Akbudak et al. EURASIP Journal on Wireless Communications and Networking 2011, **2011**:67
http://jwcn.eurasipjournals.com/content/2011/1/67

EURASIP Journal on
Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

A cross-layer resource allocation scheme for spatial multiplexing-based MIMO-OFDMA systems

Tarik Akbudak[1*], Hussein Al-Shatri[2] and Andreas Czylwik[1]

Abstract

We investigate the resource allocation problem for the downlink of a multiple-input multiple-output orthogonal frequency division multiple access (MIMO-OFDMA) system. The sum rate maximization itself cannot cope with fairness among users. Hence, we address this problem in the context of the utility-based resource allocation presented in earlier papers. This resource allocation method allows to enhance the efficiency and guarantee fairness among users by exploiting multiuser diversity, frequency diversity, as well as time diversity. In this paper, we treat the overall utility as the quality of service indicator and design utility functions with respect to the average transmission rate in order to simultaneously provide two services, real-time and best-effort. Since the optimal solutions are extremely computationally complex to obtain, we propose a suboptimal joint subchannel and power control algorithm that converges very fast and simplifies the MIMO resource allocation problem into a single-input single-output resource allocation problem. Simulation results indicate that using the proposed method achieves near-optimum solutions, and the available resources are distributed more fairly among users.

Keywords: Cross-layer optimization, Utility-based resource allocation, MIMO-OFDMA, Water-filling

I Introduction

Exploiting the channel variation across users, channel-aware resource allocation can substantially improve network performance through multiuser diversity [1]. The key idea is to select those users having the best channel condition at each individual subchannel independently. This maximizes the sum rate as well as spectral efficiency. However, sum rate maximization is sometimes unfair to cell-edge users or those with bad channel conditions [2] and thus cannot guarantee their quality of service (QoS) requirements. On the other hand, absolute fairness may decrease efficiency and system capacity. Therefore, a practical resource allocation scheme should carefully tradeoff efficiency versus fairness. As a result, joint channel- and QoS-aware resource allocation would be more beneficial compared to channel-aware resource allocation.

In this paper, we consider a single-cell of a cellular orthogonal frequency division multiple access (OFDMA) network with multiple types of services, namely best-effort and real-time, which are distinguished by their required QoS. For each service type, we introduce a utility function depending on the average transmission rate in order not only to balance fairness and efficiency but also to achieve cross-layer optimization. The overall network utility, which is the sum of the utilities of all users, is then treated as the optimization objective. For the considered problem, we propose a joint sub-carrier and power allocation algorithm that simplifies the multiple-input multiple-output (MIMO) resource allocation into a single-input single-output (SISO) resource allocation problem. By employing the proposed algorithm, it will be shown that real-time users get higher priorities than best-effort users unless their rate constraints are satisfied. On the other hand, after reaching required rates, lower priorities are given to real-time users in order to maximize the sum rate of best-effort users, thus preventing a possible waste of resources.

The rest of the paper is organized as follows. The relevance of this work to the state-of-the-art of resource allocation techniques in wireless networks is highlighted in Section II. In Section III, we describe the system model and formulate the resource allocation problem. In Section IV, we give the optimal solution for the

* Correspondence: akbudak@nts.uni-duisburg-essen.de
[1]Department of Communication Systems, University of Duisburg-Essen, Bismarckstr. 81, 47057 Duisburg, Germany
Full list of author information is available at the end of the article

subchannel and power allocation problem considered. The proposed resource allocation algorithm is presented in Section V. Next, in Section VI, we present performance evaluation results. Finally, conclusions are drawn in Section VII.

II Related work

Utility theory is a well-known theory in economics where fair and efficient resource allocation is an essential task. Utility functions are used to quantify the level of customer satisfaction or the benefit of usage of certain resources. In communication networks, utilities can be used to evaluate the degree to which a network satisfies service requirements of users' applications [3]. In wireless networks, utility-based resource allocation in code division multiple access (CDMA) networks has been analyzed in [4] and [5]. In [6], a utility-based power control in CDMA downlink for voice and data applications has been proposed.

The optimal resource allocation problem in OFDMA systems has been analyzed in [7] and [8]. In [7], the authors derived some criteria for subcarrier assignment with the goal of maximizing the instantaneous capacity. Furthermore, they converted the MIMO channel matrix into SISO channels, thus allowing a simplified resource allocation as in the SISO case. In [8], the authors proposed an algorithm which maintains proportional rates among users for each channel realization and ensures the instantaneous rates of different users to be proportional. However, due to the strict proportionality, the utilization of subcarriers is low and thus decreasing the overall sum rate. Considering the same problem formulation, two types of users' applications, best-effort (BE) and guaranteed-performance (GP), were distinguished on the basis of required QoS in [9]. The proposed method maximizes the sum capacity of BE users subject to rate constraints of GP users.

Utility-based resource allocation in OFDMA wireless networks has been studied in [10-12] and [13]. In [12] and [13], the authors considered a gradient-based scheduling algorithm which maximizes the weighted sum rate at the beginning of each scheduling interval. A user's weight is defined as the gradient of that user's utility function with respect to average throughput. Considering multiple types of traffic and QoS requirements, a joint dynamic subcarrier and power allocation scheme has been proposed in [10]. It was shown that using such a resource allocation scheme can balance efficiency and fairness. Similarly, the authors have studied different queue- and channel-aware schedulers for the 3GPP LTE downlink in [11]. They presented a practical scheduler and characterized its performance for three different traffic scenarios, namely, full-buffer, streaming video and live video. In [14] and [15], the utility is exploited to balance fairness and efficiency by jointly optimizing the physical and medium access control (MAC) layer. This results in data rate adaptation over the subcarriers with corresponding channel conditions, thus increasing throughput while simultaneously maintaining an acceptable BER. Furthermore, various utility-based optimization schemes, including the joint dynamic subcarrier assignment (DSA) and adaptive power allocation (APA), have been proposed in [14].

