at the end of the allocation procedure. The achievable power margin ($M_{med}$ - $P_i^k$) is continuously updated when a user with good channel conditions is allocated. However, this power margin is not taken from the user and then, an $M_{med}$ power level is preliminary allocated. Users whose required power exceeds $M_{med}$ will be allocated if their power excess ($P_i^k$ - $M_{med}$) is lower than the achievable power margin at that moment, that is to say, the power excess can always be retrieved from the power margin. After all the users have been allocated, we apply the power adaptation procedure: we compute the total allocated power in the cell and if it does not exceed $P_T^{max}$, the preliminary power allocation is validated. Conversely, if it exceeds $P_T^{max}$, the power excess is obtained from the users with good channel conditions. We suggest carrying this out proportionally to the power margin they experience, i.e. more power is collected from users with high power margin. In Fig. 2, the final allocated power is depicted with a dotted line. Most of the excess power will be retrieved from user 2, as it has the highest power margin.

#### 2) *FFRopa Frequency Domain Packet Scheduling*

The FFRopa resource allocation algorithm comprises the following 5 steps and implements the ICIC scheme previously defined. After each allocation within a TTI, the TD scheduling priority is updated.

*Step 1*: Order users according to the TD scheduling priority, and then, for each user *k,* obtain the number of required RBs ($n_k$) to transmit its required $T_{req,k}$ using the target *MCS* ($MCS_k$).

*Step 2:* For each user *k* and following the scheduling priority order, allocate $n_k$ consecutive RBs belonging to the *SG* of the sector (applying the power adaptation) where it experiences the best channel state, assuring that the received SINR is greater than the $MCS_k$ threshold ($SINR_{MCSk}$) on these $n_k$ RBs. Two different situations may arise here:

- There is at least one possible allocation where user *k* meets all the constraints, so this allocation is preliminary accepted and we continue to step 2 updating the scheduling priority.
- If there is no possible allocation, we must start an assignment readjustment where users already allocated are moved to other RBs in the SG (always granting the SINR requirement compliance), so that the new user can receive an allocation. If the assignment readjustment succeeds, we continue with step 2. If it fails, we finally try to allocate this user on the outer subband. If this allocation is also impossible, we repeat the step 2 for this user but allocating one RB more ($n_k$+*1*), which results in a lower desired MCS and lower SINR requirements. In any case, if the allocation it is not possible with the minimum MCS of the system, we consider that the user can not be allocated in this subframe and continue to step 2 with the next user.

*Step 3*: When the *SG* RBs of a sector has been assigned, further allocations in this sector follows the occupation ordering as stated before. In this step we also consider the occupation of RBs as in Fig. 1 when a sector has allocated all the RBs of its preferred subband. If the user can not be allocated in the inner subband, even after an assignment readjustment, we try to allocate it in the outer subband. If none of these allocations is possible, we repeat step 3 with this user but assuming that it requires one RB more to be allocated.

*Step 4:* If the total allocated power exceeds the cell limit, we perform the power adaptation procedure to collect exceeding power from users with good channel conditions.

*Step 5:* Finally, increase the MCS for the already allocated users with remaining traffic if the attainable SINR is higher than its minimum SINR threshold in all their allocated RBs and there are enough data bits on user's buffers so as to fill the new data rate up. In [4] we proved that frequently using high MCS increases the BLER, especially in low-loaded scenarios, so we only improve the MCS for those users whose allocated data rate is not greater than their required data rate ($T_{req,k}$).

#### 3) *Channel-aware Deferring Strategy*

Users which suffer from temporal bad channel conditions require more resources to deliver their data packets. To avoid starving other users of resources, we propose a channel-aware deferring (CaD) strategy to block the allocation of low-efficient users, namely those users demanding the lowest MCS in the system (MCS0). Users with good or reasonable channel conditions are allocated on the RBs where they observe the best channel conditions, so they can use efficient MCSs, and the remaining RBs are finally allocated to deferred users.

