Fairness-Based Dynamic Sub-carrier Assignment for Real Time Services in OFDMA Networks

Emanuel B. Rodrigues
Wireless Telecommunications Research Group (GTEL)
Federal University of Ceará (UFC)
Fortaleza, Brazil
Email: emanuel@gtel.ufc.br

Fernando Casadevall
Department of Signal Theory and Communications (TSC)
Universitat Politècnica de Catalunya (UPC)
Barcelona, Spain
Email: ferranc@tsc.upc.edu

***Abstract*—Fairness is a crucial aspect for resource allocation in wireless systems, particularly in the downlink of 4G Orthogonal Frequency Division Multiple Access (OFDMA)-based systems. Therefore, it would be interesting for network operators to have tools to manage fairness in their systems. This work describes the utility-based beta-rule, which is a flexible resource allocation framework based on utility theory that is able to control delay-based fairness in a scenario with real time services in an OFDMA-based system. Not only can this framework be designed to work as well-known classical policies found in the literature, but also as an adaptive policy, which is able to meet a desired system fairness target. System-level simulations demonstrate that the adaptive policy is able to guarantee a fairness requirement by dynamically adapting a fairness-controlling parameter in the sub-carrier assignment algorithm. It also presents better results than the classical policies in terms of cell capacity and user satisfaction.**

I. Introduction

Efficiency in the usage of radio resources, fairness and user satisfaction are crucial aspects for resource allocation in wireless systems, particularly in the downlink of 4G Orthogonal Frequency Division Multiple Access (OFDMA)-based systems. These three aspects are interrelated intrinsically, which generates fundamental trade-offs between them. Therefore, a better understanding of what fairness means and how it can be controlled by means of Radio Resource Allocation (RRA) techniques is of utmost importance.

Fairness is related to the notion of how equal is the number of resources allocated (resource-based) or how similar is the service quality experienced by the players (QoS-based). More insightful information about fairness can be found in [1].

Regarding Quality of Service (QoS)-based fairness, it is well known that the inherent characteristics and transmission requirements of Real Time (RT) traffic differ from those of Non-Real Time (NRT) data traffics. RT services, such as Voice over IP (VoIP), require a low and bounded delay, while NRT services, such as World Wide Web (WWW), are not delay-sensitive but require an overall high throughput. Due to these factors, the packet delay can be used as a fairness indicator in a scenario with RT services, while rate or throughput can be used as fairness indicators in a scenario with NRT services.

Utility Theory is a powerful tool that can be used to design RRA algorithms able to achieve different levels of fairness in the resource allocation process [2]. This theory can be used in communication networks to quantify the benefit of usage of certain resources, e.g. bandwidth, power; or evaluate the degree to which a network satisfies service requirement of users' applications, for example in terms of throughput and delay. In particular, this work considers the flexibility of utility functions to propose an adaptive RRA framework able to provide long-term fairness control for RT services.

II. Related Work

Some works have dealt with fairness issues for RT services implicitly. These works can be classified in two main approaches: Packet Scheduling (PSC) and utility theory.

The opportunistic PSC algorithms suitable for RT services found in the literature have priority functions that always use an efficiency indicator, such as the instantaneous transmission rate (rate maximization policy) or the ratio between the instantaneous transmission rate and throughput (proportional fairness policy), and a QoS indicator based on delay. In this way, these PSC algorithms implicitly add to the problem some kind of delay-based fairness, because the flows that presented higher delays would have more priority to use the resources. Some examplary works are [3]–[6].

The utility-based PSC algorithms adopted a more general procedure. In the case of RT services, the QoS indicator used in the priority functions is a marginal utility function based on delay. For example, [7] and [8] used z-shaped utility functions while [9] used particularly designed utility functions. Notice that the utility-based approach is more general than classical PSC priority functions because the utility functions can be freely designed to provide the desired fairness control.

As far as we are concerned, reference [10] was the first work to use the utility fairness concept to propose a general RRA policy suitable for RT services. The authors in [10] proposed a parametric utility-based RRA framework called utility-based beta-rule. This framework can be designed to work as some well-known classical RRA policies by adjusting only the fairness controlling parameter β in its parametric structure. This flexible RRA framework can be dynamically configured depending on the network conditions and the network operator's objectives. It was initially meant to manage the trade-off between efficiency in the resource usage and delay-based fairness in a scenario with RT services.