III Problem formulation

A System model

We consider the downlink of a single-cell OFDMA network, in which the transmitter (base station) is equipped with N_T transmit antennas and K receivers (users) are equipped with N_R receive antennas. At the base station, a maximum total transmission power of P_max watts and S subchannels are available for transmission.

Assuming that the total power available for subchannel s is distributed equally across spatial channels, the base station can obtain the achievable rate of user k over subchannel s, denoted by $r_{k,s}$ as

$$\begin{aligned} r_{k,s} &= \log_2\left[\det\left(\mathbf{I}_{N_\mathrm{R}} + \beta_k \frac{p_s}{N_\mathrm{T}N_{k,s}}\mathbf{H}_{k,s}\mathbf{H}_{k,s}^\mathrm{H}\right)\right]\Delta f \\ &= \sum_{i=1}^{n}\log_2\left(1+\beta_k\frac{p_s}{N_\mathrm{T}N_{k,s}}\lambda_{k,s,i}\right)\Delta f, \end{aligned} \tag{1}$$

where $\mathbf{H}_{k,s}$ and $N_{k,s}$ are the $N_\mathrm{R} \times N_\mathrm{T}$ channel frequency response matrix and noise-plus-interference power at user k and subchannel s, respectively. p_s is the transmission power allocated to subchannel s, and Δf is the bandwidth of a single subchannel. det(·) represents the determinant operator, $\lambda_{k,s,i}$ denotes the i-th eigenvalue of matrix $\mathbf{H}_{k,s}\mathbf{H}_{k,s}^\mathrm{H}$, and $n = \min(N_\mathrm{R}, N_\mathrm{T})$. Furthermore, β_k is a constant related to the target BER of user k by [16] $\beta_k = \frac{1.5}{-\ln(5BER_k)}$, and which indicates the approximated ratio between the SNR needed to achieve a certain rate for a practical system and the theoretical limit [10]. Note that the co-channel interference of neighboring cells is modeled as additive Gaussian noise in the formulation above. By Jensen's inequality [7], $r_{k,s}$ satisfies

$$r_{k,s} \le n\log_2\left(1+\gamma_{k,s}p_s\|\mathbf{H}_{k,s}\|_\mathrm{F}^2\right)\Delta f, \tag{2}$$

where $\|\cdot\|_\mathrm{F}$ is the frobenius norm and $\gamma_{k,s} = \beta_k/(N_\mathrm{T}N_{k,s}n)$. Note that (2) gives the upper bound for the achievable rate over a subchannel, thus delivering a simplified solution to (1) similar to the SISO case.

B Utility-based resource allocation

The objective of the utility-based resource allocation is to maximize the sum of the utilities U_k (·) in a network, where U_k (·) is an increasing/decreasing function of a given parameter such as instantaneous rate R_k, delay D_k, etc. of user k. From a user's point of view, the average rate $\bar{R}_k$ during a certain period of time is a relatively important QoS parameter [14] and can be smoothed by an exponentially weighted low-pass filter as

$$\bar{R}_k[\nu] = \frac{T_S}{T_W} R_k[\nu] + \left(1 - \frac{T_S}{T_W}\right) \bar{R}_k[\nu - 1], \tag{3}$$

where R_k [ν] is the instantaneous rate of user k and defined as sum of the rates over the subchannels assigned to user k at time instant ν. T_S and T_W are the time slot and the filter window length, respectively. Considering the utilities with respect to the average rate at time instant ν, $U_k(\bar{R}_k[\nu])$, the utility-based resource allocation decision can be given according to the gradient-based scheduling [12] as

$$\max_{\mathbf{R}[\nu] \in \mathcal{R}(\mathbf{H}[\nu])} \sum_{k=1}^{K} U'_k(\bar{R}_k[\nu - 1]) R_k[\nu], \tag{4}$$

where $U'_k(\cdot)$ is the derivative of U_k (·) and called the marginal utility function of user k. The objective of the above formulation is to select a rate vector $\mathbf{R}[\nu]$ = (R_1 [ν], $R_2[\nu]$, ..., R_K [ν]) from the instantaneous feasible rate region $\mathcal{R}(\mathbf{H}[\nu])$, where $\mathbf{H}[\nu]$ denotes the time-varying channel state information (CSI) available at time instant ν.

Since all $\bar{R}_k[\nu - 1]$'s are fixed at time instant ν, we can omit the time index ν to simplify the notations. Hence, the optimization problem in (4) can be considered as a weighted sum rate maximization, which can be given according to the above formulation as

$$\max_{\alpha_{k,s} p_s} \sum_{k=1}^{K} w_k \sum_{s=1}^{s} \alpha_{k,s} n \log_2 (1 + \gamma_{k,s} p_s ||\mathbf{H}_{k,s}||_F^2) \Delta f$$

subject to:

$$\begin{aligned} &\alpha_{k,s} = \{0, 1\} \forall k,\ s \\ &\sum_{s=1}^{s} p_s \le P_{\max} \\ &\sum_{k=1}^{K} \alpha_{k,s} = 1 \forall s, \end{aligned} \tag{5}$$

where $w_k \geq 0$ is a time-varying scheduling weight assigned to user k and is adaptively controlled by the marginal utility function with respect to the current average rate. $\alpha_{k,s}$ indicates whether or not subchannel s is allocated to user k. The second constraint gives an upper bound for the overall transmission power available at the transmitter, denoted by $P_{\max}$. Moreover, the last constraint states that each subchannel can only be allocated to one user at any given time.

