Resource allocation based on cross-layer QoS-guaranteed scheduling for multi-service multi-user MIMO–OFDMA systems

Samy Kambou[1] · Clency Perrine[1] · Meriem Afif[1] · Yannis Pousset[1] · Christian Olivier[1]

Abstract Cross-layer strategies for resource allocation in wireless networks are essential to guaranty an efficient utilization of the scarce resource. In this paper, we present an efficient radio resource allocation scheme based on PHY/MAC cross layer design and QoS-guaranteed scheduling for multi-user (MU), multi-service (MS), multi-input multi-output (MIMO) concept, orthogonal frequency division multiple access (OFDMA) systems. It is about a downlink multimedia transmission chain in which the available resources as power and bandwidth, are dynamically allocated according to the system parameters. Among these parameters, we can mention the physical link elements such as channel state information, spectral efficiency and error code corrector rate, and MAC link variables, which correspond to the users QoS requirements and the queue status. Primarily, we use a jointly method which parametrizes these system parameters, according to the total power, and the bit error rate constraints. Secondly, we propose a QoS-guaranteed scheduling that shares the subcarriers to the users. These users request several type of traffic under throughput threshold constraints. The main objective in this work is to adjust the average throughput per service of each user, according to their needs and likewise to satisfy a great number of connexions. Subsequently, we consider a model of moderated compartmentalization between various classes of services by partitioning the total bandwidth into several parts. Each class of service will occupy a part of the bandwidth and will be transmitted over a maximum number of sub-carriers. The simulation results show that the proposed strategy provides a more interesting performance improvement (in terms of average data rate and user satisfaction) than other existing resource allocation schemes, such as nonadaptive resource allocation strategy. The performances are also analyzed and compared for the two multi-service multi-user MIMO–OFDMA systems; with sub-carriers partitioning and without sub-carriers partitioning.

Keywords MIMO–OFDMA · Cross-layer design · Joint parameterization · Resources allocation · Multi-user · Multi-service · QoS-guaranteed scheduling

1 Introduction

Recently, the use of mobile devices such as cell-phones, tablets or laptops are increasingly penetrating into our daily lives as convenient tools for communication, work, social network, news, etc. To face the growing demand of high data rate multimedia services, the next generation wireless communication systems should be able to provide high speed broadband Internet services that have various quality of service (QoS) requirements. However, with scarce radio resource (power and bandwidth) and harsh wireless channel conditions, some techniques which lead to an efficient utilization of these resources are quite necessary. So, the new generation wireless systems should be designed, taking these factors into account. In this context, 3GPP Long Term Evolution (LTE)/LTE-Advanced [12] and WiMax (Worldwide Interoperability for Microwave Access) [4] IEEE 802.16m technologies can be considered as some suitable candidates to meet users expectation. Indeed, their radio interface easily meet high data rate requirements,

✉ Samy Kambou
  samy.kambou@univ-poitiers.fr

[1] Department of Signal Image Communication, XLIM Institute, CNRS JUR 7252, University of Poitiers, Poitiers, France

thanks to the efficient approaches, such as multi-input multi-output (MIMO) [19, 22] and orthogonal frequency division multiple Access (OFDMA) [28, 31] techniques.

1.1 Background

MIMO system, which employs multiple antennas at the transmitter and the receiver, is a promising technique in wireless environments to significantly enhance performance improvements in terms of data rate and diversity, without increasing the transmit power or bandwidth. Indeed, MIMO technology allows to combat the fading effect, exploiting multi-path propagation, and uses the radio channel efficiently by transmitting several user bit-streams in the same resource. OFDMA is about an emerging multiple access technique based on OFDM (Orthogonal Frequency Division Multiple). This technique offers high spectral efficiency and better resistance to fading environment. It is considered as an efficient method to mitigate the Inter-Symbol Interference (ISI) effects in frequency-selective fading channels by dividing the bandwidth into a number of lower-rate sub-carriers. Thus, in the multi-user communication system, different sub-carriers can be allocated to different users to exploit multi-user diversity. The association of MIMO and OFDMA techniques in multi-user configuration, called multi-user MIMO–OFDMA (MU MIMO–OFDMA) system [17, 20], is an efficient way for providing high data rate, taking into account multi-user, frequency and spatial diversities. This association is already included in several standard applications.