ISBN 978-3-8007-3621-8

The present work aims to deepen the mathematical formulation and the performance evaluation of the utility-based beta-rule framework proposed in [10]. We focus on the detailed description of the Adaptive Delay-Based Fairness (ADF) policy, which is able to perform a long-term fairness control by dynamically adapting the utility function used in the framework in order to keep the system fairness around a planned value.

III. Utility-Based Resource Allocation for Real Time Services in OFDMA Networks

The utility optimization problem considered in this work has the following objective function:

$$\max_{\mathcal{S}_j,\mathbf{p}} \sum_{j=1}^{J} U\left(d_j^{\mathrm{hol}}[n]\right). \quad (1)$$

The optimization constraints are: 1) the subsets of sub-carriers assigned to different users must be disjoint; 2) the union of all these subsets must be contained in the total set of sub-carriers available in the system; 3) the powers allocated to the sub-carriers are positive; and 4) the total sum of the powers over all sub-carriers must not surpass the total Base Station (BS) transmit power. We have that J is the total number of Mobile Terminals (MTs) in the macrocell, $\mathcal{S}_j$ is the subset of sub-carriers assigned to the jth MT, $\mathbf{p}$ is the vector of powers for all sub-carriers, and $U\left(d_j^{\mathrm{hol}}[n]\right)$ is a concave and decreasing utility function based on the current Head-Of-Line (HOL) delay $d_j^{\mathrm{hol}}[n]$ of the jth MT.

The HOL delay is the time the oldest packet in the user buffer has to wait to gain access to the wireless channel. Considering a generic MT j, it can be calculated approximately by the following recursive equation:

$$d_j^{\mathrm{hol}}[n+1] = d_j^{\mathrm{hol}}[n] + \frac{b_j^{\mathrm{hol}}[n] - R_j[n] \cdot t_{\mathrm{tti}}}{T_j[n-1]} \quad (2)$$

where $b_j^{\mathrm{hol}}[n]$ is the current number of bits in the HOL packet, t_{tti} is the duration of the Transmission Time Interval (TTI) in seconds, $T_j[n-1]$ is the average data rate (throughput) up to the previous transmission interval and $R_j[n]$ is the instantaneous achievable transmission rate. If the jth MT has not been served by any sub-carrier in the nth TTI, $R_j[n]$ is equal to zero and the HOL delay is incremented. This delay increment is calculated assuming that the remaining bits of the HOL packet will be transmitted using a rate equal to the throughput experienced so far by the MT. For sake of simplicity, we assume that the packets' interarrival time is equal to the TTI duration. Although the packet interarrival time depends on the type of application, this assumption does not invalidate the mathematical and conceptual RRA framework, and make the optimization model much more tractable. Taking this into account, one can see in (2) that if the instantaneous transmission rate is such that all remaining bits of the HOL packet are transmitted in the current TTI, the HOL delay remains constant because the previous packet in the buffer will be the HOL packet now. Finally, the HOL delay is decremented when the instantaneous achievable transmission rate is high enough to transmit the remaining bits of the HOL packet and some bits of the preceding packets in the queue.

Depending of the utility function used in (1), the optimum solution for the joint optimization problem is very difficult to be found. Most of the sub-optimum solutions are based on the problem-splitting technique, which splits the problem in two stages: first, sub-carrier assignment with fixed power allocation, and next, power allocation with fixed sub-carrier assignment. In the present work, we also use this technique.

Firstly, we will evaluate the optimization problem more carefully and make some simplifications. Assessing the objective function in (1) and the HOL delay expression in (2), we can see that the derivative of $U\left(d_j^{\mathrm{hol}}\right)$ with respect to the transmission rate R_j can be expressed as:

$$\frac{\partial U}{\partial R_j} = \frac{\partial U}{\partial d_j^{\mathrm{hol}}} \cdot \frac{\partial d_j^{\mathrm{hol}}}{\partial R_j} = \frac{\partial U}{\partial d_j^{\mathrm{hol}}} \cdot \left(-\frac{t_{\mathrm{tti}}}{T_j[n-1]}\right).$$

