Adaptive Scheduling Algorithms for Multimedia Traffic in Wireless OFDMA Systems

Marco Cecchi, Romano Fantacci, Dania Marabissi, and Daniele Tarchi
Department of Electronics and Telecommunications
University of Florence
Firenze, Italy
Email: marcocechi1982@hotmail.com, romano.fantacci@unifi.it, marabissi@lart.det.unifi.it, daniele.tarchi@unifi.it

Abstract—This paper deals with the problem of finding an optimal subcarrier allocation strategy for uplink and downlink communications in an OFDMA metropolitan wireless system. In particular a two steps approach is considered: at the first step the scheduler establishes the amount of resources to assign to users on the basis of their Quality of Service (QoS) constraints, while, at the second step, channel conditions are used to allocate subcarriers to users. A simple strategy for choosing the appropriate Modulation and Coding Scheme according to the channel conditions of the assigned subcarriers is also proposed.

I. INTRODUCTION

The consumer interest in multimedia applications and, hence, the increasing demand of high data rate services leads to a grow for research and development in wireless communications. Wireless systems have the capacity to address broad geographic areas without the costly infrastructure required to deploy cabled links. In particular IEEE 802.16 technology, supported by the WiMAX commercial consortium, is presently an interesting approach to implement an efficient global telecommunication system. Many applications have been foreseen for IEEE 802.16 based networks [1], [2]. These includes fixed or last-mile wireless access, backhauling, mobile cellular networks, etc.

Orthogonal Frequency Division Multiplexing (OFDM) is used in IEEE 802.16 networks to lower the adverse effects of frequency-selective multi-path fading by transmitting signals over a number of flat-faded narrow-band channels. However, to fully exploit the advantages of OFDM in wireless systems, dynamic allocation techniques need to be devised in order to efficiently use resources such as bandwidth, power as well as modulation schemes to increase the spectral efficiency.

Recently, the scheduling issue of multimedia traffic in wireless network has become a hot topic in research community. As a consequence, several results have been presented in the literature [3]–[9]. In particular, in [3] the authors propose an utility based function for resource allocation and scheduling for downlink traffic in an OFDM based communication by exploiting wireless channel status jointly with packet queue information. In [4] the principles of opportunistic scheduling in resource-sharing wireless communication are presented by exploiting the time varying conditions of the physical channel.

In [5], [6], the resource allocation problem in OFDMA (Orthogonal Frequency Division Multiple Access) is discussed and suitable subcarrier allocation schemes are proposed. In particular, being shown in [6] that the proposed method outperforms different alternatives, hence, we will refer in what follows to the performance results presented in [6] to highlight the advantages of the proposed approaches. Finally, in [7]–[9], the resource allocation problem in OFDMA networks is considered by taking into account also the QoS constraints of the users.

In this paper, the resource allocation problem has been analyzed by focusing the attention on the subcarriers allocation jointly with the selection of the most suitable modulation and channel coding scheme. The features offered by the IEEE 802.16e has been exploited by jointly considering the wireless channel behavior and the upper layers queue condition. The problem of finding an optimal subcarrier allocation strategy for uplink (UL) and downlink (DL) communications able to respect the traffic QoS constraints has been divided in two steps. At the first step the total amount of resources to be scheduled for each user are selected by assuming the most suitable modulation/coding scheme for the actual channel conditions. At the second step, the scheduler tries to perform an optimal subcarriers allocation by taking into account the channel propagation conditions that are different for different users due to the frequency selectivity of the channel.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The system under consideration, as specified by the IEEE 802.16e standard, is an OFDM system with Frequency Division Multiple Access (FDMA) where perfect channel state knowledge is assumed [3]–[9], i.e., the channel gain of each user on each subcarrier due to multipath fading is considered as perfectly known at the Base Station (BS) [3]–[9]. The resource allocation is based on the slot concept, i.e., on a minimum amount of allocable resources to each user. A slot is formed by a set of bins (a bin is a set of 9 subcarriers adjacent in frequency with 8 data and 1 pilot subcarriers) of the type $N_{bin} \times M = 6$, where $N_{bin}$ is the number of contiguous bins and $M$ is the number of contiguous symbols [2]. Thus,
A. AMC techniques

The term Adaptive Modulation and Coding (AMC) denotes the possibility of choosing the most suitable modulation and channel coding scheme according to the propagation conditions of the radio link (channel state) known at the transmitting end.