Applying this CaD strategy may however impede the satisfaction of their QoS requirements, as their packets can be blocked close to their maximum allowable packet delay $d_{max,k}$. (packets are discarded when they trespass this delay constraint). To prevent this fact, a user *k* can only be deferred in the n-th TTI when its head-of-the-line (HoL) packet has a delay ($W_k$ *[n]*) lower than a certain bound related to its maximum allowable traffic delay ($d_{max,k}$):

$$\frac{W_k[\mathrm{n}]}{d_{\max,k}} \le \delta \quad 0 \le \delta \le 1 \qquad (2)$$

The FFRopa algorithm with CaD can be outlined as follows:

*Step 1*: Idem to FFRopa

*Step 2 and 3:* Proceed as in FFRopa but when a user requires the MCS0 to be allocated and satisfies (2), continue with the next user.

*Step 4:* Allocate the *remaining* free RBs to deferred users following the TD scheduling priority with the same procedure of steps 2 and 3 in FFRopa.

*Step 5 and 6:* This two steps flow respectively as step 4 and 5 in FFRopa, first performing the power adaptation and then improving the MCS of already allocated users.

### *B. Time Domain Scheduling Policies*

The performance achieved by the CaD strategy in FFRopa is analyzed with the TD scheduling strategies described below. Each user *k* is assumed to generate one single data flow. Each flow has associated a mean required data rate, $T_{req,k}$, and a maximum allowable packet delay, $d_{max,k}$. Packets exceeding the delay constraint $d_{max,k}$ are dropped, so the dropping rate is the key performance parameter, along with the transmission mean delay.

- *Modified Largest Delay First (M-LDF)*: the ratio of the packet delay ($W_k[n]$) in the n-th TTI to $d_{max,k}$:

$$k^* = \arg\max_k \frac{W_k[\mathrm{n}]}{d_{\max,k}} \tag{3}$$

- *Proportional Fair (PF)* [5]: the achievable data rate for user $k$ in the n-th TTI on allocable RBs ($R_k[n]$) over its averaged received data rate in the n-th TTI ($T_k[n]$):

$$k^* = \arg\max_k \frac{R_k[n]}{T_k[n]} \tag{4}$$

$T_k[n]$ is updated for scheduling at the n-th TTI as given below, where $t_c$ is the smoothing average window, and $r_k$ the instantaneous rate allocated to user $k$ in the (n-1)-th TTI:

$$T_k[n] = \left(1 - \frac{1}{t_c}\right) T_k[n-1] + \frac{1}{t_c} r_k[n-1] \tag{5}$$

- *Modified Largest Weighted Delay First (MLWDF)* [6]: the priority takes account of the packet delay ($W_k[n]$) and the channel conditions (as in PF). The term $a_k$ (6) is related to the probability $\delta_k$ (here 5%) of a packet exceeding a certain delay bound $D_k$, (here $0.8 \cdot d_{max,k}$):

$$k^* = \arg\max_k a_k W_k \frac{R_k[n]}{T_k[n]} \quad \text{with} \quad a_k = -\frac{\log(\delta_k)}{D_k} \tag{6}$$

## III. Simulation Results

In this section, we analyze the performance of the CaD strategy over the FFRopa algorithm. The simulations have been carried out with a C++ system simulator with the set of parameters collected in Table I (following the system level parameters and the FDD frame structure of LTE). We applied a wrap-around technique to avoid border effects.

