A QoS-providing Resource Allocation Scheme in Multiuser Multicarrier Systems

Qing Bai and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Technische Universität München, Munich, Germany
Email: {bai.qing, nossek}@nws.ei.tum.de

***Abstract*—The resource allocation problem at the downlink of a wireless multiuser multicarrier system is addressed, where each end user in the cell has two QoS requirements that need to be fulfilled, namely the minimum data throughput and the maximum latency time. Under these QoS constraints and with limited available radio resources (frequency band and transmit power), the capacity of the system in terms of the number of servable users is to be maximized. A cross-layer framework is introduced where subchannel assignment, power allocation, adaptive modulation and coding, and ARQ protocol are jointly taken into account. Due to the complex structure of the optimization a suboptimal solution is proposed, which is shown by numerical experiments to be computationally tractable yet yielding relatively satisfactory system performance.**

I. Introduction

By jointly adapting variables from several layers in a communication system and thus making more efficient use of the radio resources, *cross-layer optimization* has become one of the key approaches to meet the increasing demand for better *quality of service* (QoS) in current and future wireless networks. On the other hand, due to their ability to overcome frequency selective fading and support high data rates, multicarrier systems have drawn much research and industrial attention. In this work, the application of cross-layer optimization technique to a multicarrier system that serves as a downlink transmitter is investigated, aiming at allocating radio resources to the end users in the cell in an efficient manner.

The resource allocation problem in multicarrier systems has long been studied. In the early work of [1], a multiuser OFDM subcarrier, bit, and power allocation algorithm to minimize the total transmit power is proposed. In [2], the Lagrange dual decomposition method is employed to solve both weighted sum rate maximization and weighted sum power minimization problems in OFDMA systems. When cross-layer techniques come into play, such optimizations would involve variables from different layers and can be categorized into two classes, namely the *bottom-up* optimizations, where the amount of resources is given and users' QoS are to be optimized, and the *top-down* optimizations, where the QoS requirements of users are given and the resources needed to provide such QoS are to be optimized. In [3], the maximization of a utility function takes both efficiency and fairness of resource allocation into account. Such a framework falls into the bottom-up category where the utility value serves as a QoS metric.

The cell capacity maximization problem under consideration is a bottom-up optimization, and throughout the paper we refer to the module that solves this optimization as the *resource allocator* (RA). We propose a RA algorithm that gives a suboptimal solution by solving a sequence of top-down optimizations, *i.e.*, iteratively, the minimum amount of resources needed to serve a specific subset of users is computed and compared with the total amount available. The cross-layer framework from [4] is adopted and adjusted to the special features of multicarrier systems.

II. System Model

Consider the downlink scenario of an isolated single-cell multicarrier system with D users each having one data stream to be served.[1] The data streams, or equivalently the users, are characterized by two QoS parameters each, namely *throughput* and *latency* which are defined as:

Definition 1: The throughput ρ_k of user k is the data rate that is available on top of channel coding. Let T be the length of a *transmission time interval* (TTI), *i.e.*, the time slot in which a packet is transmitted, and B_k be the number of information bits that are successfully transmitted during one TTI, then $\rho_k = B_k/T$.

Definition 2: The latency τ_k of a packet from user k is the delay it experiences until received correctly with an outage probability of no more than the predefined value $\pi^{(\mathrm{out})}$. Let $f_k[m]$ be the probability that it takes exactly m TTI's to transmit a packet error-free, then $\tau_k = (M_k - 1)(\mathrm{RTD} + T) + T$ where RTD represents *round trip delay*, and

$$M_k = \min_{M} M \quad \text{s.t.} \quad \sum_{m=1}^{M} f_k[m] \geq 1 - \pi^{(\mathrm{out})}.$$

In the following subsections, the mathematical descriptions of the regarded system components are derived which lay the basis for cross-layer optimization.

