

### **Copyright Notice**

© 2009 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.

# Adaptive Fairness Control for a Proportional Fair LTE Scheduler

Magnus Proebster, Christian M. Mueller

Universität Stuttgart, Institute of Communication Networks

and Computer Engineering, Stuttgart, Germany

magnus.proebster@ikr.uni-stuttgart.de, christian.mueller@ikr.uni-stuttgart.de

Hajo Bakker

Alcatel-Lucent Bell Labs Germany

Stuttgart, Germany

hajo.bakker@alcatel-lucent.de

**Abstract**—In cellular OFDMA networks, there exists a fundamental trade-off between the achievable cell capacity and the degree of fairness among the users in the cell. Several scheduling algorithms have been proposed which try to balance this trade-off. The parameterization of these scheduling algorithms to achieve a certain desired fairness level is non-trivial. We show that the optimal fairness parameter settings depend on the system state, such as the current cell load. Our main contribution is a design of a self-optimizing scheduler architecture which includes a controller element that dynamically adjusts the fairness parameters of the scheduler. We demonstrate that with this design, an operator-defined reference fairness level is maintained in scenarios with fluctuating load and thus cell throughput can be improved. It is applicable for a class of proportional fair scheduling algorithms and can be adapted to other algorithms allowing to influence the fairness level.

## I. INTRODUCTION

In modern cellular networks, such as the 3GPP LTE networks, channel-aware schedulers allow to exploit user diversity by preferably serving users that experience good channel conditions. A recent classification of such opportunistic schedulers can be found, for example, in [1]. A further possibility to increase throughput in OFDMA networks is achieved with frequency-selective scheduling [2]. In addition, operators seek for solutions to achieve a frequency reuse close to one to obtain a high spectral efficiency. Proposals in this direction include soft reuse schemes, reuse partitioning and other forms of inter-cell interference mitigation [3].

While all of these techniques undoubtedly increase the system capacity, care has to be taken that this improvement is not realized at the expense of users in unfavorable channel conditions. A prominent example is the max-C/I scheduler, which only allocates resources to the user with the best channel conditions. This allows achieving large gains in system capacity, but usually leads to a starvation of users at the cell borders, where channel conditions are worse than in the center area of a cell. Therefore, the max-C/I scheduling algorithm is generally considered unfair. There is thus a fundamental trade-off between a fair allocation of resources and an increase in system capacity from opportunistic scheduling.

In commercial networks, the degree of fairness is an operator's choice. Finding the parameter setting for a given scheduling algorithm such that the operator's requirements are met is non-trivial. The parameters usually depend on

scenario constraints such as the current cell load and other cell-individual boundary conditions.

In this paper, we show that the optimal scheduler parameterization for a certain class of proportional fair schedulers depends on the cell load. We argue that a base station shall be able to dynamically adapt the scheduler's parameterization to the current boundary conditions and present a design of a control element which carries out this adaptation. We show that the desired fairness level is maintained under various system conditions, irrespective of the initial parameter settings of the scheduler.

The remainder of this document is structured as follows: Section II introduces the scheduling algorithm considered here. Section III reviews some related work in this area and section IV describes our simulation model. In section V, we first present simulation results for an LTE cell with static scheduler parameterization and show the influence of varying system load. Section VI then describes the proposed controller algorithm and demonstrates that a pre-configured degree of fairness is maintained under changing load situations. Finally, section VII concludes this article.

## II. SCHEDULING ALGORITHM AND FAIRNESS

The scheduling algorithm considered here is a modified version of the Proportional Fair with Minimum/Maximum Rates (PFMR), as proposed in [4]. The key characteristics of the algorithm are briefly reproduced here:

Users are selected for a time-frequency resource (e.g. LTE resource block) by maximizing the respective scheduling weight in the following formula:

$$user = \arg \max_j \left( e^{\gamma_j T_j} \frac{R_j}{\bar{R}_j^\alpha} \right) \quad (1)$$

with  $j$  being the user index. The modification of the algorithm proposed by [4] consists in the introduction of the  $\alpha$  parameter to the proportional fair term in equation (1) to achieve tunable fairness. The  $\alpha$ -fairness parameter was initially proposed by [5] for transport layer protocols.

