Allocation algorithms for PRIMO system

Daniele Messina¹, Luca Scalia¹, Ilenia Tinnirello¹, Simone Merlin§, Andrea Zanella§, Marco Moretti♯

¹ Università di Palermo, DIE, Viale delle Scienze, 90128, Palermo, Italy
§ Università di Padova, DIE, Via Gradenigo, 35100, Padova
♯ Università di Pisa, DIET, Via Caruso Pisa

Abstract-

This paper proposes an architecture for the LL, MAC and PHY layers for a multicellular wireless system, which allows for a cross layer interaction. The PHY layer is assumed to implement an OFDM modulation, which is used at the MAC level to realize an hybrid FDMA/TDMA medium access. The complete reuse of the available resources among neighboring cells is the most challenging assumption, which brings the need for a smart resource allocation. Following a cross layer perspective, the radio resources are efficiently allocated to the different flows using information coming from the PHY layer. Each base station performs the allocation without explicit coordination with other cells, thus resulting in a distributed, dynamic allocation. Three different algorithms are presented and discussed. They are also compared with static FDMA, TDMA and random allocations.

I. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) is the most widespread and promising solution for multiple access and signaling in today’s wireless (broadband) networks [1]. Its deployments include WLAN physical layer implementations such as IEEE 802.11a/g, ETSI HIPERLAN/2 and the IEEE 802.16 standard for broadband wireless access in metropolitan area networks and the Digital Audio/Video Broadcasting (DAB/DVB) standards. The OFDM technology is based on the principle of multi-carrier transmission, originally appeared in the design of high speed digital subscriber line (HDSL) [2]. The OFDM transmission method results to be a really effective platform in multi-path environments with frequency selective fading. A significant advantage of the OFDM technology is the possibility of allocating power and rate optimally among the subcarriers, by using adaptive modulation according to instantaneous subcarrier quality, thus maintaining acceptable BER per subcarrier [3]. Moreover, in the multi-user scenario, it is possible to assign subcarriers to the less interferenced user, owing to the channel diversity among users placed in different locations (multiuser diversity gain). Optimized solutions for the multi-user case are considered in [5], [6] (centralized) and [7] (distributed).

Channel allocation together with power adaptation techniques for throughput increase are presented in [8], in the case of a generic multiple access schemes with orthogonal channels. The work in [16] provide a subcarrier allocation and power adaptation with the additional aim of providing fairness among allocated users. A joint optimization of subcarrier allocation and power adaptation is also presented in [17, 18].

The advantage of using adaptive resource allocations is evident if compared to fixed allocations, which do not account for the time-varying channel conditions perceived by the users. However, it is not trivial to find an optimized solution able to coniugate the goals of different protocol layers. The optimization problem can be intended as a i) joint optimization problem, whose complexity arises as a function of the parameters to be optimized for each layer, thus requiring a feasibility study according to the variables selected for the optimization [9, 10, 11], or ii) as a strictly layered optimization, where state information of other layers is used in order to improve specific network performance at a given layer [12].

In this work we propose a cross-layer OFDMA-based architecture mostly oriented towards the latter approach. This architecture in fact employs channel and scheduler feedbacks in order to choose the best sub-carrier allocation allocation pattern for each user, with the final objective of reducing the overall system interference while maximizing the overall system throughput.

II. SYSTEM ARCHITECTURE

A multicellular system with frequency reuse equal to one is considered [13]. Base stations (BS) are placed on an regular grid and mobiles (MS) are uniformly scattered and connected to the base station with best (average) channel.

The architecture for the radio access on each cell is shown in Fig. 1. The scheduler block takes as inputs the traffic and QoS requirements for the communications among BS and MSs while the outputs are the requests for the allocation of MAC resources (the details of scheduling procedures are out of the scope of this paper).