A. Channel Model

The downlink broadcast channel is modeled as frequency-selective fading over the total system bandwidth and frequency-flat fading over each *subchannel*, which is consist of N_c adjacent subcarriers. FDMA is employed meaning the

[1] The case where each user has more than one data stream can be easily extended from this work.

assignment of every subchannel is exclusive to one user, and *intercarrier interference* (ICI) is not taken into account. Moreover, we restrict ourselves here to the single-antenna case both at the base station (BS) and at the mobile stations (MS).

Let $H_{k,n}$ and $\sigma_{k,n}^2$ be the channel coefficient and Gaussian noise variance of user k on the nth subchannel, and p_n be the amount of power allocated on subchannel n. When assigned to user k, the *signal-to-noise-ratio* (SNR) on subchannel n can be computed as

$$\gamma_{k,n} = \frac{|H_{k,n}|^2}{\sigma_{k,n}^2} \cdot p_n. \tag{1}$$

We choose the TTI to be of length $T = 2$ ms and assume that the channel is constant during one TTI. Resource allocation is updated on a per TTI basis, and one TTI contains $N_\mathrm{s} = 16$ symbols for data transmission.

B. FEC coding and modulation

We assume that modulation and coding across the subchannels are done independently, and with reference to the WiMAX standard 8 modulation and coding schemes (MCS) are chosen as candidates, which are listed in Table I.

Table I
MODULATION AND CODING SCHEMES (MCS)

Index	Modulation Type	Alphabet Size A	Code Rate R	$R \log_2 A$
1	BPSK	2	1/2	0.5
2	QPSK	4	1/2	1
3	QPSK	4	3/4	1.5
4	16-QAM	16	1/2	2
5	16-QAM	16	3/4	3
6	64-QAM	64	2/3	4
7	64-QAM	64	3/4	4.5
8	64-QAM	64	5/6	5

Since intersymbol interference is not present in the system with the help of cyclic prefix or an equalizer, each subchannel can be modeled as a *discrete memoryless channel* (DMC) over which the *noisy channel coding theorem* [5] can be applied. Let the modulation alphabet and the coding rate on subchannel n be $\mathcal{A}_n = \{a_1, \ldots, a_{A_n}\}$ and R_n respectively. The *cutoff rate* of subchannel n with SNR $\gamma_{k,n}$ can be expressed as

$$\begin{aligned} &R_0(\gamma_{k,n}, A_n) \\ &= \log_2 A_n - \log_2 \left[1 + \frac{2}{A_n} \sum_{m=1}^{A_n - 1} \sum_{l=m+1}^{A_n} e^{-\frac{1}{4}|a_l - a_m|^2 \gamma_{k,n}} \right]. \end{aligned} \tag{2}$$

Note that the cutoff rate is monotonically increasing with SNR when the modulation alphabet is fixed, yet it is not monotone with varying modulation levels when SNR is fixed.

The noisy channel coding theorem states that there always exists a block code with block length l and binary code rate $R_n \log_2 A_n \leq R_0(\gamma_{k,n}, A_n)$ in bits per subchannel use, such that with maximum likelihood decoding the error probability $\tilde{\pi}_{k,n}$ of a code word satisfies

$$\tilde{\pi}_{k,n} \leq 2^{-l(R_0(\gamma_{k,n}, A_n) - R_n \log_2 A_n)}. \tag{3}$$

In order to apply this upper bound on code word error probability to the extensively used turbo decoded convolutional code, quantitative investigations have been done in [4] and an expression for the *equivalent block length* is derived based on link level simulations. The result from [4] shows that the performance of a turbo decoded convolutional code applied to a packet of coded length L_n in a very good approximation equals the performance of a block code with block length

$$n_\mathrm{eq} = \beta \ln L_n, \tag{4}$$

where parameter β is used to adapt this model to the specifics of the employed turbo code, and $L_n = N_\mathrm{s} N_\mathrm{c} \log_2 A_n$. Consequently, the transmission of L_n bits is equivalent to the sequential transmission of L_n / n_eq blocks of length n_eq and has an error rate of

$$\begin{aligned} \pi_{k,n} &= 1 - (1 - \tilde{\pi}_{k,n})^{\frac{L_n}{n_\mathrm{eq}}} \\ &\leq 1 - \left(1 - 2^{-n_\mathrm{eq}(R_0(\gamma_{k,n}, A_n) - R_n \log_2 A_n)}\right)^{\frac{L_n}{n_\mathrm{eq}}}. \end{aligned} \tag{5}$$