The above optimization problem is a mixed binary integer programming problem, since it involves both binary and continuous variables. Furthermore, such an optimization problem is neither convex nor concave with respect to ($\alpha_{k,s}$, p_s) and thus extremely hard to solve.

IV Optimal subchannel and power allocation

To make it easier to solve the problem, the original maximization problem in (5) can be transformed into a minimization problem as [17]

$$\min_{\alpha_{k,s}, \bar{p}_{k,s}} - \sum_{k=1}^{K} w_k \sum_{s=1}^{s} \alpha_{k,s} n \log_2 \left(1 + \tfrac{\gamma_{k,s} \bar{p}_{k,s}}{\alpha_{k,s}} ||\mathbf{H}_{k,s}||_F^2\right) \Delta f$$

subject to:

$$\begin{aligned} &\sum_{k=1}^{K} \sum_{s=1}^{s} \bar{p}_{k,s} \le P_{\max} \\ &\sum_{k=1}^{K} \alpha_{k,s} = 1 \forall s \\ &0 \le \alpha_{k,s} \le 1 \forall k,\ s\ . \end{aligned} \tag{6}$$

The first constraint in (5) is relaxed in such a way that it is a real number on the interval of 0[1]. Furthermore, we define $\bar{p}_{k,s} = \alpha_{k,s} p_s$ as the transmission power used by user k on subchannel s. The case $\bar{p}_{k,s} = 0$ corresponds to an unused subchannel for user k. The most important property of the objective function in (6) is that it is convex. The proof of convexity is given in Appendix.

Letting $\lambda \geq 0$, $\eta_s \geq 0$, $\xi_{k,s} \geq 0$ and $\mu_{k,s} \geq 0$ be the Lagrange multipliers associated with the given constraints, the Lagrangian dual of (6) can be formulated as

$$\begin{aligned} &\mathcal{L}(\bar{p}_{k,s},\ \alpha_{k,s},\ \lambda,\ \eta_s,\ \xi_{k,s},\ \mu_{k,s}) \\ &= - \sum_{k=1}^{K} w_k \sum_{s=1}^{s} n \Delta f \alpha_{k,s} \log_2 \left(1 + \tfrac{\gamma_{k,s} \bar{p}_{k,s}}{\alpha_{k,s}} ||\mathbf{H}_{k,s}||_F^2\right) \\ &\quad + \lambda \left(\sum_{k=1}^{K} \sum_{s=1}^{s} \bar{p}_{k,s} - P_{\max} \right) + \sum_{s=1}^{s} \eta_s \left(\sum_{k=1}^{K} \alpha_{k,s} - 1 \right) \\ &\quad + \sum_{k=1}^{K} \sum_{s=1}^{s} \xi_{k,s} (0 - \alpha_{k,s}) + \sum_{k=1}^{K} \sum_{s=1}^{s} \mu_{k,s} (\alpha_{k,s} - 1). \end{aligned} \tag{7}$$

The optimal solution must satisfy the Karush-Kuhn-Tucker (KKT) conditions [18], which can be given as follows:

$$\begin{aligned}&\nabla_{\alpha k,s}\mathcal{L}(\bar{p}_{k,s},\alpha_{k,s},\ \lambda,\ \eta_s,\ \xi_{k,s},\ \mu_{k,s})\\ &\quad= -w_k n\Delta f\left[\log_2\left(1+\tfrac{\gamma_{k,s}\bar{p}_{k,s}}{\alpha_{k,s}}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2\right)\right.\\ &\quad\left.-\ \tfrac{\gamma_{k,s}\bar{p}_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2}{\ln 2(\alpha_{k,s}+\gamma_{k,s}\bar{p}_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2)}\right]+\eta_s-\xi_{k,s},+\mu_{k,s}\\ &\quad=0\end{aligned} \tag{8}$$

$$\begin{aligned}&\nabla_{\bar{p}_{k,s}}\mathcal{L}(\bar{p}_{k,s},\alpha_{k,s},\ \lambda,\ \eta_s,\ \xi_{k,s},\ \mu_{k,s})\\ &\quad=\frac{-w_k\alpha_{k,s}n\Delta f\gamma_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2}{\ln 2(\alpha_{k,s}+\gamma_{k,s}\bar{p}_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2)}+\lambda=0,\end{aligned} \tag{9}$$

$$\begin{aligned}&\lambda\cdot\nabla_{\lambda}\mathcal{L}(\bar{p}_{k,s},\ \alpha_{k,s},\ \lambda,\ \eta_s,\ \xi_{k,s},\ \mu_{k,s})\\ &\quad=\lambda\left(\sum_{k=1}^{K}\sum_{s=1}^{s}\bar{p}_{k,s}-P_{\max}\right)=0,\end{aligned} \tag{10}$$

$$\begin{aligned}&\eta_s\cdot\nabla_{\eta_s}\mathcal{L}(\bar{p}_{k,s},\ \alpha_{k,s},\ \lambda,\ \eta_s,\ \xi_{k,s},\ \mu_{k,s})\\ &\quad=\eta_s\left(\sum_{k=1}^{K}\alpha_{k,s}-1\right)=0,\end{aligned} \tag{11}$$

$$\begin{aligned}&\xi_{k,s}\cdot\nabla_{\xi_{k,s}}\mathcal{L}(\bar{p}_{k,s}\alpha_{k,s},\ \lambda,\ \eta_s,\ \xi_{k,s,}\mu_{k,s})\\ &\quad=\xi_{k,s}(0-\alpha_{k,s})=0,\end{aligned} \tag{12}$$

$$\begin{aligned}&\mu_{k,s}\cdot\nabla_{\mu_{k,s}}\mathcal{L}(\bar{p}_{k,s}\alpha_{k,s},\ \lambda,\ \eta_s,\ \xi_{k,s,}\mu_{k,s})\\ &\quad=\mu_{k,s}(\alpha_{k,s}-1)=0.\end{aligned} \tag{13}$$