In the MU MIMO–OFDMA system, multi-user interference occurs, when several users communicate simultaneously. This phenomenon considerably degrades the system performance. For this reason, the authors in [23] and [24] have proposed a novel heuristic strategy which shares the users in different groups, on the basis of their average channel gain. Afterwards, channel allocations are sequentially addressed, starting from the group with the most adverse channel conditions. The spatial dimension is employed to present multiple access interferences, hindering the performance of the sequential allocation. Otherwise, the adaptive strategies with respect to power allocation [6, 8], sub-carriers distribution [13, 35], modulation and coding [21, 37], and beam-forming [30, 34] can be used in order to improve the performances. Hence, adaptive radio resource allocation schemes in MU MIMO–OFDMA systems have become an important research topic. In this context, the authors in [15, 16, 32] and [26] investigate a power and sub-carrier allocation problems for MU MIMO–ODFMA systems. They study an allocation strategy, which involves adaptive sub-carrier allocation, adaptive modulation and coding and eigen beam-forming.

This scheme achieves a significant improvement in overall system performance either by minimizing the total transmit power subject to bit error rate (BER) constraint [15, 16, 32], or by maximizing the sum-rate capacity [26] under total power budget constraint. However, this resource allocation scheme doesn't consider the various heterogeneous classes of service and different QoS requirements for different services. In addition, beside the system sum-rate capacity and efficient autonomy, QoS requirements are also significant factors to take into consideration. Therefore, it is necessary to study efficient scheduling policy that ensures the satisfaction of QoS requirements for multi-user and multi-service.

In the paper [27], the authors describe and compare some scheduling algorithms in MU MIMO–OFDM systems. The main goal in this paper is to improve the systems capacity and fairness among users, according to an efficient utilization of available bandwidth. However, this study is limited to a single service configuration and the radio resource allocation is not taken into account. Indeed, the authors only use the conventional scheduling techniques, developed in 3GPP LTE systems, such as Round Robin (RR), Max-SNR and Proportional Fair (FP) algorithms. Thus, the authors in [9] proposed an adaptive resource allocation scheme in OFDMA based multi-service WiMAX systems. This resource allocation scheme, which integrates an adaptive modulation and coding (AMC), is performed to satisfy, at first, various QoS service requirements of different users, and secondly, to improve the Multi-Service Multi-User OFDM (MS MU OFDM) system data rate next. Nevertheless, it remains suboptimal without the consideration of MIMO technology. Consequently, the authors in [3] propose a multi-service multi-user MIMO–OFDMA (MS MU MIMO–OFDMA) in order to improve the system capacity. However, the transmission power is not optimized. In fact, an equal power allocation (EPA) is used, which decreases the system performance. In order to improve the performance, an unequal power allocation (UPA) strategy is adopted in [29] and [14]. The authors optimize the power transmission with a guarantee of the user's data rate requirements, but don't consider adapted data protection also known as Unequal Error protection (UEP), neither multi-service concepts.

1.2 Contribution

The main objective of our proposal is to investigate an efficient radio resource allocation, based on PHY/MAC cross layer scheme and QoS-guaranteed scheduling for MIMO–OFDMA systems. It's about a radio system, supporting multi-user and real-time, non-real-time and best effort multimedia traffics. Algorithms are incorporating these four aspects (MIMO, OFDMA, multi-user and multi-

service) which are proposed; involving an adaptive power allocation method, an Adaptive Modulation (AM), a coding scheme, an adaptive sub-carriers assignment, and an original scheduling policy. They are executed at the beginning of every frame, to correctly allocate the radio resource to the demanding users, according to their queue status, instantaneous channel conditions and QoS requirements.