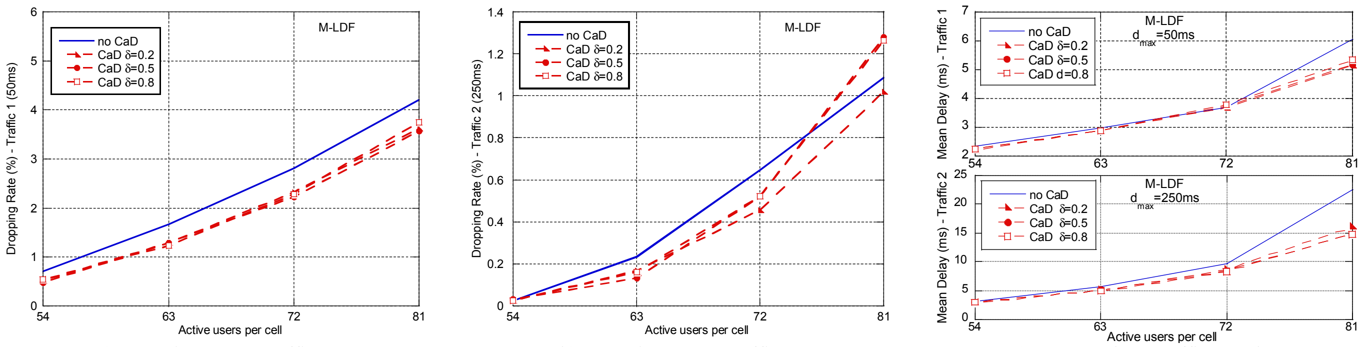
A scheduling period (TTI) of 1ms is assumed and an ARQ mechanism based on erroneous block retransmission is also considered (assuming a delay of $\Delta_{ARQ}$). As the outer subband covers one third of the cell coverage area, we assume that it comprises one third of the frequency resources (9 RBs). The set of MCSs is collected in Table I, along with their minimum SINR threshold (computed to achieve a BLER of $10^{-2}$) and their information data rate. Data services have been modeled as ON-OFF traffics with activity factor $\alpha$ (0.25). ON and OFF holding time has been modeled as an exponential random variable with mean $T_{ON}$ (30ms) and $T_{OFF}$ (90ms), respectively. In the ON state, a new packet of size $L=240$ bits is generated every TTI. This packet is divided into 2 transport units (TUs), which is the minimum data block size that can be transmitted. A maximum delay constraint of $d_{max}$ is assumed and a hard deadline is applied. Two traffic classes are considered, traffic 1 and 2, with $d_{max}$=50ms and $d_{max}$=250ms, respectively. The simulation scenario assumes that 50% of the users have traffic 1 and the other 50% with traffic 2, homogeneously distributed along the coverage area of each cell. The number of active users per cell ranges from 54 to 81. Users move at 3km/h within their respective sector to avoid handovers. They change their direction of movement within an angle of ±π/2 (chosen according to a uniform distribution) every 1s.

TABLE I. Simulation Parameters

| System Model | | | | | | |
|---|---|---|---|---|---|---|
| Cell deployment | 19 trisectorized cells | | | | | |
| Inter-site Distance | 1732m | | | | | |
| Carrier frequency | 2.5GHz | | | | | |
| Path-loss model | 128.1+37.6log$_{10}$(d(km)) | | | | | |
| Shadowing variance | 8dB | | | | | |
| Shadowing correlation distance | 50m | | | | | |
| Shadowing correlation between cells | 0.5 | | | | | |
| Noise Power | -174dBm/Hz | | | | | |
| System bandwidth / Number of RBs | 5MHz / 27 RBs | | | | | |
| RB size | 12 subcarriers for 1 TTI (1 ms) | | | | | |
| Channel model (3km/h) | Extended Pedestrian-A [8] | | | | | |
| Doppler model | Jakes | | | | | |
| Maximum output power | 43dBm | | | | | |
| Maximum transmit power per RB | 32dBm | | | | | |
| BS/UE antenna gain | 15dB(BS) / 0dB (UE) | | | | | |
| BS antenna horizontal pattern | 70deg (-3dB) with 20dB front-to-back ratio | | | | | |
| Transmission/Reception diversity gain | 3dB /3dB | | | | | |
| Receiver Noise Figure | 5dB | | | | | |
| Subframe duration (TTI) | 1ms | | | | | |
| ARQ delay ($\Delta_{ARQ}$) | 6ms | | | | | |
| Time between CQIs ($T_{CQI}$) | 1ms | | | | | |
| CQI averaging time ($W_{CQI}$) | 4ms | | | | | |
| CQI delay ($\Delta_{CQI}$) | 2ms | | | | | |
| **Power Masks (dBm/RB)** | | | | | | |
| Mask | $M_{med}$ | | $M_{max}$ | | $M_{min}$ | |
| Inner subband | 26dBm | | 29dBm | | 23dBm | |
| Outer subband | 31dBm | | 31dBm | | 31dBm | |
| **MCS Considered** | | | | | | |
| MCS index | 0 | 1 | 2 | 3 | 4 | 5 |
| Configuration | QPSK 1/2 | QPSK 3/4 | 16QAM 1/2 | 16QAM 3/4 | 64QAM 1/2 | 64QAM 3/4 |
| Info bits per RB | 120 | 180 | 240 | 360 | 480 | 540 |
| SINR threshold | 10.2dB | 14dB | 17.5dB | 21.4dB | 27.4dB | 29.5dB |