Using the result above and assuming that the TTI duration is sufficiently small, the Lagrange theorem of the mean can be used, which says that [8], [11]:

$$\begin{aligned}
&\sum_{j=1}^{J} U\left(d_j^{\mathrm{hol}}[n+1]\right) - \sum_{j=1}^{J} U\left(d_j^{\mathrm{hol}}[n]\right) \\
&\approx \sum_{j=1}^{J} \left.\frac{\partial U}{\partial R_j}\right|_{R_j = R_j[n-1]} \cdot \left(R_j[n] - R_j[n-1]\right) \\
&= \sum_{j=1}^{J} -\left.\frac{\partial U}{\partial d_j^{\mathrm{hol}}}\right|_{d_j^{\mathrm{hol}} = d_j^{\mathrm{hol}}[n]} \cdot \frac{t_{\mathrm{tti}} \cdot \left(R_j[n] - R_j[n-1]\right)}{T_j[n-1]} \\
&= \sum_{j=1}^{J} \left|\frac{\partial U}{\partial d_j^{\mathrm{hol}}}\right|\Bigg|_{d_j^{\mathrm{hol}} = d_j^{\mathrm{hol}}[n]} \cdot \frac{t_{\mathrm{tti}} \cdot \left(R_j[n] - R_j[n-1]\right)}{T_j[n-1]} \quad (3)
\end{aligned}$$

The absolute value operator was used in (3) because the utility function was assumed to be concave and decreasing, which yields negative marginal utilities and cancels the negative sign in (3). Notice that the maximization of (3) leads to the maximization of (1). In (3) we have that t_{tti} is a constant and $R_j[n-1]$ is known and fixed at the nth TTI. So a simplified optimization objective function can be expressed as:

$$\max_{\mathcal{S}_j,\mathbf{p}} \sum_{j=1}^{J} \frac{\left|U'\left(d_j^{\mathrm{hol}}[n]\right)\right|}{T_j[n-1]} \cdot R_j[n] \quad (4)$$

where $U'\left(d_j^{\mathrm{hol}}[n]\right) = \left.\frac{\partial U\left(d_j^{\mathrm{hol}}\right)}{\partial d_j^{\mathrm{hol}}}\right|_{d_j^{\mathrm{hol}} = d_j^{\mathrm{hol}}[n]}$ is the marginal utility of the jth MT with respect to its current HOL delay, and $d_j^{\mathrm{hol}}[n]$ can be obtained from the recursive equation (2). The problem (4) is a weighted sum rate maximization [12], where the weights are given by

$$w_j^{\mathrm{rt}} = \frac{\left|U'\left(d_j^{\mathrm{hol}}[n]\right)\right|}{T_j[n-1]}. \quad (5)$$

One can also notice that the optimization objective function in (4) is a linear function of $R_j[n]$. According to [8], [11],

ISBN 978-3-8007-3621-8

we can rely on this fact to state that the Dynamic Sub-carrier Assignment (DSA) problem with equal power allocation for RT services employs the following reasoning: the MT with index $m(k,n)$ is chosen to transmit on the kth sub-carrier at the nth TTI if it satisfies the condition given by (6) below:

$$m(k,n) = \arg\max_{j} \left\{ w_j^{\mathrm{rt}} \cdot c_{j,k}\left[n\right] \right\}, \tag{6}$$

where the weight w_j^{rt} is given by (5) and $c_{j,k}[n]$ denotes the instantaneous achievable transmission efficiency of the kth sub-carrier with respect to the jth MT assuming equal power allocation per sub-carrier. Notice that the users that have higher weights w_j^{rt} will have priority in the sub-carrier assignment procedure. Therefore, this user prioritization can be used to adapt the fairness in the system.

Finally, the power allocation among the sub-carriers must be done. In this work we consider Equal Power Allocation (EPA) instead of Adaptive Power Allocation (APA), since it has been shown in [10] that the utility-based beta-rule with EPA has almost the same performance than APA with less computational complexity.

IV. Adaptive Framework for Fairness Control

A. Utility-Based Beta-Rule

We consider a novel family of utility functions based on the HOL delay of the form presented below:

$$U\left(d_j^{\mathrm{hol}}\left[n\right]\right) = -\frac{\left(d_j^{\mathrm{hol}}\left[n\right]\right)^{1+\beta}}{1+\beta} \tag{7}$$

where $\beta \in [0,\infty)$ is a non-negative parameter that determines the degree of delay-based fairness.

It is clear that the longer the HOL delay a user experiences, the lower level of satisfaction the user has. Thus, we can assume that $U\left(d_j^{\mathrm{hol}}\right)$ is a decreasing and strictly concave function. This implies that the marginal utility, which is the derivative $\partial U/\partial d_j^{\mathrm{hol}}$ is a negative and decreasing function. However, the absolute value of the marginal utility is used in our proposed RRA policy, as a component of the weight w_j^{rt} given by (5). Looking at (5), one can clearly see that the higher the weight w_j^{rt}, i.e. the higher the HOL delay experienced by a given MT, the higher will be the priority of this MT to get a sub-carrier. And such priority is higher when β increases. Therefore, one can conclude that when β increases, the MTs with poorest QoS (higher HOL delay) are benefited, and so the fairness in the system becomes stricter.