Figure 1 below, summarizes the main steps of the proposed scheme. Firstly, we propose an optimal physical (PHY) and medium access control (MAC) cross layer approach based on UPA, UEP and AM techniques. These three methods are jointly considered according to the physical layer and the MAC sublayer constraints. Concerning PHY layer constraints, we consider transmission power strategy connected a QoS in terms of target bit error rate. The MAC sublayer constraint correspond to a service throughput threshold. The basic idea is that the sub-channel (ensuring the best transmission conditions), will be at first, shared by the appropriate user. In fact, we propose to classify the available sub-carriers according to their gain. Then, the more robust sub-carriers will be allocated primarily. For that reason, we can formulate the optimal PHY/MAC cross-layer design into an optimization problem with the objective to maximize the users throughput under constraints (total power and target bit error rate).

Additionally, we consider a QoS-guaranteed scheduling scheme, which is subdivided into two steps. The first one is a sub-carriers allocation scheme and the second one is a service scheduling process. We propose a best-SNR algorithm to allocate the specific sub-carriers to a given user according to the previous joint parameterization scheme, as well as the users QoS requirements. We also define a service scheduling priority function for each user and update it dynamically, depending on service parameters, each user data rate and each user status queue.

Therefore, an efficient radio resource allocation, based on PHY/MAC cross-layer approach and QoS-guaranteed

Fig. 1 Synoptic of the proposed scheme

scheduling scheme, is satisfying in order to guarantee at first, the users QoS requirements and secondly, maximizing the requested connections. Furthermore, we also consider a moderated repartition of the frequency resources between different classes of service, by partitioning the bandwidth into several parts. Thus, each service class will occupy a part of the bandwidth, therefore, it will be transmitted over a maximum number of sub-carriers. The final goal is to avoid that a priority user always requesting a priority service, saturate the available bandwidth.

The remainder of this paper is organized as follows. In Sect. 2, the system model of MS MU MIMO–OFDMA is described by presenting the PHY/MAC interaction algorithm and formulating the optimization problem. Section 3 suggests a new sub-carrier allocation and service scheduling policy approaches. The simulation results are analyzed and discussed in Sect. 4. Finally, conclusion are drawn in Sect. 5.