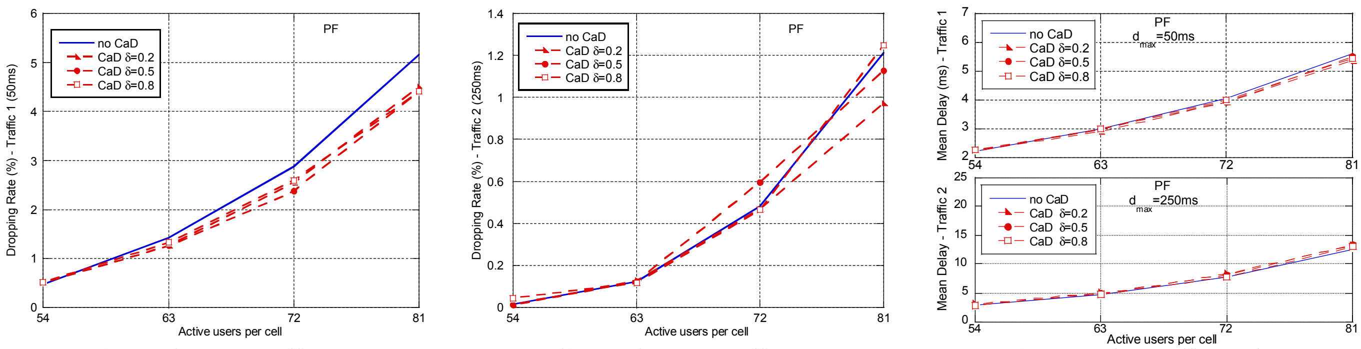
### A. Numerical results

Fig. 3-5 show respectively the performance in terms of dropping rate and mean delay for M-LDF, PF and MLWDF in FFRopa and when the CaD is applied with δ=0.2, 0.5 and 0.8. As shown in Fig. 4 for FFRopa without CaD, the dropping rate in PF rapidly rises as the cell load increases. Since it does not consider the QoS requirements to prioritize users, it cannot give more resources to the most delayed users and their dropping rate quickly rises. Consequently, as the PF term in MLWDF formula has more weight than the delay-dependent term, the dropping rate in MLWDF is higher than in M-LDF for both traffics. When it comes to the mean delay, the M-LDF scheduling policy (Fig. 3c) shows a quick increase for both traffics in high load conditions. As delayed users are more likely allocated with this policy, more packets with high delay are transmitted, which raises the mean delay in the cell. Since PF does not consider the delay constraint to prioritize users, the same allocation opportunities are given to both traffics. Users with traffic 2 benefit from this fact and their delay decreases especially for high load (Fig. 4c). In M-LDF and MLWDF, as the delay is taken into account in the TD scheduling, users with traffic 2 receive fewer allocation opportunities and their mean delay increases compared to the PF case (see Fig. 3c and 5c).