C. Protocol

At the MAC layer an *automatic repeat request* (ARQ) protocol is employed. The data sequence transmitted in one TTI on one subchannel, *i.e.*, a *packet*, is used as the retransmission unit. Since with ARQ, the corrupted packets at the receiver are simply discarded, we assume that the packet error probability (PEP) of a retransmitted packet is the same as that of its original transmission, which consequently gives

$$f_{k,n}[m] = \pi_{k,n}^{m-1}(1 - \pi_{k,n}), \quad m \in \mathbb{Z}^+.$$

The number of transmissions needed to keep the outage probability below $\pi^{(\mathrm{out})}$ is obtained as

$$M_{k,n} = \left\lceil \frac{\ln \pi^{(\mathrm{out})}}{\ln \pi_{k,n}} \right\rceil. \tag{6}$$

Now denote the set of subchannels assigned to user k as $\mathcal{S}_k$. The latency τ_k is determined by $M_k = \max_{n \in \mathcal{S}_k} M_{k,n}$ according to Definition 2. Throughput ρ_k on the other hand, equals the sum throughput on every subchannel $n \in \mathcal{S}_k$ and can be computed as

$$\rho_k = \frac{N_\mathrm{s} N_\mathrm{c}}{T} \sum_{n \in \mathcal{S}_k} R_n \log_2 A_n (1 - \pi_{k,n}) \tag{7}$$

The quantities mentioned in this section, their notations, as well as their simulation values are summarized in Table II.

III. CROSS-LAYER OPTIMIZATION

A. Problem Formulation

As introduced, the RA is provided with D data streams, each with QoS parameters $\rho_k^{(\mathrm{rq})}$ and $\tau_k^{(\mathrm{rq})}$. The bottom-up

Table II
SYSTEM PARAMETERS

Total bandwidth		10 MHz
Center frequency	f_c	2.5 GHz
FFT size		1024
Number of data subcarriers		720
Number of subchannels	N	30
Number of subcarriers per subchannel	N_c	$720/30 = 24$
Transmission Time Interval (TTI)	T	2 ms
Number of data symbols per TTI	N_s	16
Round Trip Delay (RTD)	RTD	10 ms
Turbo code dependent parameter	β	32
Outage probability	$\pi^{(\text{out})}$	0.01
Number of users in the cell	D	24

optimization of maximizing system capacity is formulated as

$$\begin{aligned} \max_{\mathcal{K},\{\mathcal{S}_k\},\boldsymbol{p},\boldsymbol{A},\boldsymbol{R}} & |\mathcal{K}| \\ \text{s.t.} \quad \mathbf{1}^{\mathrm{T}}\boldsymbol{p} &\le P_{\text{tot}}, \\ \mathcal{S}_i \cap \mathcal{S}_j &= \emptyset, \quad i,j \in \mathcal{K}, i \ne j \\ \cup_{k\in\mathcal{K}}\mathcal{S}_k &\subseteq \mathcal{N}, \\ (A_n, R_n) &\in \mathcal{M}, \quad n = 1,\dots,N \\ \rho_k &\ge \rho_k^{(\text{rq})}, \ k \in \mathcal{K} \\ \tau_k &\le \tau_k^{(\text{rq})}, \ k \in \mathcal{K}, \end{aligned} \tag{8}$$

where P_{tot} is the total available transmit power, $\mathcal{K}$, $\mathcal{N}$ and $\mathcal{M}$ are the set of users, the set of subchannels, and the set of available MCS, respectively. The top-down optimization of minimizing transmit power with respect to user subset $\mathcal{K}$ is formulated as