2 System model

2.1 System description

The downlink transmission scheme of the MS MU MIMO–OFDMA system is depicted in the Fig. 2. The Base Station (BS), which is the transmitter, communicates with K users called Mobile Equipments (MEs) (with $k \in \{1, \ldots, K\}$, the user index). The total bandwidth B is divided into N sub-carriers (with $n \in \{1, \ldots, N\}$, the sub-carrier index), which are shared by the users. The BS has M_T transmit antennas and each ME is equipped with M_R antennas. Each data user is packed into fixed length packet and a separate queue of each service is maintained for each user at the BS. These data user are modulated with M-QAM modulation to increase the spectral efficiency. Later, the RCPC (rate compatible punctured convolution) error code corrector (ECC), is used to ensure the detection and the correction of transmission errors. The modulated and coded signals are sent to inverse fast Fourier transform (IFFT) module to do OFDM modulation and the cyclic prefix (CP) is added to every OFDM symbol. Finally, the OFDM symbols are transmitted from each antenna. The OFDM modulation allows the designing of the N sub-carriers which will be distributed to different users. In order to avoid the ISI, we set that a sub-carrier is allocated to only one user during the frame time. The data, which is transmitted through one or more sub-carriers ($\Omega_k \subset \{1, 2, \ldots, N\}$ denote the set of sub-carriers allocated to a user k) of the MIMO channel, is received by the users. An imperfect knowledge of the CSI at both the transmitter side and the receiver side is considered. Indeed, the CSI is updated at regular intervals in order to consider a realistic transmission configuration. We
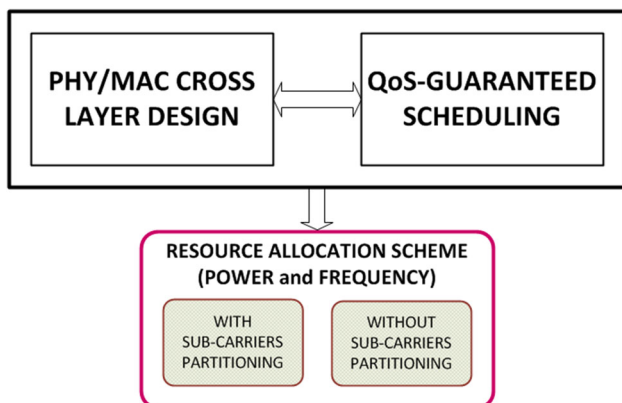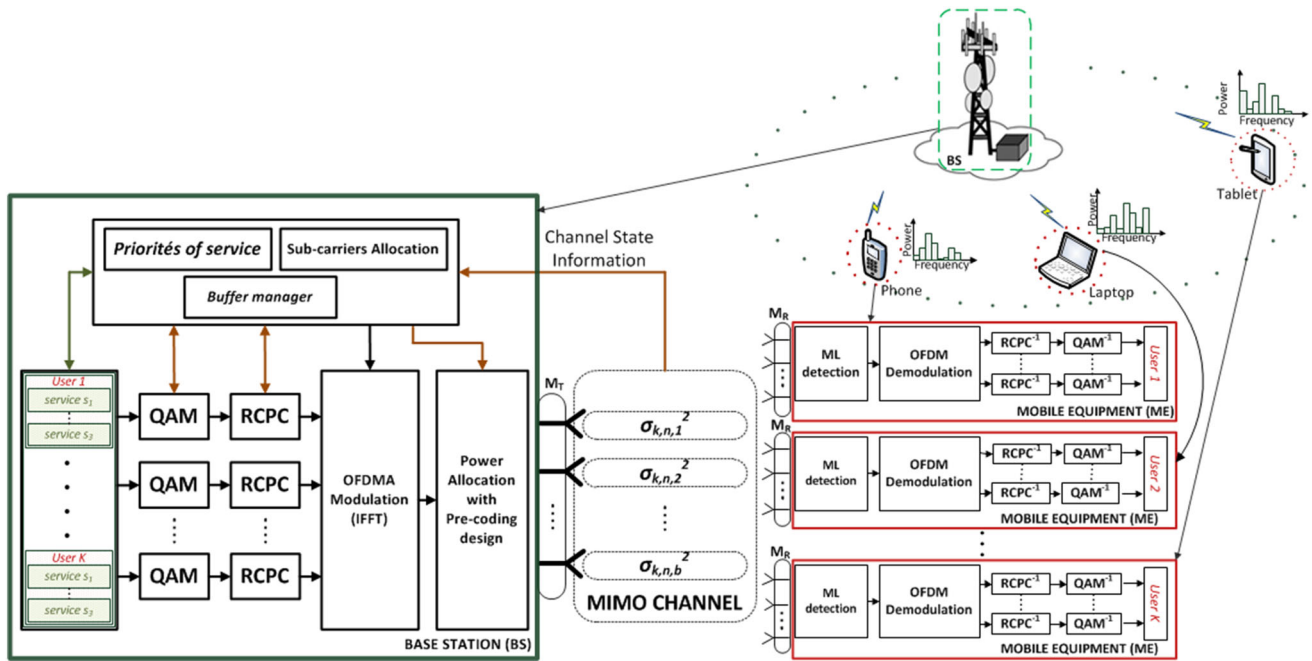
Fig. 2 Blocks diagram of proposed transmission scheme (MS MU MIMO–OFDMA model)