For our purposes we have considered that the channel quality degradation is mainly due to the path-loss and multipath fading; as for the UT mobility we have assumed that each UT can move inside the cell with a maximum speed of 70 Km/h. The mobility pattern is based on a Manhattan-like urban environment, whereas the mobility model is based on [11]. The path-loss model is based on a Hata model, while the multipath-fading effect has been considered by resorting to the Markovian model proposed in [12].

There are mainly two types of AMC techniques: maximum throughput AMC, in which the Modulation and Coding Scheme (MCS) is selected to achieve the best overall throughput without any constraint on the data reliability (i.e., bit error probability), and minimum bit error probability AMC, in which conversely, the main goal is to meet specific data reliability constraints and, hence, the MCS is selected accordingly.

In this paper we assume that data reliability is the main concern so that we refer in what follows to the second technique. Let $P_{\text{ber}}$ be the Target Bit Error Probability then the transmission rate (per Hertz) $c_{k,n}(i)$ for subcarrier $n$ of user $k$ on time slot $i$ can be expressed as a function of the local signal-to-noise ratio (SNR) $\gamma_{k,n}(i)$ and $P_{\text{ber}}$ as in [13], [14]:

$$c_{k,n}(i) = \left\lceil \log_2 \left( 1 + \frac{-1.5}{\ln(5P_{\text{ber}})} \gamma_{k,n}(i) \right) \right\rceil$$

Parameter $c_{k,n}(i)$ defined in (1) represents the maximum transmission rate that can be achieved by the user $k$ on the subcarrier $n$ by maintaining the bit error probability $P_{\text{ber}}$ under $P_{\text{ber}}$. In particular, the MCS can be selected according to the instantaneous local SNR value for user $k$ on the subcarrier $n$ from the MCSs set foreseen by the standard. Let $c$ denote the modulation scheme order defined (1), i.e., the number of bits transmitted per symbol time, by neglecting the dependence from user, subcarrier and slot number, the MCS with modulation scheme order $c$ is selected if we have [14]:

$$\Gamma_c \leq \gamma_{k,n}(i) \leq \Gamma_{c+1}$$

(2)

with $\Gamma_c$ defined as:

$$\Gamma_c = \frac{(2^c - 1) \ln(5P_{\text{ber}})}{-1.5}$$

(3)

Note that $c = 0$ means that no transmission is possible.

B. Scheduling Strategy

The aim of the scheduling strategy proposed in this paper is to perform an optimal allocation of the network resources among the users in order to maximize the overall network throughput and meet the user QoS constraints in terms of minimum bitrate needed, $P_{\text{min}}^k$, and $P_{\text{ber}}$. From above, it follows that the objective is to search for the subcarriers allocation matrix $M$ for which we have:

$$\max_k \sum_{k=0}^{K-1} \sum_{n \in M_k} r_k(n)$$

such as

$$\sum_{k=0}^{K-1} \delta[r_k(n)] = 1, \forall n$$

$$\sum_{k=0}^{K-1} \delta[r_k(n)] \leq 1, \forall n$$

(4)

where $r_k(n)$ is the bitrate achieved by user $k$ on subcarrier $n$, $\delta[.]$ is the Kronecker delta function, $M_k$ is the allocation matrix for user $k$, $K$ is the total number of users and $N$ the total number of subcarriers. Note that $M$ can be considered as a time-frequency grid with the x-axis and y-axis formed respectively by the number of OFDM symbols contained in a frame and all the subcarriers. A direct search for $M$ according to (4) is computationally too expensive [5], so that in this paper the solution of the scheduling problem has been achieved in two steps: in the first step, named Resource Allocation step, the scheduler has to define the amount of resources to allocate to users in relation to their QoS constraints while at the second step, named Subcarrier Allocation step, the scheduler searches for the optimal subcarriers allocation among users on the basis of the knowledge of the propagation conditions for each subchannel.

III. RESOURCE ALLOCATION ALGORITHMS

In this section two algorithms to fulfill resource allocation step of the scheduling strategy under consideration will be proposed. The first, identified as Weighted Fair Queuing Resource Allocation, provides a fair allocation among the users, while the second, named Maximum Throughput Resource Allocation, has to maximize the network throughput without considering fairness in performing the resource allocation. Both strategies foresee a pre-allocation phase to assign resources to users for which a minimum bitrate value has to be guaranteed.