$R_j$  is the achievable rate in the current transmission time interval (TTI).  $\bar{R}_j$  is the exponential moving average of the rate per resource block in previous TTIs, with forgetting factor  $\beta$  and  $\mu_j$  being the rate at which user  $j$  was served in the last TTI:

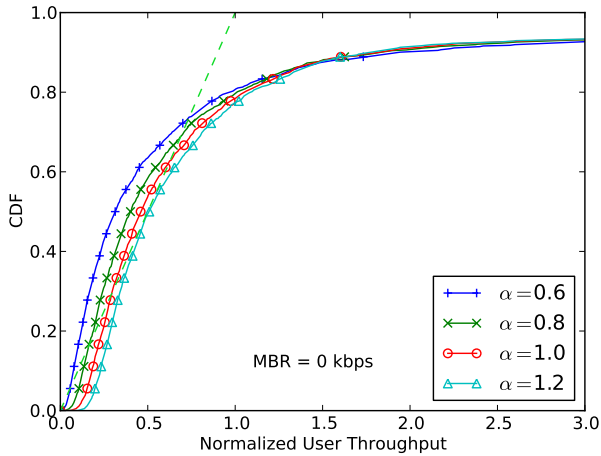


Fig. 1. Cumulative distribution function of normalized user throughput showing fairness for different  $\alpha$ -values and MBR = 0 kbps (30 Users)

$$\overline{R}_j(t+1) = (1 - \beta)\overline{R}_j(t) + \beta\mu_j(t) \quad (2)$$

As proposed in [4], the minimum bit rate (MBR) constraint is implemented with a token counter  $T_j$  and a parameter  $\gamma > 0$ , which contribute to the scheduling weight in equation (1). The token counter is updated according to:

$$T_j(t+1) = \max\{0, T_j(t) + MBR - \mu_j(t)\} \quad (3)$$

Please note that the minimum throughput requirement does not need to be met for all users at a time. It is a flexible mechanism which allows ensuring the minimum bit rate on longer time scales, i. e. multiple TTIs.

The motivation for combining PF with an MBR constraint is to allow operators to ensure a minimum degree of service to all users. The MBR has to be chosen such that basic communication is possible. This improves the availability of the operator's service even for users under unfavorable channel conditions and thus increases user satisfaction. Note that for MBR = 0 kbps the algorithm is the same as a pure PF algorithm.

While the scheduling algorithm contains several other parameters, we focus on the parameters  $\alpha$  and MBR, which both influence fairness. We expect the MBR parameter to be a fixed value that is predefined by the operator, e.g. per QoS class.

Fairness can be defined in a number of ways. Frequent assumptions on fairness constraints are a resource-fair or throughput-fair allocation of resources, meaning that all users get the same amount of frequency-time resources, respectively that all users achieve the same throughput, independent of their long-term average channel quality. The Next Generation Mobile Networks (NGMN) alliance, a consortium of different mobile network operators and vendors, has formulated a different fairness requirement [6]. It can be expressed as:

*“A system is fair, if 100-x% of the users achieve at least x% of the normalized user throughput”*

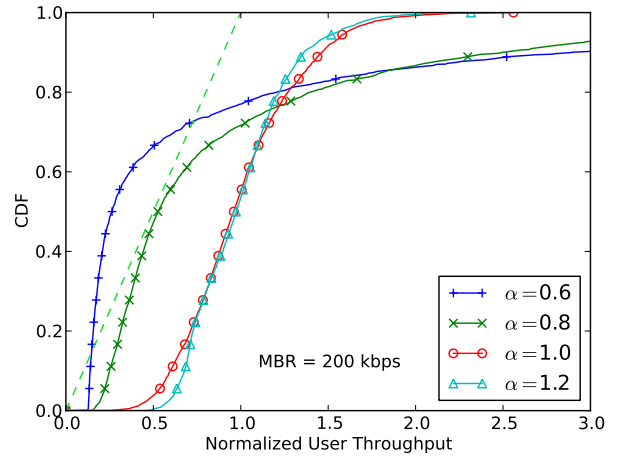


Fig. 2. Cumulative distribution function of normalized user throughput showing fairness for different  $\alpha$ -values and MBR = 200 kbps (30 Users)