carried out a new channel estimation every 20 OFDM symbols to minimize the channel variations. This technique is already used in [2], and its performance was compared to other methods of channel estimation, found in [10, 33]. These models follow a Gaussian distribution. Afterwards, a precoder design is used to decompose each sub-carrier of each user into b (with $b = \min\{M_R; M_T\}$) uncorrelated SISO sub-channels, denoted $\sigma_{k,n,sc}^2$ (with $sc \in \{1, \ldots, b\}$, the sub-channel index). These virtual links between BS and MEs can be assumed as selective frequencies Rayleigh fading. Considering the OFDM modulation, these selective frequencies fading are considered as narrow-band sub-carriers, which can be approximated as the Gaussian channels. They are considered as constant during an allocation period. At each receiver, the maximum likelihood (ML) [38] criterion is used to detect the received OFDM symbols. Thus, the receivers can be considered as linear since the virtual sub-channels are parallel and independent. In this context, the ML detection simply inverses the virtual sub-channels [10]. Onces the OFDM symbols are detected by the ML detection, the OFDM demodulation block provides digital symbols, thanks to the FFT module. These digital symbols were coded with a RCPC code. For this reason, we apply the channel decoding operation. Subsequently, the decoded digital symbols pass through the QAM demodulation blockto demodulate the information bits of each user.

The majority of these parameters are usually used in the next generation wireless systems such as IEEE 802.16, LTE or LTE/Advanced standards.

Assuming that the BS provides three types of traffics: the real-time service such as NRTV (near real time video), called real time polling service (RTPS)(s_1), the non-real-time such as the FTP (file transfer protocol), called non real time polling service (NRTPS)(s_2) and the Best Effort (BE) (s_3) services such as WEB navigation [11]. A discret Markov Modulated Poisson Process (dMMPP) model [25] is used for packet arrival process relative user. This model is more realistic due to the flexibility that it offers to manage the packets arrival of various types of traffics for each user. The dMMPP model for each user can be represented by a probability transition matrix U of the modulating Markov chain and a matrix λ modeling the Poisson arrival rate for each service.

$$U = \begin{pmatrix} U^{11} & \cdots & U^{1z} \\ \vdots & U^{ij} & \vdots \\ U^{z1} & \cdots & U^{zz} \end{pmatrix}, \quad \lambda = \begin{pmatrix} \lambda^1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda^z \end{pmatrix}. \quad (1)$$

where z is the state number of the Markov process. $U^{i,j}$ represents the service probability transition from the state i (with a number of packets pck_i) to the state j (with a number of packets pck_j), where $|pck_j - pck_i|$ are incoming or outcoming packets of the buffer.

2.2 PHY/MAC cross layer approach

In this section, the configuration steps of the MAC sublayer and the physical layer interaction is described. Indeed, the first objective of our study, is to investigate a joint PHY/MAC cross-layer design, which consider a MAC sublayer and PHY layer constraints. Therefore, we present in the subsection below, the MAC sublayer and PHY layer interactions.

2.2.1 MAC sublayer

At the MAC sublayer, we suppose that for a sampled frame time μT (where $\mu \in \mathbb{N}$ and T is the frame time), there are $m_k^s(\mu T)$ incoming packets of service s in the queue $Q_k^s(\mu T)$ of the user k and the number of transmitter bit at μT equals to $D_k^s(\mu T)$. It is assumed that, each service s has a constant number of bits per packet, noted nb_p^s.

Figure 3 illustrates an example of the evolution process of two users (k_1 and k_2) queues. These two users request the two services (s_1 and s_2). We can observe that:

- $m_{k_1}^{s_1}(\mu T) = 3$ packets (with A, B and C, which are some packets of the service s_1) are destined to the user k_1.
- $m_{k_2}^{s_1}(\mu T) = 1$ packet (with a which is a packet of the service s_2) is destined to the user k_2, and $m_{k_2}^{s_2}(\mu T) = 2$ packets (with b and c, which are some packets of the service s_2) are destined to the user k_2.
- k_1 is served with a number of bit equal to: $D_{k_1}^{s_1}(\mu T) = 2\,\text{packets} \times nb_p^{s_1}$ (A and B packets of service s_1).