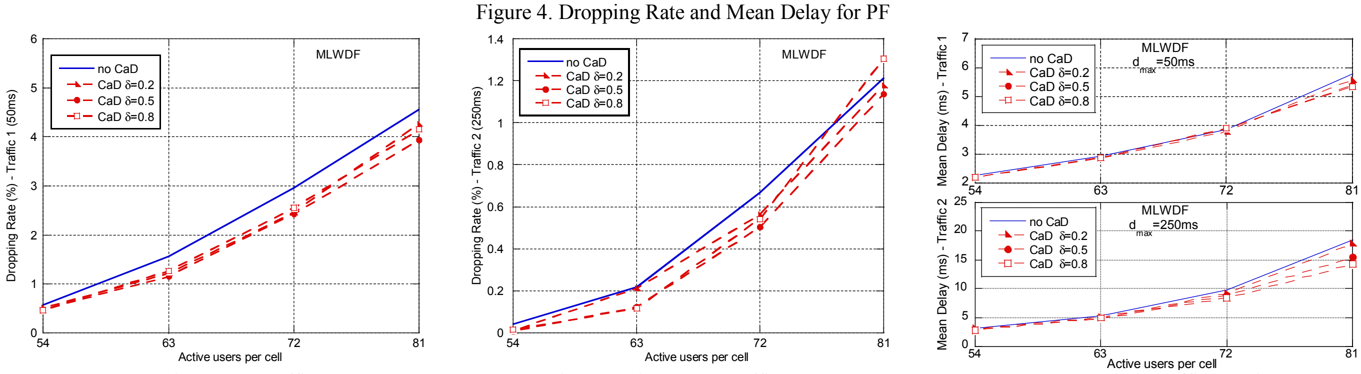

a) Dropping Rate. Traffic 1. b) Dropping Rate. Traffic 2. c) Mean Delay. Traffic 1 and 2.

Figure 3. Dropping Rate and Mean Delay for M-LDF




a) Dropping Rate. Traffic 1. b) Dropping Rate. Traffic 2. c) Mean Delay. Traffic 1 and 2.

Figure 4. Dropping Rate and Mean Delay for PF




a) Dropping Rate. Traffic 1. b) Dropping Rate. Traffic 2. c) Mean Delay. Traffic 1 and 2.

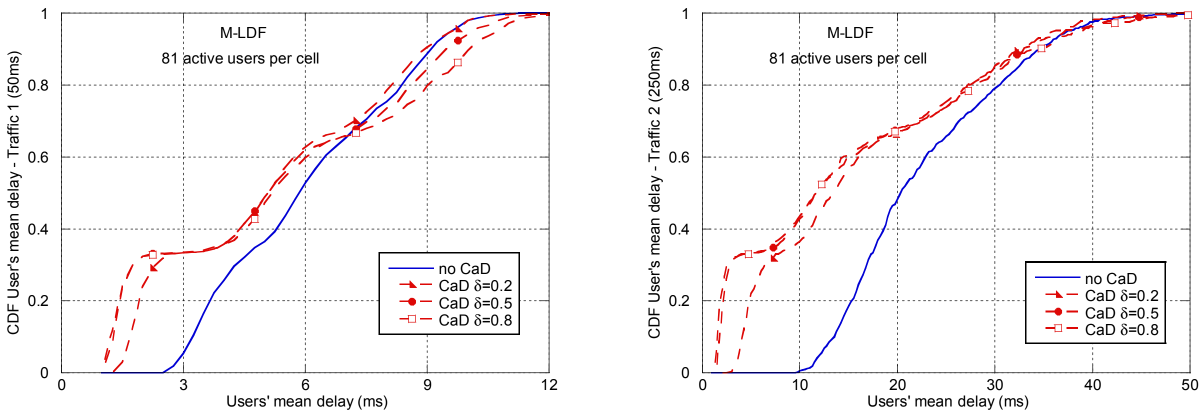
Figure 5. Dropping Rate and Mean Delay for MLWDF

When the CaD strategy is applied, a neat reduction in the dropping rate for the most restrictive traffic (traffic 1) can be achieved. As more resources are allocated to users with fairly good channel conditions (more frequently cell-center users), their dropping rate directly decreases at the expense of an increment in the delay of bad channel users. Thus, we can say that the CaD transfers some delay from users with good channel state to users with bad channel.

With the CaD, the users with good channel state are allocated first, so they can use more efficient MCSs and occupy the minimum number of RBs possible. Therefore, it is very likely that after the allocation of the users with good channel conditions, there are still free RBs to allocate some deferred users, so the impact of the CaD on such users is not critical. When comparing the possible values of δ, we realize that the lowest dropping rate for traffic 1 is obtained for δ=0.5. This value allows deferring bad users long enough to get a performance improvement for users with good channel conditions and prevents deferred users from accumulating high delays so as to impede the satisfaction of their maximum delay limits.