From (8), we define

$$\begin{aligned}\Psi_{k,s}&=w_k n\Delta f[\Phi(\bar{p}_{k,s},\alpha_{k,s})-\phi(\bar{p}_{k,s},\alpha_{k,s})]\\ &=\eta_s-\xi_{k,s}+\mu_{k,s},\end{aligned} \tag{14}$$

where $\Phi(\bar{p}_{k,s},\alpha_{k,s})$ is the logarithmic function and $\phi(\bar{p}_{k,s},\alpha_{k,s})$ is the rest function of the first term in (8). From (12) and (13), if subchannel s is allocated to user k, i.e., $\alpha_{k,s} = 1$, then $\xi_{k,s} = 0$ and $\mu_{k,s} \geq 0$. On the other hand, if subchannel s is not allocated to user k, i.e., $\alpha_{k,s} < 1$, then $\xi_{k,s} = 0$ and $\mu_{k,s} = 0$. Thus, we can write

$$\Psi_{k,s}\begin{cases}\geq\eta_s,\ \alpha_{k,s}=1\\ =\eta_s,\ \alpha_{k,s}<1.\end{cases} \tag{15}$$

From (11) and (15), it can be concluded that η_s is a constant for subchannel s of all users and subchannel s can be allocated to the user $u(s)$, who has the maximum $\Psi_{k,s}$ on that subchannel, i.e.,

$$u(s)=\arg\max_k\Psi_{k,s}. \tag{16}$$

The objective in (16) is equivalent to finding the maximum $w_k n\Delta f\Phi(\bar{p}_{k,s},\ \alpha_{k,s})$. Hence, considering (2), we can conclude that

$$\alpha_{u(s),s}=\begin{cases}1,\ u(s)=\ \arg\ \max_k\{w_k\cdot r_{k,s}\}\\ 0,\ \text{otherwise}.\end{cases} \tag{17}$$

Note that the condition in (17) corresponds to selecting the user with the maximum weighted rate for subchannel s and given the transmit power levels.

Similarly, from (9) and (10), we may obtain the well-known water-filling solution as

$$\begin{aligned}\bar{p}_{k,s}&=\frac{w_k\alpha_{k,s}}{\lambda'}-\frac{\alpha_{k,s}}{\gamma_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2}\\ &=\begin{cases}\max\left\{0,\ \frac{w_k}{\lambda'}-\frac{1}{\gamma_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2}\right\}, & \alpha_{k,s}=1\\ 0, & \alpha_{k,s}<1,\end{cases}\end{aligned} \tag{18}$$

where λ' is a constant which is a function of λ and can be obtained through substituting (18) into (10) which yields

$$\lambda'=\frac{\sum_{k=1}^{K}w_k|\Omega_k|}{\sum_{k=1}^{K}\sum_{s\in\Omega_k}\dfrac{1}{\gamma_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2}+P_{\max}}, \tag{19}$$

where Ω_k ($|\Omega_k| \leq S$) is the set of subchannels assigned to user k.

V Suboptimal power and subchannel allocation

Ideally, the subchannels and power levels must be allocated jointly to achieve the optimal solution to the optimization problem in (6). However, it is not possible to solve the considered problem in a closed form due to a prohibitive computational burden at the base station. Since the base station has to rapidly allocate the available resources as the time-varying radio channel varies, low-complexity algorithms should be chosen for effective implementations. Therefore, we propose a suboptimal resource allocation algorithm which is able to jointly allocate subchannels and power levels with a low computational complexity.

A The proposed algorithm

In order to obtain $\mathbf{U}$ and $\mathbf{P}$, which are the $S \times K$ subchannel allocation matrix with binary entries $\alpha_{k,s}$ and the power assignment matrix with continuous entries $\bar{p}_{k,s}$, respectively, the proposed algorithm requires a channel condition matrix $\mathbf{G}$ which is defined as

$$\mathbf{G} = \begin{bmatrix} \gamma_{1,1}||\mathbf{H}_{1,1}||_{\mathrm{F}}^2 & \gamma_{2,1}||\mathbf{H}_{2,1}||_{\mathrm{F}}^2 & \cdots & \gamma_{K,1}\ ||\mathbf{H}_{K,1}||_{\mathrm{F}}^2 \\ \gamma_{1,2}||\mathbf{H}_{1,2}||_{\mathrm{F}}^2 & \gamma_{2,2}||\mathbf{H}_{2,2}||_{\mathrm{F}}^2 & \cdots & \gamma_{K,2}\ ||\mathbf{H}_{K,2}||_{\mathrm{F}}^2 \\ \vdots & \ddots & \ddots & \vdots \\ \gamma_{1,S}||\mathbf{H}_{1,S}||_{\mathrm{F}}^2 & \gamma_{2,S}||\mathbf{H}_{2,S}||_{\mathrm{F}}^2 & \cdots & \gamma_{K,S}\ ||\mathbf{H}_{K,S}||_{\mathrm{F}}^2 \end{bmatrix}$$

where each row and column correspond to a subchannel and user, respectively. In the following, the various steps involved in the proposed algorithm are described:

1. Construct an $S \times K$ matrix $\tilde{\mathbf{G}}$ which is the permuted version of $\mathbf{G}$ such that the maximum entry in each row, i.e., of each subchannel, is greater than the maximum entry of the following row. This permutation allows us to start with the subchannels having better channel conditions and thus a fast convergence can be obtained.