The allocation block, which acts as a cross layer adaptation entity, exploits information coming from upper and lower layers to map the scheduler requests on the PHY layer. In the following, the PHY, MAC and allocation layers are described in more details and their interaction are highlighted.
Each scheduler request uler requests setting also the power and the bitload for each PBU.

scenario.

dynamic in the adjacent cells may change the current interference users face during the upcoming frame, since traffic and channel

ter possesses information about channel and interference state per-

The physical layer (PHY) is assumed to use an OFDM mod-
ulation with $N_s$ subcarriers. Each subcarrier can be loaded with an independent amount of bits (by using different modulation and coding schemes) and power. The medium access is organized in a hybrid FDMA-TDMA. The MAC frame is composed by $N_f$ sub-
sequent OFDM symbols of duration $T_f$ each, resulting in a frame duration of $T_f$ s. Since different subcarriers can be allocated to different users, the addressable space is composed of $N_s \times N_f$ physical bearers. In order to reduce the addressing overhead, bear-
ers which are close in time and frequency (i.e. which fall within the channel coherence in frequency and time) are grouped together. In the following these groups are indicated as physical base units (PBU). Thus, the addressable resources are reduced to a matrix of $M_f \times M_s$ PBUs. Physical carriers belonging to the same group are equally bit and power loaded. Bitload and power can be set independently on each PBU.

We assume that at the beginning of each frame the transmit-
er possesses information about channel and interference state ac-

The allocation block takes as input such information, and at
the beginning of each frame, assigns a subset of PBUs to the sched-
uler requests setting also the power and the bitload for each PBU. Each scheduler request $k$ is represented by a minimum rate $r_{k,min}$ and maximum rate $r_{k,max}$ to be allocated. The allocated resources can be used by the scheduler in order to transmit one or more pack-
sets and a single request may require several successive frames to
be completed. Moreover, we assume that the aggregated rate al-
duced rate $r_{tot}$ is set according to policies of admission and load control out of the

Each BS autonomously performs the allocation and no explicit
communication takes place among cells.

III. ALLOCATION ALGORITHMS

Many algorithms can be defined to solve the resource allocation problem choosing different objective function to be optimized and adopting different techniques to solve the optimization problem. In this section, three different allocation algorithms are presented.

Let $c_{k,s,t}$ denote the Shannon capacity associated to the $(s, t)$ $(1 \leq s \leq M_s, 1 \leq t \leq M_t)$ PBU when the $k$-th user (i.e., the user that has generated the $k$-th rate-request, $1 \leq k \leq K$) transmits with power $p_{k,s,t}$. Considering the transmission from a BS to a MS, the capacity $c_{k,s,t}$ is

$$c_{k,s,t} = F_{PHU} \log_2 \left(1 + \frac{p_{k,s,t} G_{k,s,t}}{p_n + p I_{k,s,t}}\right)$$

where $F_{PHU}$ is a scalar factor depending on the bandwidth and the time duration of a PBU, $G_{k,s,t}$ is the channel gain of user $k$ on the $(s, t)$ PBU, $p_n$ is the noise power and $p I_{k,s,t}$ is the power of interference. Let $I_{s,t}$ be the set containing the BSs interfering over the $(s, t)$ PBU. For each interferer $i \in I_{s,t}$ let $p_{i,s,t}$ be the value of the transmitted power and $G_{i,k,s,t}$ the channel gain between the inter-
ferer and the receiving MS; then the value of the total interfering power is

$$p I_{k,s,t} = \sum_{i \in I_{s,t}} p_{i,s,t} \cdot G_{i,k,s,t}.$$  \hspace{1cm} (2)

Let $b_{k,s,t}$ be the bitload on the PBU $(s, t)$ for the request $k$.