$$\begin{aligned} \min_{\{\mathcal{S}_k\},\boldsymbol{p},\boldsymbol{A},\boldsymbol{R}} & \mathbf{1}^{\mathrm{T}}\boldsymbol{p} \\ \text{s.t.} \quad \mathcal{S}_i \cap \mathcal{S}_j &= \emptyset, \quad i,j \in \mathcal{K}, i \ne j \\ \cup_{k\in\mathcal{K}}\mathcal{S}_k &\subseteq \mathcal{N}, \\ (A_n, R_n) &\in \mathcal{M}, \quad n = 1,\dots,N \\ \rho_k &\ge \rho_k^{(\text{rq})}, \ k \in \mathcal{K} \\ \tau_k &\le \tau_k^{(\text{rq})}, \ k \in \mathcal{K}, \end{aligned} \tag{9}$$

and its optimum value is denoted by $P_{\min}$. Note that since $\mathcal{M}$ is a finite set of MCS, (9) can be infeasible.

As the latency requirement bounds the maximum number of transmissions to

$$M_k = \left\lfloor \frac{\tau_k^{(\text{rq})} - T}{\text{RTD} + T} + 1 \right\rfloor, \tag{10}$$

the maximum tolerable PEP is computed as $\pi_k^{(\text{rq})} = \sqrt[M_k]{\pi^{(\text{out})}}$. This means, constraint $\tau_k \le \tau_k^{(\text{rq})}$ is equivalent to $\pi_{k,n} \le \pi_k^{(\text{rq})}$, $n \in \mathcal{S}_k$. As $R_n \log_2 A_n$ is within the range from 0.5 to 5, the minimum number of subchannels required by user k, $N_k^{(\text{l})}$, and the maximum number of subchannels user k can possibly use, $N_k^{(\text{u})}$, are computed as

$$N_k^{(\text{l})} = \left\lfloor \frac{\rho_k^{(\text{rq})} T}{5 \cdot N_s N_c} \right\rfloor + 1, \quad N_k^{(\text{u})} = \left\lceil \frac{\rho_k^{(\text{rq})} T}{0.5 \cdot N_s N_c} \right\rceil. \tag{11}$$

Therefore a sufficient condition for (9) to be feasible is $\sum_{k\in\mathcal{K}} N_k^{(\text{l})} \le N$.

As it is intractable to enumerate all user subsets, strict priority order is imposed on the D users such that the reduction and expansion of $\mathcal{K}$ is simplified. The general RA procedure is summarized in Algorithm 1.

Algorithm 1 Resource Allocation Procedure by RA

Require: Prioritized D data streams with QoS requirements
Ensure: Maximum set of data streams $\mathcal{K}$ that can be served
  Compute $N_k^{(\text{l})}$, $N_k^{(\text{u})}$
  $K \leftarrow \max K'$ s.t. $\sum_{k=1}^{K'} N_k^{(\text{l})} \le N$
  $\mathcal{K} \leftarrow \{1,\dots,K\}$, solve (9) with $\mathcal{K}$
  while $P_{\min} > P_{\text{tot}}$ **do**
    $\mathcal{K} \leftarrow \mathcal{K} \setminus \{K\}$, $K \leftarrow K-1$, solve (9) with $\mathcal{K}$
  end while
  for $K = K+2,\dots,D$ **do**
    if $N_K^{(\text{l})} + \sum_{k\in\mathcal{K}} N_k^{(\text{l})} \le N$ **then**
      Solve (9) with $\mathcal{K} \cup \{K\}$
      if $P_{\min} \le P_{\text{tot}}$ **then**
        $\mathcal{K} \leftarrow \mathcal{K} \cup \{K\}$
      end if
    end if
  end for

B. The Three-step Approach

The non-convex top-down optimization has intrinsically a complicated structure in that the optimization variables are tightly coupled causing a direct decomposition of the original problem impossible. In [6], a non-iterative method is proposed to solve the sum-rate maximization problem in OFDMA systems subject to the constraints on the proportionality on user data rates. Based on this idea we propose a three-step approach to solve (9), *i.e.*, first the subchannel assignment (SA) is determined, then power allocation (PA) is performed, at last SA and PA are adjusted if there are free subchannels left. At each step, some variables are kept fixed while some others are being optimized: in the SA step, the MCS on every subchannel is fixed and different assignments are compared by computing the power needed on each subchannel to achieve the required PEP; in the PA step, the subchannel assignment is fixed and the best combination of MCS on each subset of subchannels is found.