2. For each row (subchannel) in $\tilde{\mathbf{G}}$ (i.e., s = 1, 2,..., S), letting $\alpha_{k,s} = 1$ for k = 1, 2,..., K,

(a) while considering the current subchannel s in conjunction with the previous channel allocations, get the power levels $\bar{p}_{k,s}$ for k = 1, 2,..., K according to the condition in (18) using

$$\bar{p}_{k,s} = \max\left\{0, \frac{w_k}{\lambda'} - \frac{1}{\gamma_{k,s}||\mathbf{H}_{k,s}||_{\mathrm{F}}^2}\right\}.$$

(b) While considering the current power levels $\bar{p}_{k,s}$ for k = 1, 2,..., K, allocate the current subchannel to a user according to the condition in (17) using

$$\alpha_{u(s),s} = \begin{cases} 1, u(s) = \arg\ \max_k\{w_k \cdot r_{k,s}\} \\ 0, \text{otherwise}. \end{cases}$$

3. After obtaining the subchannel allocation matrix $\mathbf{U}$ and the power assignment matrix $\mathbf{P}$, calculate the sum rate R using

$$R = \sum_{k=1}^{K} R_k = \sum_{k=1}^{K}\sum_{s\in\Omega_k} r_{k,s}.$$

4. Considering the current subchannel allocation, repeat Step (2) and Step (3) to obtain another subchannel allocation matrix $\tilde{\mathbf{U}}$, power assignment matrix $\tilde{\mathbf{P}}$ as well as the new total weighted sum rate $\tilde{R}$.

5. Check the difference between R and $\tilde{R}$.

(a) If, by doing this, the desired accuracy is reached, i.e., $|\tilde{R} - R| \leq \varepsilon$, stop the iteration and return the last allocation matrices $\mathbf{U}$ and $\mathbf{P}$.
(b) Otherwise, repeat the whole cycle from Step (2) until fulfilling the condition in Step (5a).

B Complexity analysis

Assume that the channel condition matrix $\mathbf{G}$ is previously available at the base station. The complexity of the matrix permutation in Step (1) is $\mathcal{O}(S\log S)$. The complexity of Step (2a) and Step (2b) (after all subcarriers are assigned) are $\mathcal{O}(SK)$ and $\mathcal{O}(S\log K)$, respectively. Step (3) requires $\mathcal{O}(S)$ additions and thus has a complexity of $\mathcal{O}(S)$. Therefore, the overall complexity of the posed algorithm can be roughly given as $\mathcal{O}(SK)$, which is still efficient compared to the complexity of the brute-force search over all possible combinations, $\mathcal{O}(K^S)$.

VI Performance evaluation

A QoS differentiation among users

The utility functions can be derived quantitatively through characterization of the traffic statistics of given service classes [19]. Hence, in order to maintain a stable queue for a given user k, we can derive a utility function with respect to the average rate $U_k(\bar{R}_k)$ considering the traffic statistics of the given service class.

In the following, we derive the utility functions for best-effort and real-time applications considering three normalizations: $U_k\ (0) = 0$, $U_k(\bar{R}_{\mathrm{th}}) = U_0$ and $U_k(\bar{R}_{\max}) = U_{\max}$. Here, U_0 is the basic utility when user k has a threshold average rate $\bar{R}_{th}$, and $\bar{R}_{\max}$ is the maximum average rate which fully satisfies the QoS requirement of user k.

A.1 Best-effort applications

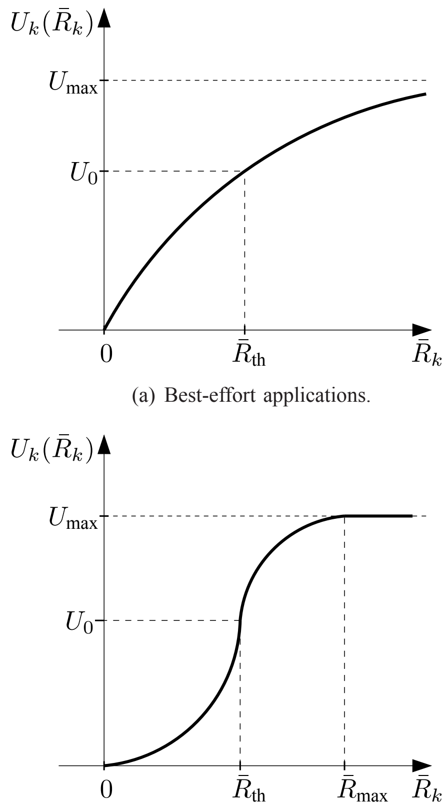
Best-effort applications, e.g., e-mail and file transfer, are delay-tolerant and thus considered as elastic applications. The elasticity of these applications can be modeled by concave utility functions [3]. Hence, we can define a utility function for best-effort applications by the following equation (see Figure 1a):

$$U_k(\bar{R}_k) = U_{\max}\left(1 - \exp\left(\log\left(\frac{(U_{\max}-U_0)}{U_{\max}}\right)\frac{\bar{R}_k}{\bar{R}_{\mathrm{th}}}\right)\right).$$

Note that it holds $\bar{R}_{\max} = \infty$ for the above function and implies that a best-effort user is fully satisfied when the average data rate goes to infinity.