Since $c_{k,s,t}$ is the theoretic limit for the number of information bits that can be transmitted on the PBU $(s, t)$, it is:

$$b_{k,s,t} \leq \alpha c_{k,s,t},$$  \hspace{1cm} (3)

where $\alpha < 1$ accounts for limits of actual implementation. Let $\chi_{k,s,t}$ be the indicating function for the allocation of PBU $(s, t)$ to request $k$, i.e.

$$\chi_{k,s,t} = \begin{cases} 1 & \text{if the (s, t) PBU is assigned to user k,} \\ 0 & \text{otherwise.} \end{cases}$$  \hspace{1cm} (4)

Thus, the aforementioned rate constraints can be formalized as:

$$r_{k,min} \leq \frac{1}{T_f} \sum_{s=1}^{M_s} \sum_{t=1}^{M_t} b_{k,s,t} \chi_{k,s,t} \leq r_{k,max} \ .$$  \hspace{1cm} (5)

i.e., the bit rate allocated to each request has to be within the imposed limits and

$$\frac{1}{T_f} \sum_{k=1}^{K} \sum_{s=1}^{M_s} \sum_{t=1}^{M_t} b_{k,s,t} \chi_{k,s,t} = r_{tot} \ .$$  \hspace{1cm} (6)

i.e., the total aggregated allocated rate has to equal the one re-
quested by the scheduler.

A further constraint on the allocation is that, inside each cell, each PBU can be allocated to a single user only:

$$\sum_{k=1}^{K} \chi_{k,s,t} \leq 1 \ \forall (s, t) .$$  \hspace{1cm} (7)
Thus the problem of allocation can be formulated as the problem of jointly finding the optimal set of values of \( \chi_{k,s,t} \) (channel allocation), \( b_{k,s,t} \) (bit loading) and \( p_{k,s,t} \) (power loading) that enforces the constraints in (5)-(7) and optimizes a proper objective function (usually transmitted power). The complexity of this problem is very large and thus we propose three practical algorithms that solve a reduced and linearized version of it.

### A. Power minimization by a greedy heuristic

The allocation problem can be seen as the selection of a certain subset in a given space of events. An event is the assignment of a PBU to a given request, with a certain bitload. We call these events Transmission Hypothesis (TxHp). Thus, a TxHp is determined once a 4-tuple of values is given. The elements of a 4-tuple are:
- \( k \): the request identifier;
- \( s,t \): the coordinates of the PBU;
- \( b \): the amount of bitload.

The whole space of events is composed of all the TxHps corresponding to unused PBUs, i.e. if \( K \) is the number of pending requests and \( N_B \) is the number of available bitloads, the space of events is constituted by \( K \times N_B \) TxHps for each unused PBU.

The allocator determines the requested transmission power corresponding to a TxHp on the base of the given bitload and of the measured interference. Let \( \chi_{k,s,t,b} \) be the indicating function for the selection of the TxHp \((k,s,t,b)\) and \( p_{k,s,t,b} \) the corresponding requested transmission power, computed using (1)-(3). Since only one format can be used on a single PBU, the following constraint applies:

\[
\sum_{b \in B} \chi_{s,t,r,b} = \chi_{s,t,r} \quad \forall (s,t). \quad (8)
\]

The task of the allocator is to find the feasible selection of TxHp that meets the rate and power constraints and minimize the sum of transmission powers. Thus, the allocation problem can be formulated as

\[
\chi = \arg \min_{\chi} \sum_{k=1}^{K} \sum_{s,t} \sum_{b \in B} p_{k,s,t,b} \chi_{k,s,t,b} \quad \text{subject to} \quad \sum_{b \in B} \chi_{k,s,t,b} \leq 1 \quad (C.1)
\]

\[
r_{k,\text{min}} \leq \frac{1}{T_f} \sum_{s,t} \sum_{b \in B} b \cdot \chi_{k,s,t,b} \leq r_{k,\text{max}} \quad (C.2)
\]

\[
\frac{1}{T_f} \sum_{k=1}^{K} \sum_{s,t} \sum_{b \in B} b \cdot \chi_{k,s,t,b} = r_{\text{tot}} \quad (C.3)
\]

The problem in (9) can be solved with an iterative greedy heuristic technique: at each iteration run the allocator increments the transmitted load by choosing a new TxHp until all the rate constraints are satisfied. At each run, the t-uple selected is the one that requires the least increment of the overall transmitted power.