The expression of the weight w_j^{rt} is particularized for the utility function (7) as follows:

$$w_j^{\mathrm{rt}} = \frac{\left(d_j^{\mathrm{hol}}\left[n\right]\right)^{\beta}}{T_j\left[n-1\right]}. \tag{8}$$

This particular weight must be used in the DSA algorithm, which is given by (6).

Different performances in terms of resource efficiency and delay-based fairness can be achieved depending on the value of the fairness controlling parameter β. Varying β, the beta-rule framework presented above can be designed to work as different classical RRA policies suitable for RT services, such as Proportional Fairness (PF) [13], Modified Largest Weighted Delay First (M-LWDF) [3] and First-In-First-Out (FIFO) [4]. Moreover, this work also proposes a novel adaptive policy called Adaptive Delay-Based Fairness (ADF), which is described in section IV-B. Table I summarizes the main characteristics of the beta-rule framework and the four particular RRA policies contemplated by this framework.

TABLE I
FEATURES OF THE UTILITY-BASED BETA-RULE FRAMEWORK

Policies	β	w_j^{rt}	**Characteristics**
PF	0	$\frac{1}{T_j[n-1]}$	High efficiency and low fairness
M-LWDF	1	$\frac{d_j^{\mathrm{hol}}[n]}{T_j[n-1]}$	Static trade-off b/w efficiency and fairness
FIFO	$\beta \to \infty$	$\lim_{\beta\to\infty} \frac{d_j^{\mathrm{hol}}[n]^{\beta}}{T_j[n-1]}$	Low efficiency and high fairness
ADF	adaptive	$\frac{d_j^{\mathrm{hol}}[n]^{\beta}}{T_j[n-1]}$	Dynamic trade-off b/w efficiency and fairness

B. Adaptive Delay-Based Fairness (ADF)

It was shown that a general RRA technique based on (7) is able to provide several degrees of delay-based fairness. The ADF policy explores this flexibility in order to perform a delay-based fairness control planned by the network operator in a scenario with RT services. This is done by means of the adaptation of the fairness controlling parameter β within the utility-based beta-rule framework shown in Table I.

Notice that it is not necessary to use the whole range of β values in order to achieve suitable fairness degrees. In fact, the delay-based fairness control is very sensitive to the value of β. Thus, we have that small values are sufficient to provide desired fairness degrees on the ADF DSA algorithm.

The ADF policy is based on the definition of the User Fairness Index (UFI) ϕ_j^{rt}, which depends on the filtered HOL delay and is calculated for each MT in the cell. The UFI changes with time and is defined as:

$$\phi_j^{\mathrm{rt}}\left[n\right] = \frac{d_j^{\mathrm{req}}}{d_j^{\mathrm{hol,filt}}\left[n\right]} \tag{9}$$

Normally, the delay requirement of the jth MT d_j^{req} is the same for all users of the same type and is equal to the delay budget of the RT service (maximum time that a packet can spend in the buffer before being discarded). Note that in (9), a filtered version of the HOL delay using a low-pass exponential filtering was considered. This was done in order to smooth the time series and allow a more stable control of the fairness.

Next, a fairness index for the whole cell comprising all RT flows is defined by

$$\Phi_{\mathrm{cell}}^{\mathrm{rt}}\left[n\right] = \frac{\left(\sum_{j=1}^{J} \phi_j^{\mathrm{rt}}\left[n\right]\right)^2}{J \cdot \sum_{j=1}^{J} \left(\phi_j^{\mathrm{rt}}\left[n\right]\right)^2}, \tag{10}$$

ISBN 978-3-8007-3621-8

where J is the number of MTs in the cell. This proposed Cell Fairness Index (CFI) is a particularization of the well-known Jain's fairness index proposed by Jain et al. in [14].

Notice that $1/J \leq \Phi_{\text{cell}}^{\text{rt}}[n] \leq 1$, where J is the total number of MTs in the cell. A perfect fair allocation is achieved when $\Phi_{\text{cell}}^{\text{rt}}[n] = 1$, which means that the HOL packet delays of all MTs are equally proportional to their delay requirements (all user fairness indexes are equal). The worst allocation occurs when $\Phi_{\text{cell}}^{\text{rt}}[n] = 1/J$, which means that all sub-carriers were allocated to only one MT, i.e. the HOL packet delay of one user is very low while the others are very high. It is relevant to emphasize that the fairness calculation procedure presented above is general in the sense that different classes of RT users with different delay requirements can be contemplated.