A.2 Real-time applications

Compared to best-effort applications, real-time applications, e.g., voice and video applications, are rather delay-sensitive and thus considered as delay/rate-adaptive applications. Such applications can be modeled by sigmoidal-like [3] utility functions, for which a part of the utility curve is convex, representing the fact that, once the average data rate is below a certain threshold rate $\bar{R}_{th}$, satisfaction of a real-time user drops dramatically. We can define a utility function for real-time



(b) Real-time applications.

Figure 1 Utility functions with respect to average transmission rate. **a** Best-effort applications, **b** real-time applications.

applications by the following equation (see Figure 1b):

$$U_k(\bar{R}_k) = \begin{cases} U_0\left(1-\sqrt{1-\frac{\bar{R}_k^2}{\bar{R}_{\text{th}}^2}}\right), & 0 < \bar{R}_k \le \bar{R}_{\text{th}} \\ U_0 + (U_{\max}-U_0)\sqrt{1-\frac{(R_{\max}-\bar{R}_k)^2}{(R_{\max}-\bar{R}_{th})^2}}, & \bar{R}_{\text{th}} < \bar{R}_k \le \bar{R}_{\max} \\ U_{\max}, & \bar{R}_k \le \bar{R}_{\max}. \end{cases}$$

B Simulation assumptions

In all simulations we present in this paper, it is assumed that the wireless channel is a frequency-selective channel consisting of six independent Rayleigh multipaths modeled by the power delay profile of the ITU Pedestrian-B outdoor to indoor channel model [20]. Depending on the simulation scenario, each user is assumed to be stationary or moving at a speed of 3 km/h. For simplicity, co-channel interference is neglected and only receiver noise is taken into account. The length of a time slot T_{S} and the averaging filter window T_{W} are 1 ms and 1 s, respectively. All simulations are averaged over 60, 000 time slots, which correspond to 1 min in reality. Assuming an infinite number of bits for each user's queue, we consider both best-effort and real-time services and let each user have a corresponding utility function described in Section VI-A. Real-time users are assumed to have a mean source rate ($\bar{R}_{th}$) of 96 kbps and a maximum source rate ($\bar{R}_{\max}$) of 144 kbps. For best-effort users, there are no rate requirements. However, we assume a threshold rate of 512 kbps for the minimum user satisfaction. Furthermore, we set U_0 = 5 and $U_{\max}$ = 10 for both service classes. Other important simulation parameters are given in Table 1. Note that the non-concavity of the utility functions may affect the solutions. Hence, such functions can be modified to deal with this problem as in [14].

C Simulation results

Firstly, we evaluate the optimality of the proposed iterative resource allocation algorithm. To this end, we compare the performance of the proposed algorithm to that of Algorithm 4 in [14], whose computational complexity was also given as $\mathcal{O}(SK)$. The desired accuracy for both algorithms (ε) is assumed to be 10^{-3}. Furthermore, we compare the performance of the proposed algorithm to the brute-force search, which delivers the optimal solution among K^S possible resource allocation combinations, and to that of the case, where the water-filling solution in (18) is used assuming a fixed subchannel allocation which is selected randomly among all possible combinations at each time slot. Since this resource allocation scheme requires no iteration, we call it "*non-iterative selection*".

Due to the computational overhead caused by the brute-force search, the number of users in this simulation is fixed to 6. Each user is assumed to be stationary, thus has fixed path-loss and shadowing values. We divide the 6 users into 2 groups, best-effort and real-time users. Each group consists of 3 users which are sorted according to their distances to the base station so

Table 1 Simulation parameters

Parameter	Value
Cell radius	1 (km)
Channel bandwidth	1.08 (MHz)
Total number of subcarriers	72
Total number of subchannels (S)	6
Maximum Tx power ($P_{\max}$)	20 (W) (43 (dBm))
Antenna gain	0 (dBi)
Log-normal shadowing (σ)	8 (dB)
Path-loss factor (d in [m])	28.6 + 35 log(d)(dB)
Noise figure	9 (dB)
Thermal noise density	-174 (dBm/Hz)
Target BER	1%
Antenna configuration ($N_{\text{R}} \times N_{\text{T}}$)	2 × 2

that the path-loss difference between the closest to and farthest from the base station is 22 dB.

From Figures 2 and 3, it is clear that the proposed resource allocation algorithm outperforms Algorithm 4 in [14] and achieves a performance quite close to that of the brute-force search, which always delivers the optimal solution to the optimization problem considered. Furthermore, it can be seen from the figures that due to different path-loss values, different best-effort users experience different rates. However, this difference is quite low for the real-time users. This confirms the fact that utility-based resource allocation is able to differentiate between different types of users. Even for the case where random subchannel allocation is assumed, i.e., non-iterative selection, a certain degree of fairness between users can be obtained by using the utility-based water-filling.

Next, we evaluate the fairness and efficiency of the proposed iterative resource allocation algorithm considering a more realistic scenario. During this simulation, we assume that the number of users is always an even integer and half of users are using the same service class. Furthermore, each user is assumed to be moving at a speed of 3 km/h in a random direction. Assuming 4 randomly placed users initially, we increase the number of users up to 36 by randomly placing 2 additional users at a time.

It can be seen from Figures 4 and 5 that as the number of users increases and the average rate and utility of best-effort users drop dramatically since the resources get more and more scarce. However, there is only a minor decrease for real-time users. This shows that the proposed iterative algorithm gives higher priorities to real-time users and thus can maintain the performance of the users having QoS requirements.