This metric has also been adopted by 3GPP as one of the performance metrics for the verification of the LTE performance gains compared to HSPA [7]. For the work presented here, we use the NGMN fairness metric as our reference.

### III. RELATED WORK

The task of a well known PF scheduler proposed e.g. in [8] is the balancing of the trade-off between fairness and system capacity: the current channel quality is weighted with a long-term average of the channel quality seen by a user. Such a scheduler is considered long-term fair. The authors of [9] extend a PF scheduler by additional weighting factors to adjust the level of fairness. They discuss the trade-off between fairness and system throughput for an OFDMA system, but do not adapt these weighting factors in a running system. In [4], the authors show that the parameterization of PF-like schedulers depends on cell load, but also do not consider dynamic adaptation of these parameters.

Within the work item Self-Organizing Networks (SON), 3GPP is currently studying and specifying concepts for self-configuring and self-optimizing of the Evolved Universal Terrestrial Radio Access Network (E-UTRAN) [10]. Base station internal algorithms, i.e. the problem addressed here are not yet covered by 3GPP or NGMN.

The EU funded project Socrates presented the results of scheduler parameter optimization to the NGMN SON-subgroup [11]. They analyzed the impact of the moving average parameter  $\beta$  and came to the conclusion that the potential for self-optimization of  $\beta$  is not very significant. This is consistent to the findings of the authors and the reason why we focus on  $\alpha$  and MBR parameters. To the best of our knowledge, so far no other publication deals with adaptive configuration of  $\alpha$ -fairness parameters for wireless schedulers.

### IV. SIMULATION MODEL

Our simulator is based on the IKR Simulation Library [12]. We use a seven-site scenario consisting of a center cell and

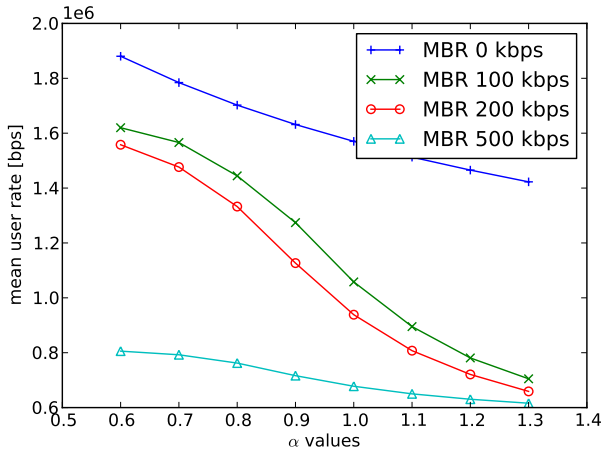


Fig. 3. Mean user rate for varying MBR and  $\alpha$  (30 Users)

one tier of hexagonally arranged interfering cells. In the simulation, we evaluate the scheduling in the center cell. We assume omnidirectional antennas and a reuse one scenario with full buffers, which means that all base stations transmit at full power and always have enough data to occupy the whole bandwidth. As we want to analyze the scheduling behavior in the base station, only the downlink is considered. Feedback which is sent in the uplink direction is assumed to be ideal.

The simulation is organized in so-called “drops”. For each drop, user locations in the center cell are chosen randomly with uniform distribution. The simulation time of a drop consists of a warm-up phase of 150 ms and a measurement phase of 1 s, respectively 1000 TTIs. On this short time-scale, it is possible to assume a fixed user location, while user mobility is modeled for the radio channel by means of the Doppler shift. To get statistically meaningful results, a large number of drops (60) are simulated for each parameter setting.