The value of δ has more influence on the dropping rate for traffic 2. This traffic has higher delay constraint (250ms), so these users can be deferred for longer periods when δ is high. Moreover, while a user remains deferred, it may generate more data packets that also accumulate high delays. In most cases, these users are not able to recover from this situation and their dropping rate increases, especially in high load conditions. That is why the lowest dropping rate in traffic 2 is generally obtained for δ=0.2.

When we analyze the impact of CaD on the mean delay, we realize that it avoids the steep growing in the delay for both M-LDF (Fig. 3c) and MLWDF (Fig 5c) as the cell load increases. Since more resources are allocated for users with




a) Traffic 1. b) Traffic 2.

Figure 6. CDF user's mean delay for M-LDF and 81 active users per cell

good channel, the use of efficient MCSs improves. This implies that more packets are transmitted per TTI and the mean delay decreases. This effect does not take place in PF since this scheduling policy already prioritizes the allocation of users with good channel. A value of δ=0.8 provides the best performance in terms of delay because a higher number of non-efficient users are deferred and more packets are delivered with low delay.

Finally, to test how the CaD strategy modifies the user's performance, we depict in Fig. 6 the cumulative distribution function (CDF) of the user's mean delay for M-LDF policy and 81 active users per cell. When the CaD is applied, almost 65% of the users with traffic 1 experience less mean delay whereas this percentage rises until 90% for traffic 2. These results confirm that the CaD transfers a great part of the delay experienced in the cell to a certain group of users whose delay was already high. As the QoS is continuously tracked, the impact of the CaD on these last users is significantly lower than the benefit provided to the users with good channel conditions.

To summarize, the CaD improves the global performance in term of dropping rate and delay at the expense of a slight degradation in the mean delay observed by a small group of users with bad channel state. A value of δ=0.5 shows a good trade-off for both traffics as it provides the lowest dropping rate for the most restrictive traffic while the performance degradation for traffic 2 is not very important. This value also avoids the fast rising in the mean delay for M-LDF and MLWDF as the cell load increases.

## IV. Conclusion

In this paper we study the performance improvement that can be achieved in a mobile OFDMA network with a CaD strategy that defers the allocation of users with bad channel conditions in order to save more resources to cell-center users. This CaD, along with a delay-dependent criterion to decide whether a user can be deferred or not, allows reducing the dropping rate and the mean delay in the cell, even for the most restrictive traffic in the scenario.

## Acknowledgment

This work has been supported by the Spanish Government through the grant TEC2008-04946 from the Ministerio de Ciencia e Innovación (MICINN) and Fondos Europeos de Desarrollo Regional (FEDER), and by the project CENIT CEN2007-1036: TelMAX from Ministerio de Industria (MITYC).

## References

[1] Y. Xiang, J. Luo, C. Hartmann, "Inter-cell Interference Mitigation through Flexible Resource Reuse in OFDMA based Communications Networks", Proc. European Wireless, EW'07, pp. 1-7, April 2007.
[2] H. Lei, X. Zhang, D. Yang, "A novel frequency reuse scheme for multi-cell OFDMA systems", Proc. IEEE Vehicular Technology Conference (VTC'07-Fall), Baltimore, MD, 2007, pp. 347–351.
[3] 3GPP, TSG RAN WG1#42 R1-050507, "Soft frequency reuse scheme for UTRAN LTE", Huawei, May 2005.
[4] A. Hernández, I. Guío, A. Valdovinos, "Downlink Scheduling for Intercell Interference Fluctuation Mitigation in Partial-Loaded Broadband Cellular OFDMA Systems". Proc. IEEE International Conf. on Ultra Modern Telecommunications 2009 (ICUMT'09), pp.1-6.
[5] A. Pokhariyal, et al, "HARQ Aware Frequency Domain Packet Scheduler with Different Degrees of Fairness for the UTRAN Long Term Evolution". Proc. VTC2007-Spring, pp. 2761-2765, April 2007
[6] M. Andrews, et al., "Providing quality of service over a shared wireless link," IEEE Commun. Mag., vol. 39, no. 2, pp. 150-154, Feb. 2001.
[7] 3GPP TR 36.942: "Radio Frequency system scenarios" (Release 8).
[8] 3GPP TS 36.104. "Base station (BS) radio transmission and reception (FDD)".