- k_2 is served with a number of bit equal to: $D_{k_2}^{s_2}(\mu T) = 2\,\text{packets} \times nb_p^{s_2}$ (b and c packets of service s_2).
- At the end of this frame time μT, users queues status are equal to:

  - $Q_{k_1}((\mu+1)T) = Q_{k_1}^{s_1}((\mu+1)T) + Q_{k_1}^{s_2}((\mu+1)T) = 1\,\text{packet} + 1\,\text{packet}$ (where C is a packet of the service s_1 and a is a packet of service s_2).
  - $Q_{k_2}((\mu+1)T) = 0$. All packets destined to the user k_2, have been served.

During the next frame time $(\mu+1)T$, there are new packets (D, E, F, G, d and e) in addition to the buffer. Once the users k_1 and k_2 are respectively served with a number of bits $D_{k_1}^{s_1}((\mu+1)T) + D_{k_1}^{s_2}((\mu+1)T)$ and $D_{k_2}^{s_1}((\mu+1)T) + D_{k_2}^{s_2}((\mu+1)T)$, the users queues status ($Q_{k_1}((\mu+2)T)$ and $Q_{k_2}((\mu+2)T)$) are computed at the end of this frame time. The same process is used to determine the users queues status at the end of the following frame time, and so on.

Generally, we can follow the evolution of users queues, thanks to the use of the dMMPP model. So, the queue length at the end of the frame time μT of user k, requesting a service s, is given as:

$$Q_k^s((\mu+1)T) = Q_k^s(\mu T) + m_k^s(\mu T) - \frac{D_k^s(\mu T)}{nb_p^s} \quad (2)$$

According to equation (1), the arrival rate of the Poisson process $\lambda_k^s((\mu+1)T)$ of a user k, requesting a service s, can be predicted with the given parameter $\lambda_k^s(\mu T)$ such as:
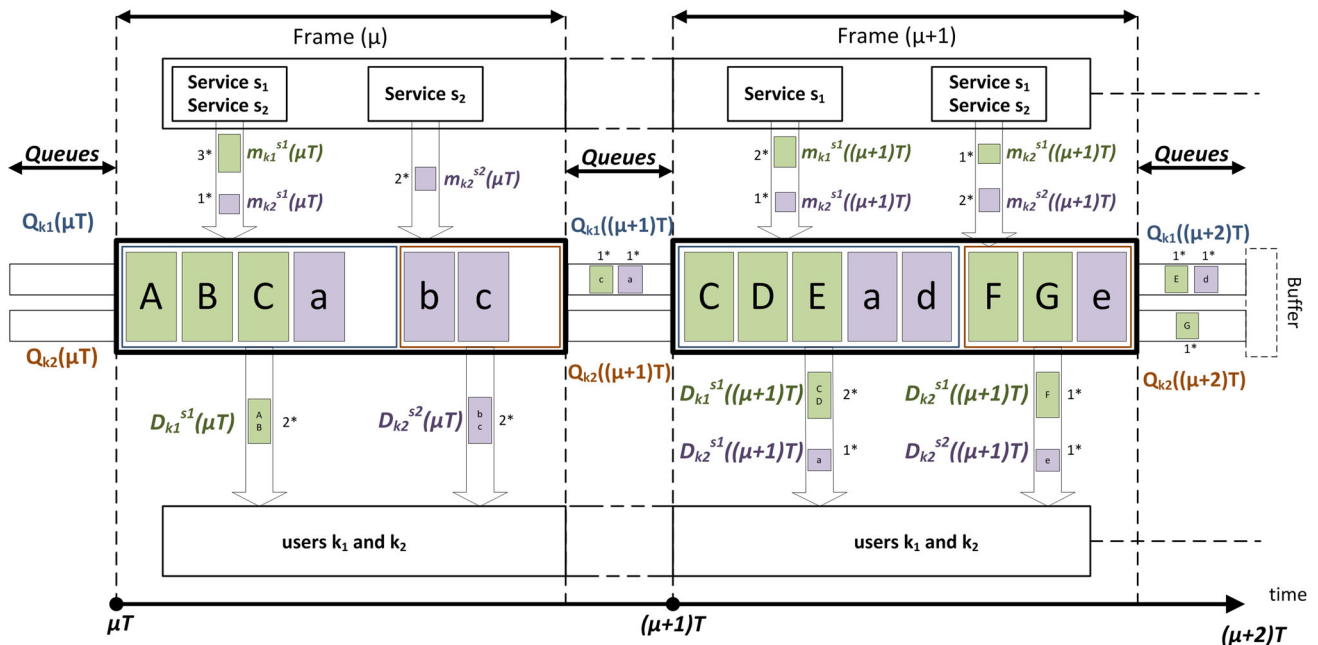