1) Subchannel Assignment (SA): In this step the same MCS is assumed on every subchannel. If fact, whichever specific MCS is chosen will yield the same SA. According to (2), (4), (5), the minimum power required to achieve $\pi_k^{(\text{rq})}$ on a subchannel can be computed by using a binary search on the cutoff rate curve, and the results are recorded in matrix $\boldsymbol{P} \in \mathbb{R}_+^{|\mathcal{K}|\times N}$. The SA problem is formulated as

$$\begin{aligned} \min_{\{\mathcal{S}_k\}} & \sum_{k\in\mathcal{K}} \sum_{n\in\mathcal{S}_k} p_{k,n} \\ \text{s.t.} \quad \mathcal{S}_i \cap \mathcal{S}_j &= \emptyset, \quad i,j \in \mathcal{K}, i \ne j \\ \cup_{k\in\mathcal{K}}\mathcal{S}_k &\subseteq \mathcal{N}, \\ |\mathcal{S}_k| &\ge N_k^{(\text{l})}, \ k \in \mathcal{K} \\ |\mathcal{S}_k| &\le N_k^{(\text{u})}, \ k \in \mathcal{K}, \end{aligned} \tag{12}$$

i.e., from each column of $\boldsymbol{P}$ one entry is picked such that the kth row has between $N_k^{(\mathrm{l})}$ and $N_k^{(\mathrm{u})}$ picked entries, and the sum of all picked entries is minimized.

Intuitively the minimum entry from every column in $\boldsymbol{P}$ is picked up. If $N_k^{(\mathrm{l})} \leq |\mathcal{S}_k| \leq N_k^{(\mathrm{u})}, k \in \mathcal{K}$ is not fulfilled, then a set of unsatisfied users $\mathcal{K}_\mathrm{u}$ with $|\mathcal{S}_k| < N_k^{(\mathrm{l})}$, and a set of oversatisfied users $\mathcal{K}_\mathrm{o}$ with $|\mathcal{S}_k| > N_k^{(\mathrm{l})}$ are obtained. We deprive from the oversatisfied users the subchannels with the least advantage assigning to them, until $\forall k \in \mathcal{K}_\mathrm{o}, |\mathcal{S}_k| = N_k^{(\mathrm{l})}$. All the deprived subchannels form a set of extra subchannels $\mathcal{N}_\mathrm{e}$. The same assigning procedure is then repeated on $\mathcal{K}_\mathrm{u}$ and $\mathcal{N}_\mathrm{e}$. The recursion stops when there are no more unsatisfied users. The SA algorithm is summarized in Algorithm 2. It is a greedy algorithm in the sense that during the assignment, the unsatisfied and exactly-satisfied users never give up the subchannels already assigned to them.