In the Weighted Fair Queuing (WFQ) Resource Allocation a weight, proportional to the bandwidth (i.e., data rate) needs,
is assigned to each user as:

$$w_k = \frac{B_k}{\sum_{k=0}^{K-1} B_k}$$  \hspace{1cm} (5)$$

where $B_k$ denotes the additional bandwidth with respect to that achieved by the pre-allocation phase needed to the $k$-th user to empty his queue. This weight represents the percentage of the total resources available, after completion of the pre-allocation phase, that this user will get by the subcarrier allocation step.

In the Maximum Throughput (MT) Resource Allocation strategy, after having performed the pre-allocation phase, the scheduler aim is to maximize the overall throughput of the wireless network by allocating all the required resources to users having favorable channel propagation conditions and hence, able to achieve the highest throughput. After completion of this task, if resources are still available, they are assigned to users seeing bad channel propagation conditions. It is considered that the minimum amount of resources that can be allocated to a user is one slot (i.e., not a single subcarrier) and that a same slot cannot be allocated to more than one user at the same time.

IV. SUBCARRIER ASSIGNMENT ALGORITHMS

This section deals with the subcarrier assignment step. In particular, three novel approaches will be presented. In addition to this, the static allocation algorithm will be also considered in order to have a benchmark in performing performance comparisons.

The first proposed algorithm has been derived as a modified version of the well known Rate Craving Greedy (RCG) proposed in [6], by considering the slot allocation instead of the subcarrier allocation. The resources to be allocated into the subchannels have been selected by using the WFQ strategy described in Section III. This algorithm operates in two steps: as the first step, for each subchannel $s$ the algorithm searches for the user $\tilde{k}$ that has the best channel gain $|\tilde{H}_k(s)|^2$ for that subchannel [6] (if any):

$$\tilde{k} = \arg \max_{0 \leq k \leq K-1} |\tilde{H}_k(s)|^2, \hspace{1cm} (6)$$

where:

$$|\tilde{H}_k(s)|^2 = \frac{\sum_{n=1}^{48} |H_k(\tilde{n})|^2}{48}, \hspace{1cm} \tilde{n} \in s.$$  \hspace{1cm} (7)$$

Hence, all the idle slots (if any) belonging to the subchannel are assigned to user $\tilde{k}$ without considering the slots allocated after the completion of the resource allocation step. From (7) it follows that $|\tilde{H}_k(s)|^2$ is derived as the average of the single channel gain of the user $k$ on each subcarrier $\tilde{n}$ that belongs to the subchannel $s$. We can note from Algorithm 1, outlined below, that after the first step completion, users could have an excess or lack of allocated slots. As a consequence, the second step has as goal that of solving and lowering this unfairness. In particular, for each user $\tilde{k}$ with an excess of allocated slots, the algorithm searches for the user $\hat{l}$ belonging to the group $\Omega$ of users with lack of allocated slots, for which it exists a subchannel $\hat{s}$ allowing the lowest difference in channel gain with respect to any other user belonging to $\Omega$ and possible subchannels belonging to $k$. Whenever user $\hat{l}$ is found, a slot $\hat{n}$ (belonging to subchannel $\hat{s}$) is removed from user $\tilde{k}$ and assigned to user $\hat{l}$. This procedure is started over again until there are users with lack of allocated slots.

**Algorithm 1 RCG Algorithm**

**Ensure:** slot$_k$ is the number of slot that the algorithm has to allocate to user $k$ and $C_k$ is the set of subcarrier allocated to the user $k$.

**for all** $\tilde{k}$ such that $0 \leq \tilde{k} \leq S - 1$ **do**

$$\hat{k} = \arg \max_{0 \leq k \leq K-1} |\tilde{H}_k(s)|^2$$

$$C_k \leftarrow C_k \cup \{s\}$$

**end for**

**for all user** $\tilde{k}$ such that $\#C_k > \text{slot}_k$ **do**

$$\hat{l} = \arg \min_{l \in \Omega} \min_{0 \leq \hat{s} \leq \hat{S}-1} \{ |\hat{H}_k(\hat{s})|^2 - |\tilde{H}_k(\tilde{s})|^2 \}$$

$$\hat{s} = \arg \min_{0 \leq \hat{s} \leq \hat{S}-1} \{ |\hat{H}_k(\hat{s})|^2 - |\tilde{H}_k(\tilde{s})|^2 \}$$

$$C_k \leftarrow C_k \setminus \hat{n}, \hspace{1cm} C_{\hat{l}} \leftarrow C_{\hat{l}} \cup \hat{n}$$