Fig. 3 User queues evolution

$$\lambda_k^s((\mu+1)T) = \sum_{z'=1}^{z} U^{z'z}.\lambda_k^s(\mu T) \qquad (3)$$

The expression of the incoming packets is deduced as follows:

$$E[m_k^s(\mu T)] = \lambda_k^s(\mu T) \qquad (4)$$

We adopt a finite queue model with a limited capacity equal to L^s. As shown in the Fig. 3, an incoming packet may be lost, if the number of packets in the queue is upper to L^s. Therefore, the real queue length at $((\mu+1)T)$ should be determined as follows:

$$Q_{k,real}^s((\mu+1)T) = \min\{L^s, Q_k^s((\mu+1)T)\} \qquad (5)$$

Since this limited capacity of the buffer, the number of dropped packets $Q_{k,drop}^s(\mu T)$ at the end of frame time (μT) due to the overflow can be predicted as:

$$Q_{k,drop}^s(\mu T) = \max\left\{0, Q_k^s(\mu T) + m_k^s(\mu T) - \frac{D_k^s(\mu T)}{nb_p^s}\right\} \qquad (6)$$

Let us define $PLR_k^s(\mu T)$, the packet loss rate as follows:

$$\begin{aligned} PLR_k^s(\mu T) &= \frac{Q_{k,drop}^s(\mu T)}{E[m_k^s(\mu T)]} \\ &= \frac{\max\left\{0, Q_k^s(\mu T) + m_k^s(\mu T) - \frac{D_k^s(\mu T)}{nb_p^s}\right\}}{E[m_k^s(\mu T)]} \end{aligned} \qquad (7)$$

2.2.2 PHY layer

At the physical layer, the MIMO–OFDMA aspect is considered. For each user k (with $k \in \{1, \ldots, K\}$), the MIMO $(M_R \times M_T)$ channel state matrix on sub-carrier n (with $n \in \{1, \ldots, N\}$) is defined by $H_{k,n}$. Then, we can deduce the state matrix H of the whole system as following:

$$\begin{aligned} H_{k,n} &= \begin{pmatrix} h_{1,1} & \cdots & h_{1,M_R} \\ \vdots & \ddots & \vdots \\ h_{M_T,1} & \cdots & h_{M_T,M_R} \end{pmatrix} \\ H &= \begin{pmatrix} H_{1,1} & \cdots & H_{1,N} \\ \vdots & H_{k,n} & \vdots \\ H_{K,1} & \cdots & H_{K,N} \end{pmatrix} \end{aligned} \qquad (8)$$

To exploit the MIMO–OFDMA characteristics, we consider a precoder design which assume a knowledge of the CSI at both sides: the receiver side and the transmitter side. This scheme allows the decomposition of the MIMO–OFDMA channel into b (with $b = \min\{M_R; M_T\}$) virtual parallel independent SISO (Single Input Single Output) sub-channels with decreasing gains, and thereby eliminates the Inter-Antenna Interference (IAI). It works like a pre-equalizer by adjusting the transmission power. Thus, we

consider an UPA algorithm under the constraint of the total transmitted power E_T.