VII Conclusion

In this paper, we investigated the resource allocation problem for the downlink of a spatial-multiplexing-

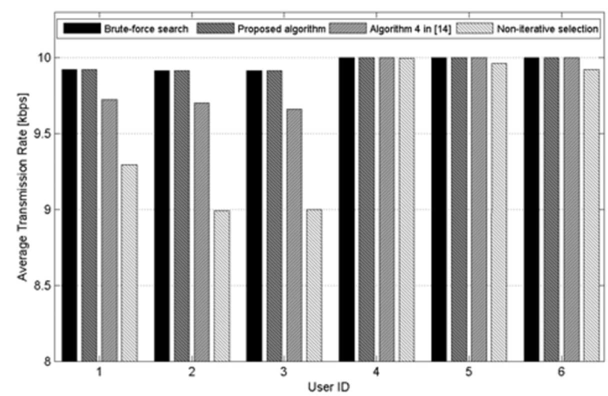
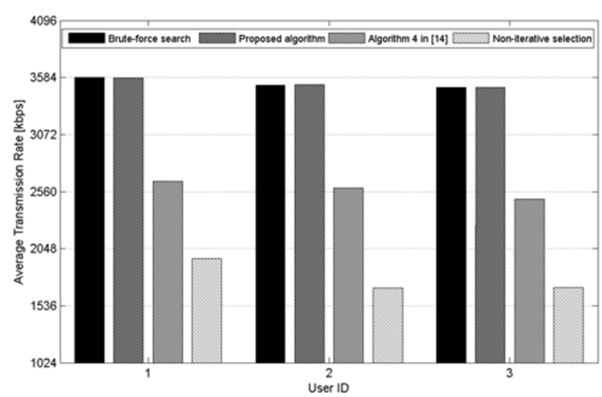
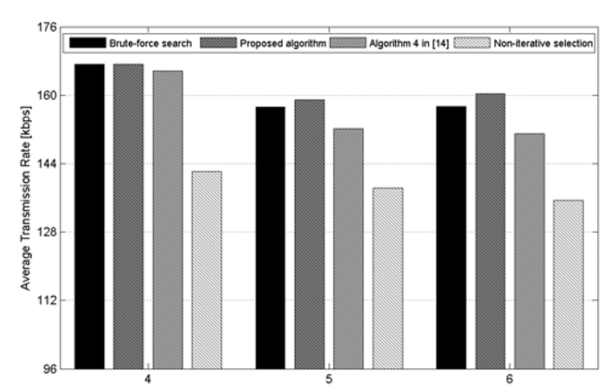


Figure 3 Average utility of various resource alloction schemes. First and last three users correspond to best-effort and real-time users, respectively.



(a) Best-effort users.



(b) Real-time users.

Figure 2 Average rate of various resource allocation schemes for different types of users. **a** Best-effort users, **b** real-time users.

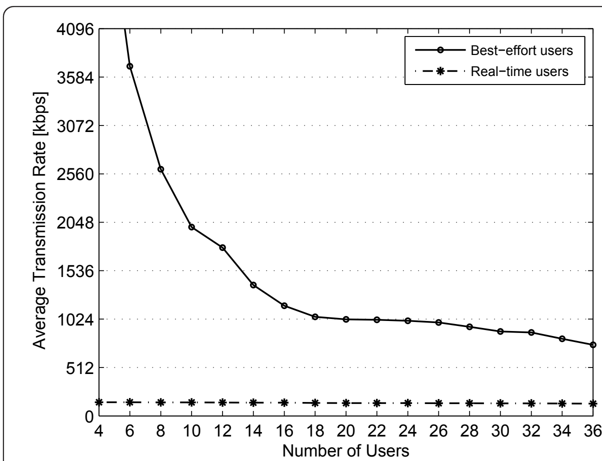


Figure 4 Average rate when increasing the number of users in a network.

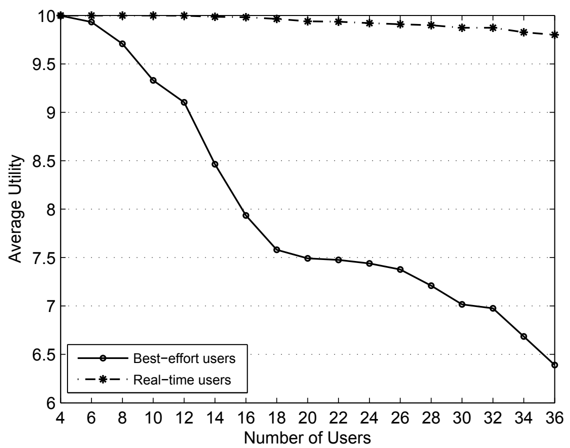


Figure 5 Average utility when increasing the number of users in a network.

based cellular MIMO-OFDMA system. Considering utility functions for individual users in a network, we formulated an optimal resource allocation problem, which simplifies the MIMO resource allocation problem into a SISO resource allocation problem. This problem was shown to be convex. We have presented a low-complexity resource allocation algorithm, which was shown to deliver near-optimum solutions. Furthermore, it was shown that using the proposed algorithm can maintain the performance of real-time users in case of network congestion.