Algorithm 2 Subchannel assignment

Require: $\boldsymbol{P} = (p_{k,n})$, $N_k^{(\mathrm{u})}$ and $N_k^{(\mathrm{l})}$
  $\mathcal{S}_k \leftarrow \emptyset, k \in \mathcal{K}$
  for $n = 1, \ldots, N$ **do**
    $k \leftarrow \mathrm{argmin}_k\, p_{k,n}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \cup \{n\}$
  end for
  $\mathcal{K}_\mathrm{u} \leftarrow \{k : |\mathcal{S}_k| < N_k^{(\mathrm{l})}\}$, $\mathcal{N}_\mathrm{e} \leftarrow \emptyset$
  while $\mathcal{K}_\mathrm{u} \neq \emptyset$ **do**
    for each $k \in \{k : |\mathcal{S}_k| > N_k^{(\mathrm{l})}\}$ **do**
      while $|\mathcal{S}_k| > N_k^{(\mathrm{l})}$ **do**
        $n \leftarrow \mathrm{argmin}_{n \in \mathcal{S}_k} \left(\min_{k' \in \mathcal{K}_\mathrm{u}} (p_{k',n} - p_{k,n})\right)$
        $\mathcal{N}_\mathrm{e} \leftarrow \mathcal{N}_\mathrm{e} \cup \{n\}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \setminus \{n\}$
      end while
    end for
    for each $n \in \mathcal{N}_\mathrm{e}$ **do**
      $k = \mathrm{argmin}_{k \in \mathcal{K}_\mathrm{u}} p_{k,n}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \cup \{n\}$
    end for
    $\mathcal{K}_\mathrm{u} \leftarrow \{k : |\mathcal{S}_k| < N_k^{(\mathrm{l})}\}$, $\mathcal{N}_\mathrm{e} \leftarrow \emptyset$
  end while
  $\mathcal{N}_\mathrm{e} \leftarrow \{n :$ extra subchannels assigned to the last $k \in \mathcal{K}_\mathrm{u}\}$
  $\mathcal{K}_\mathrm{a} \leftarrow \{k : |\mathcal{S}_k| < N_k^{(\mathrm{u})}\}$
  for each $n \in \mathcal{N}_\mathrm{e}$ **do**
    $k = \mathrm{argmin}_{k \in \mathcal{K}_\mathrm{a}} p_{k,n}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \cup \{n\}$, update $\mathcal{K}_\mathrm{a}$
  end for

2) Power Allocation (PA): With the SA result as input, power allocation is no longer coupled among the users. The top-down optimization (9) can therefore be decomposed into $|\mathcal{K}|$ independent optimizations as

$$\min \sum_{n \in \mathcal{S}_k} p_n \quad \text{s.t.} \quad \rho_k \geq \rho_k^{(\mathrm{rq})}, \tau_k \leq \tau_k^{(\mathrm{rq})}. \qquad (13)$$

Firstly we look for all *efficient* MCS combinations on $n \in \mathcal{S}_k$ such that

$$\sum_{n \in \mathcal{S}_k} R_n \log_2 A_n > \frac{\rho_k^{(\mathrm{rq})} T}{N_\mathrm{s} N_\mathrm{c}} \qquad (14)$$

is fulfilled. A MCS combination M_c is said to be *efficient* if $\nexists M_\mathrm{c}' \neq M_\mathrm{c}$, such that $\forall n \in \mathcal{S}_k, (R_n \log_2 A_n)_{M_\mathrm{c}'} \leq (R_n \log_2 A_n)_{M_\mathrm{c}}$ but still fulfills (14). Note that to achieve the same PEP, the most power-saving MCS combination is always an efficient one.

For each MCS combination, the optimal power allocation is the solution to problem

$$\begin{aligned} &\min_{\{\pi_{k,n} : n \in \mathcal{S}_k\}} \sum_{n \in \mathcal{S}_k} p_n \\ &\text{s.t.} \quad 0 \leq \pi_{k,n} \leq \pi_k^{(\mathrm{rq})}, \quad n \in \mathcal{S}_k, \\ &\qquad \sum_{n \in \mathcal{S}_k} R_n \log_2 A_n (1 - \pi_{k,n}) \geq \frac{\rho_k^{(\mathrm{rq})} T}{N_\mathrm{s} N_\mathrm{c}}. \end{aligned} \qquad (15)$$

Since the mapping from PEP to transmit power is monotonic, the second constraint in (15) must be active at optimality. What is more, if the MCS combination satisfies

$$\sum_{n \in \mathcal{S}_k} R_n \log_2 A_n > \frac{\rho_k^{(\mathrm{rq})} T}{(1 - \pi_k^{(\mathrm{rq})}) N_\mathrm{s} N_\mathrm{c}},$$

then $\pi_{k,n} = \pi_k^{(\mathrm{rq})}, n \in \mathcal{S}_k$ is optimum. However, other properties of the mapping which provide more insight on solving (15), such as differentiability and convexity, are hard to prove. As the expense to further reduce the PEP when it is already small can be very high, central points in the polyhedron defined by the constraints of (15) could serve as good approximations to the optimal solution. In the implementation we choose the *analytic center* [7] of the polyhedron in particular.