The objective of the ADF policy is guarantee that the instantaneous CFI $\Phi_{\text{cell}}^{\text{rt}}[n]$ converges to a planned value $\Phi_{\text{target}}^{\text{rt}}$ by adapting the parameter β in the utility-based beta-rule framework. The parameter β is adapted using a feedback control loop, as indicated below.

$$\beta[n] = \beta[n-1] - \eta_{\text{rt}} \cdot \left(\Phi_{\text{filt}}^{\text{rt}}[n] - \Phi_{\text{target}}^{\text{rt}}\right) \tag{11}$$

where $\Phi_{\text{filt}}^{\text{rt}}[n]$ is a filtered version of the CFI $\Phi_{\text{cell}}^{\text{rt}}[n]$ using an exponential smoothing filtering; $\Phi_{\text{target}}^{\text{rt}}$ is the desired value for the CFI (Cell Fairness Target (CFT)); and the parameter η_{rt} is a step size that controls the adaptation speed of β.

The ADF technique is an iterative and sequential process. At each TTI, the steps indicated in Fig. 1 are executed. After some iterations (TTIs), the ADF technique reaches a stable convergence of the fairness pattern defined by the target CFI. The simplicity of the ADF policy makes it a robust and reliable way to control the trade-off between resource efficiency and delay-based fairness among RT flows. By keeping the cell fairness around a planned target value, the network operator can have a stricter control of the network QoS and also have a good prediction about the performance in terms of system capacity.

In order to illustrate the convergence of the ADF technique, we present Fig. 2. It shows an example of the convergence of the filtered CFI $\Phi_{\text{filt}}^{\text{rt}}$ when considering a CFT equal to 0.8 (Fig. 2(a)). In this figure, two different step sizes η_{rt}, which are used in the control loop of the parameter β (see expression (11)), are assessed. It can be noticed that a higher value of η_{rt} allows a faster convergence in comparison with a lower value, but the steady-state error is also higher.

Fig. 2(b) shows how the parameter β responds to the variation of the filtered CFI $\Phi_{\text{filt}}^{\text{rt}}$. When $\Phi_{\text{filt}}^{\text{rt}}$ is higher than 0.8, which is the chosen CFT, the ADF technique decreases β in order to decrease the fairness in the system to a value closer to the desired target. Otherwise, when the filtered cell fairness index is not sufficiently high, β must be increased by the control loop, which forces the resource allocation to be fairer. This control is done indefinitely, which allows the convergence of $\Phi_{\text{filt}}^{\text{rt}}$, and consequently a strict control of the delay-based fairness in the system.