### B. Power minimization by linear programming

The allocation problem as formulated in (9) has a linear cost function and is subject to a set of linear constraints, so that it can be solved by means of linear programming (LP) solvers that are very fast and widely used [14]. Moreover, in the particular case when there is only one possible transmitting format \( (N_B = 1) \), it is possible to model the radio resource allocation problem as a network flow problem. The advantage of this approach is that it can be solved with the network simplex method (NSM), which is an adaptation for the network flows of the simplex method. The NSM is the most efficient solver for min-cost-max-flow network problems [15] and it outperforms other existing techniques like for example LP relaxation used in [6] to solve the problem in (9). After solving the problem in (9) with \( N_B = 1 \), the transmitted power can be further reduced by removing the condition that the modulation format transmitted on each subcarrier must be exactly one. In facts, the \( N(k) \) PBUs allocated to request \( k \) are as many parallel Gaussian channels and the optimal power distribution can be found formulating the waterfilling algorithm [4] with a rate constraint.

### C. Efficiency maximization

The whole problem of bit loading, power loading and allocation is linearized by fixing \textit{a priori} the transmission power \( p_{k,s,t,b} \), thus fixing also the corresponding bitload \( b_{k,s,t} \) for each PBU and each request. In particular, we assume an equal power loading \( p_{k,s,t} = p' \forall k,s,t \).

Let define the \textit{efficiency} of a PBU allocation as follows

\[
e_{k,s,t} = \frac{b_{k,s,t}}{\sum_{k=1}^{N} b_{k,s,t}}. \quad (10)
\]

This index permits to compare the advantage of allocating the PBU \((s,t)\) to the user \( k \), rather than to any other user.

In addition to the constraints (3), (5), (6) and (7), the optimization problem is completed considering the objective function

\[
\min_{k=1}^{K} \sum_{s=1}^{M_s} \sum_{t=1}^{M_t} \chi_{k,s,t}(1 - e_{k,s,t}), \quad (11)
\]

which aims at obtaining the maximum aggregated efficiency. If the problem has no solution, the optimization problem is repeated using an higher power level \( p'^{i+1} = p' + \delta p \).

Since the formalized problem presents a high computational complexity, a greedy heuristic has been developed.

The algorithm chooses iteratively a request \( k \) and then allocates a PBU to this request. Hence, rate requests and available resource set are refreshed accordingly and the procedure is repeated until the constraints of the allocation problem are satisfied.

At each stage of the algorithm, all users requests are sorted using the metric

\[
\eta_k = \frac{b_{k,av}}{r_{k,rem}}. \quad (9)
\]

Here, \( r_{k,rem} \) is the bitrate the flow \( k \) still need in order to reach the minimum rate request \( r_{k,\text{min}} \) for the current frame and \( b_{k,av} \) is the total bitload that might be allocated to the \( k \)-th request, if all the PBUs still available were assigned to it. The user with maximum \( \eta_k \) is chosen for the allocation of a PBU. The selected request will be assigned the PBU with the highest efficiency \( e_{k,s,t} \) among those not yet allocated.

### IV. Simulation results

We evaluated the performance behaviour of our cross-layer radio resource allocation algorithms through a C++ based simulation...
tool. Our simulator creates a cellular network scenario, with 9 BSs distributed over a toroidal surface (this surface does not have edges, so that border effects are not present). The BSs are placed in a regular geometry with a minimum distance between two BSs of 1000 meters. The MSs are uniformly distributed and are associated each with the BS from which the best channel is sensed.

The channels are time varying with a Doppler frequency which corresponds to an average speed of the MSs equal to 1 m/s and a delay spread of $C_s = 10\,\text{ms}$. We have considered a downlink traffic scenario in which the MSs are the receivers while the interferers are the BSs.

Each MS has a single downlink data flow with loose delay requirements characterized by a packet length of 1500 bytes and variable interarrival rate.

Since robust coding techniques are supposed to be employed at the PHY layer, we evaluate the probability of packet error according to the measured values of signal and interference strength on all the PBUs used to carry a packet.

Each simulation is made of 10 batches which run for 300 frame times. In each batch the distribution of the MSs is renewed. The first 100 frames of a batch are assumed as transient and do not contribute to the outcomes. The symbol duration is $T_s = 16\,\mu\text{s}$.