3) Adjustment: The outcome of PA might indicate zero MCS on some subchannels, which means these subchannels are released from occupation, and can be assigned again to other users. As higher MCS are much more power consuming than lower MCS, we find the subchannels using the relatively highest MCS as well as their possessors, and compare each alternative of assigning the empty subchannels to these users. The adjustment phase is in fact an amendment to SA.

IV. Simulation Results

24 users uniformly located in a cell of radius 2 km are assumed, each having one data stream to be served. The data streams are characterized by the typical QoS parameters from 3 classes of traffic, namely voice traffic, video streaming, and other interactive streaming applications such as gaming. The QoS requirements of the data streams used in the simulations are listed in Table III according to their priority order, where the unit of ρ is kbit/s and the unit of τ is ms.

Table III
QoS requirements of 20 users for simulation

User	ρ_k	τ_k	User	ρ_k	τ_k
1-6	128	20	13-18	384	40
7-12	1600	40	19-24	200	50

The wireless channel is modeled as a frequency-selective fading channel consisting of six independent Rayleigh multipaths with an exponentially decaying power profile. The delay

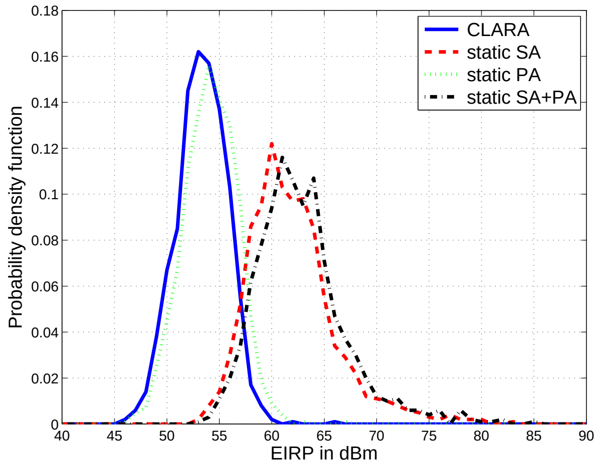

Figure 1. Probability density of the minimum power to serve all 20 users

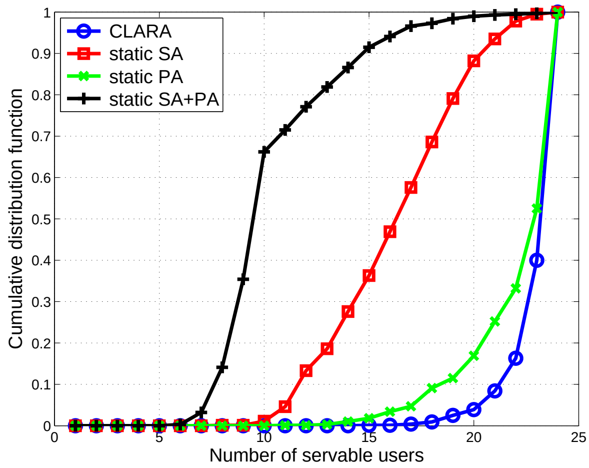

Figure 2. Cumulative distribution of the number of servable users

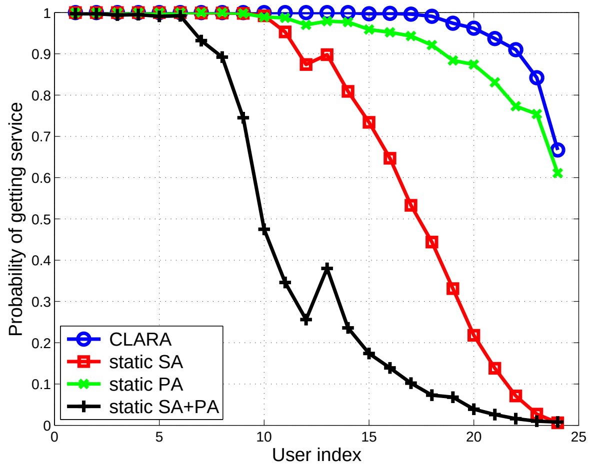

Figure 3. Probability of each user to get served

spreads are uniformly distributed within 1 μs, resulting in a rms delay spread of about 0.3 μs which is consistent with the assumed channel coherence bandwidth. The path loss in dB is computed as $PL(d) = 140.6 + 35.0 \log_{10} d$ following the COST-Hata model, where d is the distance between MS and BS in km. We assume a receiver noise level of -174 dBm/Hz and a total transmit power of $P_{\text{tot}} = 10$ W, which plus antenna gain and minus BS internal loss results in an *effective isotropic radiated power* (EIRP) of 54 dBm.