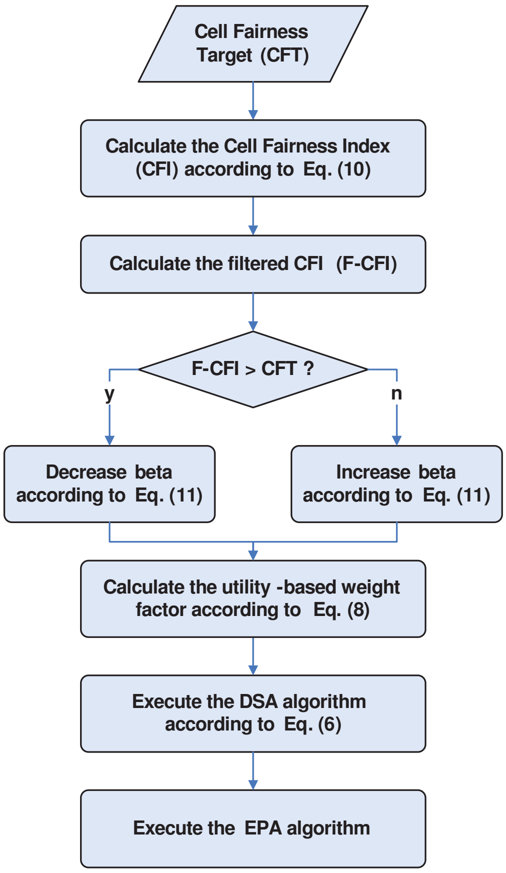


Fig. 1. Block diagram of the Adaptive Delay-Based Fairness (ADF) technique

V. Simulation Results

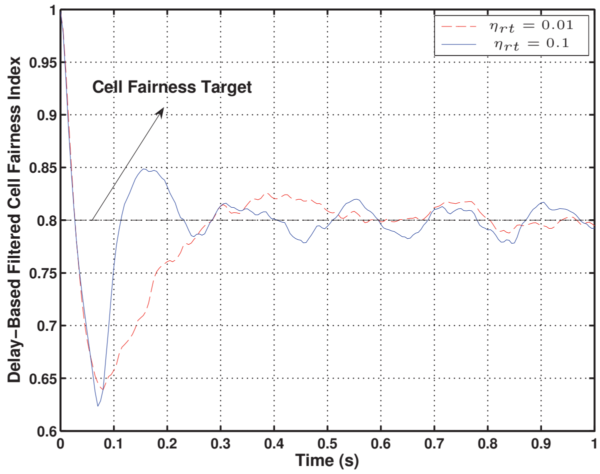
The performance of the utility-based beta-rule framework is evaluated by means of system-level simulations. The simulations took into account the main characteristics of an OFDMA system. The simulation parameters are depicted in Table II.

A. Fairness Analysis

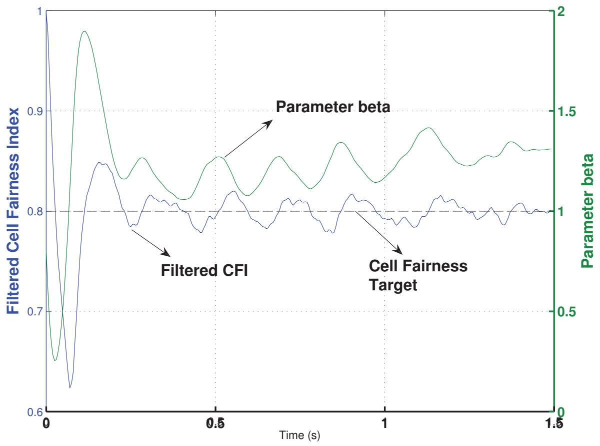
In Fig. 3 the fairness results are shown. We can see that the two extremes are clear: FIFO provides the highest fairness, while PF is the worst. In particular, PF shows a monotonically behavior, where the fairness index decreases with the increase of the number of users. When there are more users in the system, PF is able to better explore the multi-user diversity and a more opportunistic allocation causes a fairness decrease. FIFO presents the highest fairness, which means that it is the best policy at equalizing the values of the HOL delays of the users. M-LWDF and ADF have an intermediate behavior between FIFO and PF, in accordance with our expectations.

Notice that M-LWDF presents a fairness curve that slightly increases with the system load. This is due to the fact that the system is becoming congested for the range of loads considered in the simulations. Overload situations in a scenario with RT services is characterized by high delay-based fairness and poor QoS (compromise between fairness and QoS). This is the case of the FIFO policy that provides the maximum

ISBN 978-3-8007-3621-8



(a) Example of convergence of the delay-based filtered cell fairness index for various step sizes η_{rt}



(b) Example of variation of the delay-based filtered cell fairness index and adaptation of the parameter β

Fig. 2. Convergence analysis of the ADF Policy

fairness at the expense of unacceptable delays. Since PF uses the resources more efficiently, it is able to avoid congestion for the range of loads considered. We have that M-LWDF is a static trade-off between FIFO and PF, so it presents worse QoS and higher delay-based fairness as the system becomes more congested. On the other hand, the ADF policy is very precise at controlling the fairness levels and efficient at preventing the system from being congested, as can be observed in Fig. 3. Due to the structure of the utility-based beta-rule framework and the limited range for the adaptation of the parameter β, the performance of the ADF policy is constrained by the performances of FIFO (maximum β) and PF (minimum β).

B. Efficiency Analysis

In Fig. 4, the total cell throughput for the RRA policies is given. As expected, the worst performance is presented by FIFO, because it only takes the delay information into account, which may lead to an inefficient resource allocation. The other policies, which give more importance to the channel quality

TABLE II
SIMULATION PARAMETERS

Parameter	Value
Number of cells	1
Maximum BS transmission power	1 W
Cell radius	500 m
MT speed	static
Carrier frequency	2 GHz
Number of sub-carriers	192
Effective sub-carrier bandwidth	14 kHz
Path loss[a]	$L = 128.1 + 37.6 \log_{10} d$
Log-normal shadowing standard dev.	8 dB
Small-scale fading	3GPP Typical Urban (TU)
AWGN power per sub-carrier	-123.24 dBm
BER requirement	10^{-6}
Link adaptation	Continuous using effective Shannon capacity formula [11]
Transmission Time Interval (TTI)	0.5 ms
RT traffic model	Packets of 32 bytes with interarrival time of 2 ms
HOL delay filtering constant (f_{delay})	100
β range	[0,10]
ADF control time window	0.5 ms
ADF fairness target ($\Phi^{\text{rt}}_{\text{target}}$)	Variable
ADF step size (η_{rt})	0.1
ADF filtering time constant	10
FER threshold	2%
RT delay budget	100 ms
Simulation time span	5 s
Number of independent runs	70