Detailed settings for the considered scenarios are provided in Tab.I.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>BASE</th>
<th>TDMA</th>
<th>PDMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_s$</td>
<td>256</td>
<td>256</td>
<td>256</td>
</tr>
<tr>
<td>$M_t$</td>
<td>32</td>
<td>32</td>
<td>256</td>
</tr>
<tr>
<td>$N_c$</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>$M_c$</td>
<td>16</td>
<td>128</td>
<td>16</td>
</tr>
</tbody>
</table>

| $B$ | \{512, 1024, 1536\} | \{4096, 8192, 12388\} |

It is provided a comparison between the allocation algorithms A, B, C and a system in which the allocation is done in a random manner, i.e. without considering this information in the decision of assigning a resource to a given user, but only in the determination of the transmission power. Fig. 2 shows the results of this comparison. The average offered traffic inside a single cell is in the $x$ axis, the throughput, determined by the packets delivered without error, is on the $y$ axis. All systems definitely outperform the random allocator with a similar behavior either with 108 and 216 users in the whole network (12 and 24 users per cell respectively). Alg. A resulted to perform better when the system is not saturated, while in conditions of saturation Alg. C reaches higher values of throughput. The curves related to the random allocator start with a small increment and gradually degrade as the offered load increases.

The value of throughput for the cross-layer allocators is on the average under a half of the offered load. However this does not mean that the rest of the packets suffer transmission errors: in fact the limit applied to the transmission power is an implicit admission control mechanism and most of the lost traffic is blocked at the scheduler/allocator interface and it is not actually transmitted on the wireless medium. Fig. 3 the error free delivered traffic and the amount of traffic accepted by the allocator and so transmitted over the radio resources are compared. The ratio between error free packets and transmitted packets results regularly larger than 0.85.

In Fig. 4 the average value of transmitted power at the BS is plotted against the offered load. Alg. B is the one that on average requires the least power. The first part of the curves shows how the cross-layer systems self-regulate the transmission power with the growth of the transmitted traffic so that the inter-cell interference is reduced. With all algorithms, a smaller power is needed in the case of 216 users, due to the effect of the user diversity. In both cases the transmitted power drastically decreases with the highest values of accepted load. This happens because when the network is near to saturation only the users with the best channels are served and the transmission power needed at the BS is low.

In order to point out the different behavior of the simulated algorithms, in Fig. 5 it is shown the average number of PBUs allocated to a single request. As can be seen the three allocation mechanisms end up with quite different choices on the assignment of PBUs to the requests. Alg. A tends to assign more PBUs and to use them with lower power, Alg. B shows the opposite behaviour, Alg. C stays in the middle.

We also considered two scenarios in which an OFDMA and a TDMA access are simulated. In these cases the PBUs have been...
aggregated. For example in the case of TDMA all the PUs with the same \( x \)-coordinate are joint and constitute a single PBU which can carry 4096, 8192 and 12388 bits. We compared the mixed OFDMA/TDMA access with equivalent OFDMA and TDMA systems. We used the cross-layer allocation of section A in all the three systems. Fig. 6 shows the average throughput achieved in a single cell versus the offered load. The mixed OFDMA/TDMA system performs better especially for medium loads. Even if the amount of bandwidth used for signalling, which is lower in the OFDMA and TDMA cases, should be taken into account this result is promising.

V. CONCLUSIONS

In this paper an architecture for the PHY, MAC, and LL layer of a multicellular system with a complete reuse of the available radio resources has been presented. The multiple access technique is a hybrid OFDMA-TDMA based on multicarrier modulation. The requirements on the spectral efficiency of such a system brings to the need of an efficient cross-layer MAC-PHY allocation. The allocation problem has been defined to guarantee the QoS requirements of the various users while optimizing the resource usage. We propose here three different resource assignment algorithms aimed to solve the allocation problem. Simulation results for the performance of two of the proposed algorithm has been shown. Further work is in progress to better test the system and to improve the allocation algorithms under different load scenario.

REFERENCES