The channel is modeled by path loss, shadowing and fast fading. Hereby, path loss and shadowing are constant during one drop. Fast fading is modeled by a varying Rayleigh fading trace and the assumption of the ETSI Vehicular A model [13]. A summary of simulation parameters can be found in table I.

In LTE, the number of users which can be scheduled in one TTI is limited due to signaling overhead. For our simulations, we assume that at most 10 users can be scheduled in the same time-slot for a system bandwidth of 10 MHz. For all users, a scheduling weight is calculated for each resource block depending on the channel quality the respective user perceives in the current resource block. The mechanism to select the users to be scheduled in a TTI compares the top five resource block weights of all users. The users with the largest sum of these weights are selected for scheduling. Resource allocation is done in a subsequent step and compares just the weights of the respective resource block.

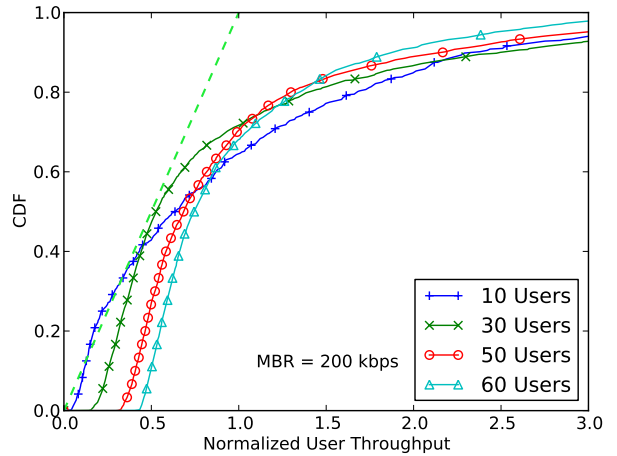


Fig. 4. Influence of system load on throughput and fairness ( $\alpha = 0.8$ , MBR = 200 kbps)

TABLE I  
SIMULATION PARAMETERS

Parameters	Model Assumptions
Baseline parameters	3GPP [14] / NGMN [6]
Inter-site distance	1000 m
Channel model	COST 231 HATA path loss Gudmundson shadowing Rayleigh Fading (Vehicular A)
Shadowing std. deviation	8 dB
Bandwidth	10 MHz
Central frequency	2 GHz
Doppler shift	10 Hz
SINR clipping	-5 dB < SINR < 26 dB
Scheduling granularity	1 ms · 180 kHz (1 resource block)

## V. SCHEDULER PERFORMANCE WITH STATIC PARAMETERIZATION

Before describing the design of a controller which is able to adjust the scheduling parameters, we first evaluate the influence of the parameters MBR and  $\alpha$  on the scheduler performance, especially throughput and fairness.

### A. Influence of scheduling parameters

The selected scheduling algorithm has two parameters, namely MBR and  $\alpha$ , which both influence fairness and which complement each other. This behavior can be seen in Fig. 1 and 2 which show the cumulative distribution function (CDF) of the user rates, normalized to the mean user rate. In the CDF plot of the normalized user throughput, the NGMN fairness requirement is represented by a line through the origin (dashed line in Fig. 1). A CDF lying completely on the right hand side of the NGMN fairness requirement is considered fair.

Fig. 1 shows the results for a pure PF scheduler without MBR-constraints and 30 active users in the center cell. We can see that fairness increases with increasing  $\alpha$ . To achieve a fair system in this scenario,  $\alpha \approx 1.2$  is needed. For Fig. 2, the