[a] d is the distance to the BS in km.

and use the resources more efficiently, show similar behavior with higher system capacity. Notice that the ADF policy configured with a CFT equal to 1.0 presents a cell capacity much higher than FIFO, even though both are equivalent in terms of delay-based fairness (see Fig. 3).

C. Satisfaction Analysis

A RT user is considered satisfied if its Frame Erasure Rate (FER) is lower than a threshold. In our simulation model, we assume that a frame is lost if a packet arrives at the MT receiver later than the delay budget of the RT service. The

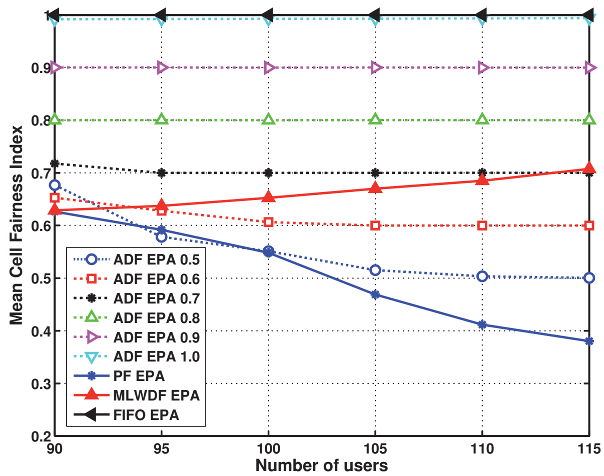


Fig. 3. Mean cell fairness index as a function of the number of users

ISBN 978-3-8007-3621-8

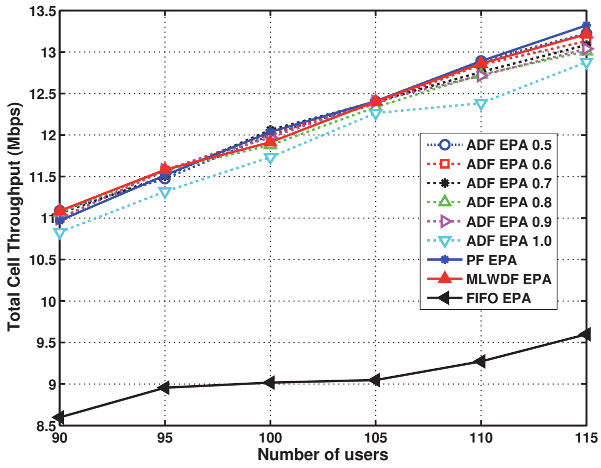


Fig. 4. Total cell throughput as a function of the number of users

values of the FER threshold and the delay budget considered in the simulations are presented in Table II.

The satisfaction of the RT users, which is based on the percentage of packets discarded due to excessive delay, is presented in Fig. 5. The QoS degradation provoked by the FIFO policy is clear. Although FIFO takes into account the delay in its allocation criterion, it is the one that presents the highest packet delays. This shows that the fact of not exploiting the OFDMA diversities is not beneficial in terms of QoS. Furthermore, when the system load increases, it causes the system to become stuck, i.e. the majority of the packets are discarded because they have a delay greater than 100 ms. All other RRA policies present similar performance regarding user satisfaction. Although the ADF policy with CFT equal to 1.0 shows a performance similar to FIFO in terms of delay-based fairness, it provides a much better satisfaction.

VI. Conclusions

The utility-based beta-rule framework described in this work is able to provide several degrees of fairness based on HOL

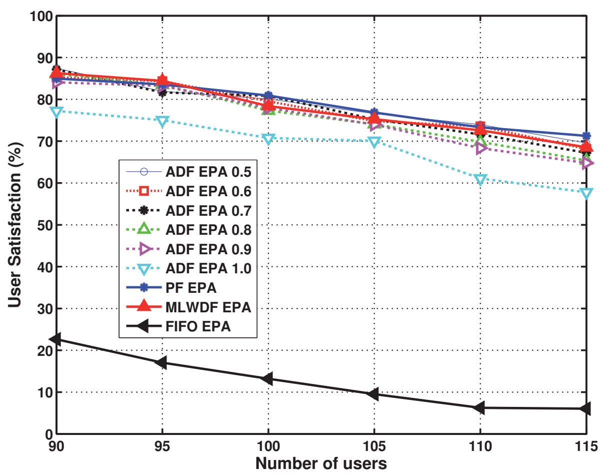


Fig. 5. User satisfaction as a function of the number of users

delay. The adaptive ADF policy controls the fairness in the system by setting a fairness index target and letting the utility-based adaptive policy control the fairness dynamically.