**end while**

**end for**

The second algorithm proposed is named as the Best Slot Selection (BSS) algorithm and is outlined below in Algorithm 2. Likewise the RCG algorithm, the BSS alternative is based on a two steps procedure and it follows the resource allocation phase performed according to the WFQ algorithm. As first step, the BSS algorithm searches for the user $\bar{k}$ having at that time the lowest percentage of assigned slots defined as the ratio between the slots $s_k$ assigned to a certain user in previous cycles by the BSS algorithm, and the slots (slot$_k$ in Algorithm 2) that $\bar{k}$ has to have allocated after the completion of the resource allocation phase. When BSS starts, all users have a 0 percentage of assigned slots because $s_k$ is zero for all of them so that the algorithm selects randomly the user $\bar{k}$ from which to start. At the second step BSS first finds the best subchannel $\tilde{s}$ for the user $\bar{k}$ that has at least one idle slot $\tilde{n}$ (it chooses the subchannel in which $\bar{k}$ has the highest channel gain), and then the scheduler assigns that slot to the user $\bar{k}$. BSS iterates this two steps procedure until there are idle slots. Note that this algorithm has lower computational complexity than RCG.

The last algorithm proposed is named Maximum Throughput Allocation (MTA); it exploits the MT resource allocation defined in Section III and aims to maximize the overall throughput also in the subcarrier assignment. This alternative is outlined below as Algorithm 3. The MTA algorithm needs an additional information about the terms $\Gamma_c$ defined by (3).
Algorithm 2 BSS Algorithm

\begin{algorithm}
\textbf{for all} \ n \ \textbf{such that} \ 0 \leq \ n \leq \ N - 1 \ \textbf{do}
\begin{align*}
\hat{k} &= \arg\min_{0 \leq k \leq K-1} \frac{s_k}{\text{slot}_k} \\
\hat{s} &= \arg\max_{0 \leq s \leq S-1} |\hat{H}(s)|^2 \\
C_{\hat{k}} &\leftarrow C_{\hat{k}} \cup \{\hat{s}\}
\end{align*}
\textbf{end for}
\end{algorithm}

In particular, MTA starts from the highest MCS (i.e., the one allowing the highest transmission bitrate) and then, by analyzing each subchannel \( s \), searches for a user \( k \) that has a \( \gamma_{k,n}(i) \) high enough to allow data transmissions on \( s \) with that MCS. If there are more than one user for which this condition is fulfilled the algorithm selects the one with the lowest total SNR (\( \bar{\gamma}_{\text{TOT},k} \)), defined as:

\[ \bar{\gamma}_{\text{TOT},k} = \frac{\sum_{s=1}^{S} \hat{\gamma}_{k,s}}{S}. \] (8)

According to this procedure, we make possible to allocate to users those slots on which they can transmit at the highest bitrate, hence, the overall network throughput is maximized. Then MTA proceeds to assign to \( \hat{k} \) the idle slots belonging to that subchannel that \( \hat{k} \) has to receive after the resource allocation phase. If \( \hat{s} \) does not contains enough idle slots to satisfy the need of \( \hat{k} \), all the slots will be assigned \( \hat{k} \) and the missing ones will be assigned to him by the next iteration of the MTA algorithm (the set of slots belonging to \( \hat{s} \) that are assigned to \( \hat{k} \) are denoted in Algorithm 3 below as \( \Lambda \)).

Algorithm 3 MTA Algorithm

\begin{algorithm}
\textbf{for} \ MCS = MCS_{\text{max}} : MCS_{\text{min}} \ \textbf{do}
\begin{algorithm}
\textbf{for all} \ s \ \textbf{such that} \ 0 \leq s \leq S - 1 \ \textbf{do}
\begin{align*}
\hat{k} &= \arg\max_{0 \leq k \leq K-1} \bar{\gamma}_{k,s} \\
C_{\hat{k}} &\leftarrow C_{\hat{k}} \cup \Lambda
\end{align*}
\textbf{end for}
\end{algorithm}
\textbf{end for}
\end{algorithm}