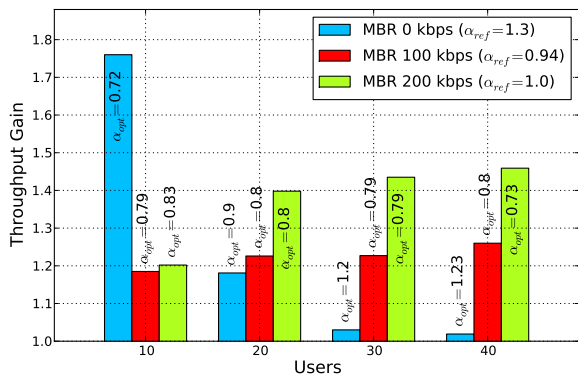


Fig. 5. Potential controller gains w.r.t. system throughput in comparison to a static parameterization of  $\alpha$

MBR parameter has been changed to 200 kbps. This largely influences the impact of  $\alpha$  on the system's fairness. With this MBR setting,  $\alpha = 0.8$  is already sufficient for a fair system.

Fairness always comes at the cost of a loss in overall cell throughput. This can be seen in Fig. 3, where the mean user rate for different MBR settings is plotted against  $\alpha$ . Both, an increase of MBR and  $\alpha$ , lead to a reduction in the mean user rate and thus also to a reduction of the total cell throughput. The different slope of the curves in Fig. 3 shows the interworking between MBR and  $\alpha$ . The more fairness is ensured by a higher MBR, the lower is the impact of an increased  $\alpha$  (and the other way round). For a pure PF-scheduler, the influence of increasing  $\alpha$  is still visible but less dominant.

### B. Influence of system load

System load here is defined as the number of active users in the cell. With an increasing number of users, the scheduler has more difficulties in satisfying all user demands.

Fig. 4 demonstrates what happens, if the scheduling parameters remain constant while the load changes. The influence of the MBR constraint increases with the number of users, because the amount of granted resources, given by the product users  $\cdot$  MBR, increases. This means that for a constant  $\alpha$  and MBR, the system tends to get fairer with increasing number of users. As a consequence, the operator-defined fairness requirement is exceeded and the cell throughput is lower than it could be with a smaller value of the  $\alpha$  parameter.

System load is just one boundary condition. Other boundary conditions like cell geometry, user distribution or sudden changes in channel conditions, for example, can influence the optimal parameter setting as well. The set of boundary conditions is individual for each cell.

### C. Estimated Gain from Adaptive Parameterization

For the case of changing system load, we can derive a rough estimate of the potential gain in terms of system throughput from adaptive parameterization under the assumption that the operator wants to ensure a certain fairness level all the time. To get a reference value of  $\alpha$  in a static system for a given

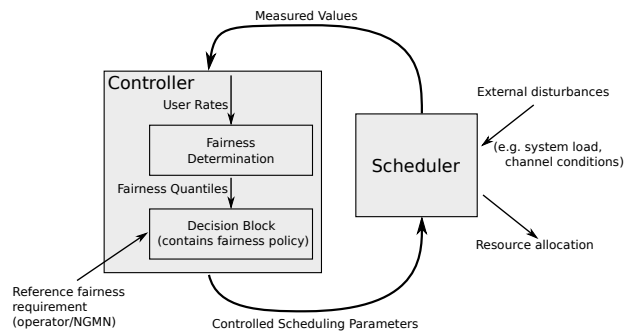


Fig. 6. Design of controller internal blocks and the closed-loop feedback system

MBR, we choose the  $\alpha$ -value that showed to be fair for all user configurations we simulated. This corresponds to a very conservative parameter setting, such that the system is fair over a wide range of scenarios. We calculate the gain of adaptive parameterization as the ratio between the cell throughput achieved with optimal parameter setting  $\alpha_{opt}$  and the cell throughput of the static setting  $\alpha_{ref}$ .

The estimated gains for various system loads in terms of different numbers of users are given in Fig. 5. Please note that these gains constitute an upper bound due to idealized signaling assumptions and the rather conservative choice of the reference  $\alpha$ -value.

## VI. SCHEDULER PERFORMANCE WITH ADAPTIVE PARAMETERIZATION

From section V, we have seen that the optimal value of  $\alpha$  for a given MBR depends on the system load. The results suggest that significant gains in cell throughput can be achieved with an adaptive  $\alpha$  parameter. In this section, we demonstrate the design of a controller that dynamically adjusts  $\alpha$  and present simulation results showing its benefits.