The ADF policy is able to converge the cell fairness index to any target fairness index defined by the network operator. This fairness control is bounded by the structure of the beta-rule framework, i.e. the minimum and maximum fairness performance depends on the allowed range of values for the parameter β. The ADF policy is robust in terms of cell capacity and user satisfaction for all CFTs considered in our study.

Acknowledgment

Emanuel B. Rodrigues would like to thank CAPES (Brazil) for a Ph.D. Scholarship support, and CNPq (Brazil) and FUNCAP (Ceará, Brazil) for a Postdoctoral Scholarship support.

References

[1] T. Lan, D. Kao, M. Chiang, and A. Sabharwal, “An axiomatic theory of fairness in network resource allocation,” in *Proc. IEEE 29th Conference on Computer Communications - INFOCOM*, Mar. 2010, pp. 1–9.

[2] E. B. Rodrigues, “Adaptive Radio Resource Management for OFDMA-Based Macro- and Femtocell Networks,” Ph.D. dissertation, Universitat Politècnica de Catalunya - UPC, Barcelona, Spain, 2011.

[3] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, “Providing quality of service over a shared wireless link,” *IEEE Commun. Mag.*, vol. 32, no. 2, pp. 150–154, Jun. 2001.

[4] S. Shakkottai and A. L. Stolyar, “Scheduling algorithms for a mixture of real-time and non-real-time data in HDR,” in *Proc. of 17th International Teletraffic Congress - ITC*, 2001, pp. 793–804.

[5] A. R. Braga, E. B. Rodrigues, and F. R. P. Cavalcanti, “Packet scheduling for VoIP over HSDPA in mixed traffic scenarios,” in *Proc. IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications - PIMRC*, Sep. 2006, pp. 1–5.

[6] R. Srinivasan and J. S. Baras, “An analysis of delay-constrained opportunistic scheduling for cellular wireless systems,” Center for Satellite and Hybrid Commun. Networks - CSHCN, Tech. Rep. TR 2004-2, 2004.

[7] S. Ryu, B. Ryu, H. Seo, and M. Shin, “Urgency and efficiency based wireless downlink packet scheduling algorithm in OFDMA system,” in *Proc. IEEE 61st Vehicular Technology Conference - VTC Spring*, vol. 3, May 2005, pp. 1456–1462.

[8] H. Lei, L. Zhang, X. Zhang, and D. Yang, “A packet scheduling algorithm using utility function for mixed services in the downlink of OFDMA systems,” in *Proc. IEEE 66th Vehicular Technology Conference - VTC Fall*, Sep. 2007, pp. 1664–1668.

[9] G. Song and Y. G. Li, “Utility-based resource allocation and scheduling in OFDM-based wireless broadband networks,” *IEEE Communications Magazine*, vol. 43, no. 12, pp. 127–134, Dec. 2005.

[10] E. B. Rodrigues and F. Casadevall, “Control of the trade-off between resource efficiency and user fairness in wireless networks using utility-based adaptive resource allocation,” *IEEE Communications Magazine*, vol. 49, no. 9, pp. 90–98, 2011.

[11] G. Song, “Cross-layer resource allocation and scheduling in wireless multicarrier networks,” Ph.D. dissertation, Georgia Institute of Technology, Georgia, USA, 2005.

[12] L. M. C. Hoo, B. Halder, J. Tellado, and J. M. Cioffi, “Multiuser transmit optimisation for multicarrier broadcast channels: asymptotic FDMA capacity region and algorithms,” *IEEE Transactions on Communications*, vol. 52, no. 6, pp. 922–930, Jun. 2004.

[13] F. Kelly, “Charging and rate control for elastic traffic,” *European Transactions on Communications*, vol. 8, pp. 33–37, 1997.

[14] R. Jain, D. Chiu, and W. Hawe, “A quantitative measure of fairness and discrimination for resource allocation in shared computer systems,” DEC Research, Tech. Rep. TR-301, Sep. 1984.

ISBN 978-3-8007-3621-8