Appendix: Proof of convexity

Without loss of generality, we can rewrite the objective function as

$$f(x,\ y) = -x\log_2\left(1 + \frac{cy}{x}\right), \tag{20}$$

where $c > 0$ is a constant. The gradient of $f\,(x,\ y)$ can calculated as

$$\nabla f(x,\ y) = \begin{bmatrix} \frac{1}{\ln 2}\left[\frac{cy}{x+cy} - \ln(1 + \frac{cy}{x})\right] \\ -\frac{1}{\ln 2}(\frac{cx}{x+cy}) \end{bmatrix}. \tag{21}$$

Similarly, the Hessian of $f\,(x,\ y)$ can be obtained from (21) as

$$\nabla^2 f(x,\ y) = \frac{c^2 y}{\ln 2(x + cy)^2}\begin{bmatrix} \frac{y}{x} & -1 \\ -1 & \frac{x}{y} \end{bmatrix}. \tag{22}$$

Since x and y are also positive, it can be shown that the eigenvalues of $\Delta^2 f\,(x,\ y)$ are non-negative, representing the fact that the Hessian of $f\,(x,\ y)$ is positive semi-definite. Thus, the convexity of the objective function is proven.

Acknowledgements
This work has been funded by the German Federal Ministry of Economics and Technology (BMWi) and carried out within the framework of the research project OPTIFEMTO in cooperation with our partners mimoOn GmbH, Duisburg and Heinrich Hertz Institute, Berlin.

Author details
[1]Department of Communication Systems, University of Duisburg-Essen, Bismarckstr. 81, 47057 Duisburg, Germany [2]Institute of Communications Engineering, University of Rostock, Richard-Wagner-Str. 31, 18119 Rostock, Germany

Competing interests
The authors declare that they have no competing interests.

Received: 16 June 2011 Accepted: 18 August 2011
Published: 18 August 2011

References

1. G Song, Y Li, Utility based resource allocation and scheduling in OFDM-based wireless broadband networks. IEEE Commun Mag. **43**, 127–134 (2005)
2. T Akbudak, M Simsek, B Zhao, A Czylwik, Symmetric capacity of multi-user MIMO downlink under per-base station power constraints. in *Proceedings of the ITG Workshop on Smart Antennas (WSA) 2011*, (Aachen, Germany, Feb 2011)
3. S Shenker, Fundamental design issues for the future internet. IEEE J Sel Areas Commun. **13**, 1176–1188 (1995)
4. CU Saraydar, NB Manam, DJ Goodman, Pricing and power control in a multicell wireless data network. IEEE J Sel Areas Commun. **19**, 1883–1892 (2001). doi:10.1109/49.957304
5. M Xiao, NB Shroff, EKP Chong, A utility-based power-control scheme in wireless cellular systems. IEEE/ACM Trans Netw. **11**, 210–221 (2003). doi:10.1109/TNET.2003.810314
6. P Liu, ML Honig, S Jordan, Forward-link cdma resource allocation based on pricing, in *Proceedings of IEEE Wireless Communications and Networking Conference*. **3**, 1410–1414 (Sept 2000)
7. X Xiao, Z Hu, G Hu, L Li, Adaptive subcarrier allocation for increasing the capacity of multiuser spatial multiplexing based OFDM systems, in *Proceedings of IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications* (Berlin, Germany, Sept 2005)
8. Z Shen, JG Andrews, BL Evans, Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints. IEEE Trans Wirel Commun. **4**, 2726–2737 (2005)
9. M Anas, K Kim, S Shin, K Kim, QoS aware power allocation for combined guaranteed performance and best effort users in OFDMA systems, in *Proceedings of IEEE International Symposium on Intelligent Signal Processing and Communication Systems* (Nov 2004)
10. M Katoozian, K Navaie, H Yanikomeroglu, Utility-based adaptive radio resource allocation in OFDM wireless networks with traffic prioritization. IEEE Trans Wirel Commun. **8**, 66–71 (2009)
11. B Sadiq, R Madan, A Sampath, Downlink scheduling for multiclass traffic in LTE. EURASIP J Wirel Commun Netw. **2009**, 9–9 (2009)
12. R Agrawal, R Berry, J Huang, V Subramanian, Optimal scheduling for OFDMA systems, in *Proceedings of 40th Annual Asilomar Conference on Signals, Systems, and Computers*. (Pacific Grove, CA, USA, Nov 2006)
13. J Huang, VG Subramanian, R Agrawal, RA Berry, Downlink scheduling and resource allocation for OFDM systems. IEEE Trans Wirel Commun. **8**, 288–296 (2009)
14. G Song, Y Li, Cross-layer optimization for OFDM wireless networks-Part II: algorithm development. IEEE Trans Wirel Commun. **4**, 625–634 (2005)
15. G Song, Y Li, Cross-layer optimization for OFDM wireless networks-Part I: theoretical framework. IEEE Trans Wirel Commun. **4**, 614–624 (2005)
16. X Qiu, K Chawla, On the performance of adaptive modulation in cellular systems. IEEE Trans Commun. **47**, 884–895 (1999). doi:10.1109/26.771345
17. CY Wong, RS Cheng, KB Lataief, RD Murch, Multiuser OFDM with adaptive subcarrier, bit, and power allocation. IEEE J Sel Areas Commun. **17**, 1747–1758 (1999). doi:10.1109/49.793310
18. S Boyd, L Vandenberghe, *Convex Optimization*, (Cambridge University Press, Cambridge, 2004)

19. G Miao, N Himayat, Low complexity utility based resource allocation for 802.16 OFDMA systems. in *Proceedings of WCNC 2008*, (Las Vegas, NV, 2008)
20. Recommendation ITU-R M.1225: guidelines for evaluation of radio transmission technologies for IMT-2000, Technical Report (1997)

doi:10.1186/1687-1499-2011-67
Cite this article as: Akbudak *et al.*: A cross-layer resource allocation scheme for spatial multiplexing-based MIMO-OFDMA systems. *EURASIP Journal on Wireless Communications and Networking* 2011 **2011**:67.