### A. Controller Design

Fig. 6 shows the basic structure of our controller. An important issue is to design the algorithms to be robust against external disturbances. It is thus designed as a closed-loop feedback system.

The scheduler constitutes the executing and observing entity. It reports the user rates to the controller, which in-turn adapts the parameter setting of the scheduler.

The advantage of measuring the user rates is that no additional measurements are needed, given that the user rates are already known to the scheduler. The user rates are processed in the controller to obtain the fairness metric. For the evaluation of the NGMN fairness requirement, the quantiles of the user rates are needed, but other metrics are imaginable, too. The fairness metric is then forwarded to the decision block, where it is matched against the fairness policy which is provided by the operator. The decision block reacts on a fairness mismatch by changing the  $\alpha$  parameter of the scheduler.

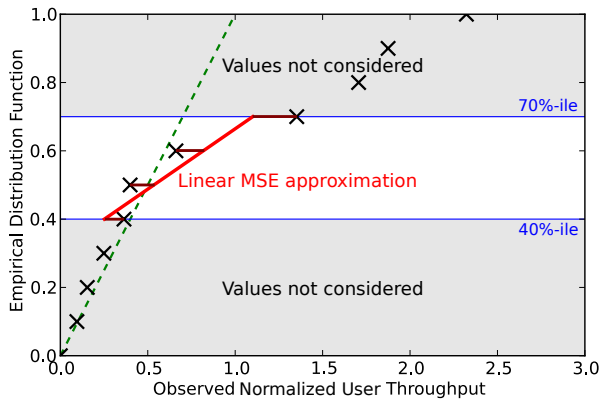


Fig. 7. Fairness determination with user rates observed during a sampling interval

It is important to state that the controller operates on a longer timescale than the scheduler. The controller’s task is to adapt the scheduler’s parameters to changes in cell load and other external constraints. It should hence operate on a timescale comparable to the variation of these external constraints. If both the scheduler and the controller were operating at the same timescale, the whole system would be prone to instability and oscillations.

A further possible enhancement of the controller is the handling of different traffic classes with different fairness/QoS-requirements. In this case, multiple instances of the controller exist in parallel and the scheduler instance has to perform the reporting of user rates on a per traffic class basis.

### B. Fairness determination & regulation

In our implementation, we transform the observed user rates into rate quantiles to check them against the fairness requirement. This is done continuously by dividing the observations into sampling intervals. At the end of an interval, the quantiles of the user rates measured in this interval are calculated and matched against the fairness requirement. If necessary,  $\alpha$  is decreased or increased.

During one sampling interval it may occur that only very few users (e.g. in the order of ten) are active in the cell. This means that the obtained empirical distribution function is very coarse-grained. To cope with this problem, we do a linear approximation of the CDF between the 40%-ile and the 70%-ile, where the slope of the CDF is known to fit rather well to the fairness requirement. The approximation is done by determining the linear function minimizing the mean square error to the observed quantiles in the specified range as demonstrated in Fig. 7. The advantage of this method is, that outliers of very good and very bad channels are neglected, which makes the controller decision more robust.

We determine the step size for changes of  $\alpha$  in relation to the area between the fairness observation and the fairness requirement. If the approximated line crosses the fairness