For the purpose of comparison, the allocation schemes with static SA, where each user is assigned with adjacent fixed subchannels, with static PA where the same MCS and PEP are applied on every subchannel the user occupies, and with static SA+PA which combines the former two are also simulated. The program is executed with 1000 independent channel realizations. In Fig. 1-3 it is clearly shown that the allocation scheme proposed in this paper, under the name *CLARA*, outperforms the other three static schemes in reducing power consumption and providing service to more users, whereas static SA+PA performs the worst as expected. The importance of an efficient dynamic SA scheme is also stressed by the larger gain obtained going from static SA to CLARA.

From Fig.3 where the probabilities of each user getting served under the four allocation schemes are shown, it can be seen that fairness is a big issue for the proposed scheme due to the strict priority imposed on users. This situation can be improved by adding another scheduling unit on top of RA, which by delicate algorithms can provide the RA with consecutive user lists that maintain fairness to a certain extent.

V. Conclusion and Outlook

In this paper a cross-layer assisted resource allocation scheme for multicarrier systems is proposed. Based on a three-step approach the transmit power minimization subject to target QoS values is solved, which is employed by an outer capacity maximization algorithm to validate the ability of the system to serve a certain subset of QoS demanding users. Numerical experiments have demonstrated the effectiveness of the proposed scheme, which can serve as a good starting module for a more sophisticated cross-layer framework.

References

[1] C. Y. Wong, R. S. Cheng, K. B. Letaief and R. D. Murch, *Multiuser OFDM with adaptive subcarrier, bit and power allocation*, IEEE Journal on Select. Areas of Commu., pp. 1747-1758, vol. 17, 1999.

[2] K. Seong, M. Mohseni and J. M. Cioffi, *Optimal Resource Allocation for OFDMA Downlink Systems*, in IEEE Int. Symposium on Information Theory, July 2006, pp. 1394-1398.

[3] G. Song and Y. Li, *Cross-layer optimization for OFDM wireless networks-part I: theoretical framework*, in IEEE Trans. on Wireless Communications, March 2005.

[4] B. Zerlin, M. T. Ivrlač, W. Utschick, J. A. Nossek, I. Viering and A. Klein, *Joint optimization of radio parameters in HSDPA*, in IEEE 61st Vehicular Technology Conference VTC 2005-Spring, vol. 1, pp. 295-299.

[5] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Son, 1968.

[6] I. C. Wong, Z. Shen, B. L. Evans and J. G. Andrews, *A low complexity algorithm for proportional resource allocation in OFDMA systems*, in IEEE Workshop on Signal Processing Systems, Oct. 2004, pp. 1-6.

[7] R. M. Freund, *Notes for Class: Analytic Center, Newtons Method, and Web-Based ACA*.