V. NUMERICAL RESULTS

We focus here on an IEEE 802.16 network with 10 MHz channels, a TDD transmitting structure and an OFDMA frame with time duration equal to 8 ms. Moreover, we assume to have divided the frame in two equal parts, so that the uplink and downlink subframe length is 4 ms. The 10 MHz channel choice means that for each symbol there are 768 available data subcarriers. As for the number of users in the cell, we consider 20 UTs with VoIP traffic modeled according to the G.729 codec with Voice Activity Detection, and 50 UTs with HTTP traffic according to a trunked Pareto distribution. Moreover, we have chosen \( P_{\text{ber}} \) equal to \( 1 \cdot 10^{-3} \), and hence, from (1), we have:

\[ r_{k,n} = \left[ \frac{\Delta f \cdot T_f}{N_{\text{symb}}} \cdot \log_2 \left( 1 + \frac{-1.5}{ln(5 \cdot P_{\text{ber}} \cdot \gamma_{k,n})} \right) \right] \] (9)

where \( \Delta f \) is the bandwidth associated to the subcarrier \( n \); \( T_f \) is the OFDMA frame length and \( N_{\text{symb}} \) is the number of OFDM symbols in a frame. Starting from (9) it has been possible to find the local SNR thresholds that allows to decide for a given user which is the modulation scheme to transmit with. In particular by (3) and according the channel model outlined in Section II-A, the parameter \( \tilde{\gamma}_{k,s} \) results to be:

- \( 10.3 \, \text{dB} \leq \tilde{\gamma}_{k,s} < 17.25 \, \text{dB} \Rightarrow \text{QPSK} \)
- \( 17.25 \, \text{dB} \leq \tilde{\gamma}_{k,s} < 23.5 \, \text{dB} \Rightarrow \text{16QAM} \)
- \( \tilde{\gamma}_{k,s} \geq 23.5 \, \text{dB} \Rightarrow \text{64QAM} \)

By taking into account the fact that QPSK is the modulation scheme with the lowest admitted number of bits per symbol, if \( \tilde{\gamma}_{k,s} \) is lower than 10.3 dB the constraint on \( P_{\text{ber}} \) cannot be satisfied so that we consider as lost all the transmitted bits in that slot.

In Fig. 1 the overall network throughput at the BS side as a function of the traffic load is shown for the downlink case. The results for the uplink case have not reported here due to space limitations; however we can say that, their behavior is very similar to the downlink case, mostly due to the fact that uplink and downlink sub-frames have the same length and, so, the same amount of resources.

The performance of the proposed schemes are compared in Fig. 1 with the static subcarrier allocation approach. The performance of the RCG algorithm, previously proposed, has been considered as benchmark. It is evident in Fig. 1 that all the dynamic algorithms outperform the static case. Moreover, Fig. 1 highlights that MTA achieves the best overall throughput performance (about 60% better than the Static algorithm) at the expense, as outlined before, of an unfair resource allocation among users. Conversely, RCG and BSS algorithms allow a fair allocation among users but, unfortunately, also a lower throughput, due to the fact that use a weighted fair queuing resource allocation (about 30% higher than the Static case).

In Fig. 2 the average queue length at the BS side as a function of the traffic load is shown for the downlink case. Once again this figure shows that the three proposed dynamic algorithms outperforms the Static alternative. As before the best performance are achieved by the MTA algorithm, while the RCG and BSS highlight similar performance.

Fig. 3 shows the performance in terms of number of times that the BS cannot transmit during downlink communications (outage) as a function of the traffic load toward MSs. As a remark, a user is in outage if the SNR is below the threshold of 10.3 dB considered as the minimum value for which the QPSK communication can be received. By analyzing the behavior of the algorithms we can see that for all the algorithms the curves has a maximum for a certain traffic load value, while for higher traffic load the number of outages is lower. This is due to the fact that a low traffic load means a lower number...
of users that have traffic to send. Because we have assumed that all the resources have to be allocated, when the number of users is lower it happens more frequently that bad channel slots are assigned to users, while for a higher number of users the probability that, given a slot, there is at least one user for which the channel is good is higher. Once again, the MTA algorithm outperforms all the other alternatives.

VI. CONCLUSION

In this paper the resource allocation problem for an IEEE 802.16 wireless network has been considered. The scenario has been composed of a variable number of users inside a cell served by a central BS. The problem of finding an optimal subcarrier and bit allocation strategy for uplink and downlink communications was divided in two steps: resource allocation and subcarriers (or slots) allocation. Three dynamic algorithms, belonging to two different families of resource allocation strategies, were proposed and analyzed. In particular, it has been clearly pointed out in the paper that all the proposed techniques have better performance than the static allocation solution. Finally, the proposed MTA algorithm has resulted to be the best solution; on the other hand the BSS algorithm has performance slightly worse than the RCG alternative but with a significantly lower computational complexity.

REFERENCES