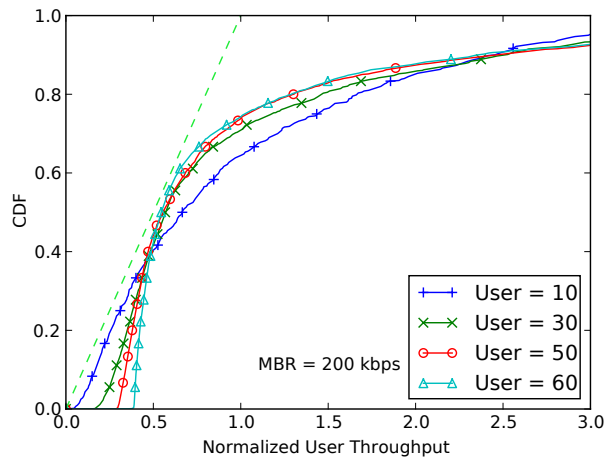


Fig. 8. Cumulative distribution function showing throughput and fairness with adaptive  $\alpha$  (MBR = 200 kbps)

requirement between the specified quantiles or lies completely left of the fairness requirement, the area left of the requirement is considered and  $\alpha$  is increased relatively. If the line lies completely right of the fairness requirement in between the regarded quantiles,  $\alpha$  is decreased proportionally to the lowest distance between observation and requirement. This way, it is possible to stabilize the control-loop around the optimal point of operation.

### C. Simulation Results

To evaluate the performance of our controller, we conducted simulations with different load situations. The controller sampling interval is 100 ms. In order to allow the feedback system to stabilize, we simulated each drop with 1 s warm-up phase and 10 s simulation time. The scenario considered for the CDFs in Fig. 8 is identical to the one of Fig. 4, except for the adaptive parameterization. It can be seen that the controller achieves a fairness level very close to the NGMN requirement, irrespective of the initial settings.

Figure 9 demonstrates the capability of the scheduler to adapt to changing load situations. It shows the evolution of  $\alpha$  over the simulation time with a sampling interval of 1 s. After 50 s, the number of active users changes from 15 to 30, requiring a different fairness setting. For a pure PF-scheduler (without MBR),  $\alpha$  needs to be increased to maintain the fairness requirement. For an MBR of 200 kbps or 400 kbps,  $\alpha$  can be reduced, which increases system throughput while the fairness requirement is preserved. The gain with respect to a static configuration ranges between 7% and 14% in this scenario. This is because a static configuration is sub-optimal half of the time.

The evolution of the controlled  $\alpha$ -parameter shows that the controller is able to adapt the scheduler to changes in the load situation and to long-term changes in radio propagation conditions.

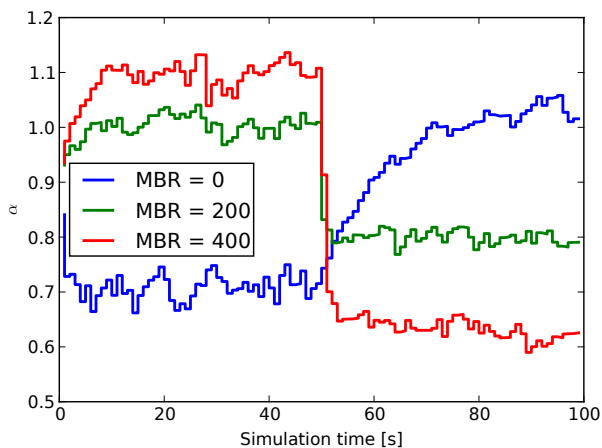


Fig. 9. Evolution of the  $\alpha$ -parameter; The number of active users changes at  $t = 50$ s from 15 to 30.

#### D. Stability of the Controller

In order to test the stability of the controller, we also simulated fluctuating traffic instead of full buffer traffic. It showed that fast fluctuations of users, i.e. many users being switched off and new users being switched on, may impair the fairness measurement and lead to a distorted regulation. To cope with this problem, our simulations showed that it is sufficient to increase the sampling interval to a reasonable value. It is enough to have the sampling interval in the order of seconds. Then, all channel oscillations and user fluctuations will be attenuated sufficiently.

Furthermore, the stability of the controller depends on the stability of the scheduling algorithm. If the fairness adjusting parameter of the scheduler is not able to increase or decrease fairness as expected anymore, also the controller actions will not have desired effect anymore. In this case, however, this will not lead to an unstable behavior. If the system is very fair (e.g. because all users have comparable channel conditions),  $\alpha$  will be decreased to zero where it is bounded. This means the scheduler behaves as a max-C/I scheduler, which is the best choice in this case. In the other direction, a larger  $\alpha$  always leads to a fairer system up to equal-rate scheduling (which is inherently fair), so the controller will remain stable. If the controller is to be applied for different scheduling algorithms, the stability of the respective scheduler needs to be considered.

## VII. CONCLUSIONS

The parameterization of a scheduler has strong impact on the fairness perceived by the users. In accordance with other authors, we showed that the optimal parameterization of a scheduler that meets certain operator-defined fairness requirements depends on external, cell-individual constraints such as the current number of users.

We proposed the application of a controller that is able to automatically adjust the scheduler parameters and estimated the potential gain from such an adaptive parameterization. We presented a basic implementation of a controller which

continuously adapts fairness parameters of an  $\alpha$ -fair scheduler to the current load situation of a cell. This allows a cell to always deliver maximum throughput, while the service-class dependent fairness requirements set by the operator are preserved. The estimated gains were in good accordance with the results of the actual controller implementation. Please note that the demonstrated controller concept is not limited to the specific scheduling algorithm considered here. Any scheduling algorithm offering parameters to control the shape of the fairness CDFs can be controlled in a similar way.

This self-optimization of scheduler parameters is especially advantageous in actual network deployments with irregular cell shapes, where the optimal parameter setting differs from cell to cell. Consequently, high gains can be expected for the overall system throughput while still maintaining a predefined fairness level.

## REFERENCES

- [1] B. Sadiq, R. Madan, and A. Sampath, "Downlink scheduling for multiclass traffic in LTE," *EURASIP Journal on Wireless Communications and Networking*, 2009.
- [2] G. Wunder, C. Zhou, H.-E. Bakker, and S. Kaminski, "Throughput maximization under rate requirements for the OFDMA downlink channel with limited feedback," *EURASIP Journal on Wireless Communications and Networking*, 2008.
- [3] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4g networks - [lte part ii: 3gpp release 8]," *Communications Magazine, IEEE*, vol. 47, no. 4, pp. 74–81, 2009.
- [4] M. Andrews, L. Qian, and A. Stolyar, "Optimal utility based multi-user throughput allocation subject to throughput constraints," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 4, March 2005, pp. 2415–2424 vol. 4.
- [5] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *Networking, IEEE/ACM Transactions on*, vol. 8, no. 5, pp. 556–567, 2000.
- [6] R. Irmer, *Radio Access Performance Evaluation Methodology*, Next Generation Mobile Networks Std. V 1.3, January 2008.
- [7] *LTE physical layer framework for performance verification*, 3GPP TSG-RAN148 Std. R1-070 674, February 2007.
- [8] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of cdma-hdr a high efficiency-high data rate personal communication wireless system," in *Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo. 2000 IEEE 51st*, vol. 3, 2000, pp. 1854–1858 vol.3.
- [9] C. Wengert, J. Ohlhorst, and A. von Elbwart, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA," in *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, vol. 3, may 2005, pp. 1903–1907.
- [10] *E-UTRAN Self-configuring and self-optimizing network use cases and solutions*, 3GPP Std. TR 36.902, Rev. V 9.0.0, October 2009.
- [11] R. Litjens, I. F. Diaz, H. van den Berg, K. Spaey, E. U. Warriach, and D. Dimitrova, "Packet scheduling parameter optimisation use case," <http://www.fp7-socrates.org/?q=node/25>, February 2010.
- [12] Institute of Communication Networks and Computer Engineering, "IKR Simulation Library," <http://www.ikr.uni-stuttgart.de/Content/IKRSimLib/>, Stuttgart, Germany, 2010.
- [13] *Selection procedures for the choice of radio transmission technologies of the UMTS*, ETSI Std. TR101 112/UMTS30.03 V3.2, April 1998.
- [14] *Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA)*, 3GPP Std. TR 25.814, Rev. V7.1.0, September 2006.