Let us define $Y_{k,n}$ the received signal of the user k on the sub-carrier n over the MIMO–OFDM $(M_R \times M_T)$ system. The system equation, using precoder design is given as:

$$Y_{k,n} = G_{k,n}H_{k,n}F_{k,n}X_{k,n} + G_{k,n}n_o \qquad (9)$$

where $X_{k,n}$ is the $(b \times 1)$ transmitted symbol vector, $H_{k,n}$ is the $(M_R \times M_T)$ MIMO–OFDM channel matrix, $F_{k,n}$ is the $(M_T \times b)$ linear precoder matrix, $G_{k,n}$ is the $(b \times M_R)$ linear decoder matrix and n_o is the $(M_R \times 1)$ zero-mean additive noise vector.

By applying the Singular Value Decomposition (SVD) method [7], the expression of the MIMO–OFDM system is obtained as follows:

$$Y_{k,n} = G_{k,n}^d H_{k,n}^v F_{k,n}^d X_{k,n} + G_{k,n}^d n_o^v \qquad (10)$$

where $H_{k,n}^v = G_{k,n}^v H_{k,n} F_{k,n}^v$ is the diagonal MIMO channel matrix, $G_{k,n}^v$ and $F_{k,n}^v$ are unitary matrices and $n_o^v = G_{k,n}^v n_o$ is the transformed additive white Gaussian noise vector with covariance matrix $R_{n_o^v} = I_b$ (I_b is b size identity matrix), $F_{k,n}^d$ and $G_{k,n}^d$ are, respectively, the precoding and decoding matrices.

Using the SVD operation, the MIMO channel matrix $H_{k,n}$ of the user k using the sub-carrier n is consisting of singular vectors values $\sigma_{k,n,sc}^2$ (with $sc \in \{1, \ldots, b\}$) (cf. to Fig. 2) and is expressed as follows:

$$H_{k,n} = \begin{pmatrix} \sigma_{k,n,1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{k,n,b}^2 \end{pmatrix} \qquad (11)$$

Recall that the ML criterion is used to detect the received symbols. Actually the ML detection is not necessary. The MIMO channel decomposition technique simplifies the ML detection, which is independently performed on each SISO subchannel. In fact, without any loss of generality, we consider the decoding matrix $G_{k,n}^v$ as an unitary matrix I_b (with b the number of SIS0 sub-channels). Thus, the precoding matrix $F_{k,n}^v$ are only defined. We can determine the precoding parameters, denoted $\omega_{k,n,sc}^2$ (coefficients of $F_{k,n}^v$ matrix) using a UPA strategy. This technique makes the correspondence between many system elements. Therefore, we can calculate the gain of channel $\gamma_{k,n,sc}$ after the precoding step as follows:

$$\gamma_{k,n,sc} = \omega_{k,n,sc}^2 \sigma_{k,n,sc}^2 \qquad (12)$$

We adapt the UPA algorithm developed in [1], in the context of multi-service and multi-user transmission. For each user who simultaneously requests the three services (worst case), optimal precoding coefficients $[\omega_{k,n,sc}^2]^s$ are jointly computed, according to the spectral efficiency $M_{k,n,sc}$, the error corrector code rate $r_{k,n,sc}$ and the SNR of

each SISO sub-channel $\sigma^2_{k,n,sc}$ under total power constraint E_T and BER of the physical layer, required by services. These BER are denoted ber^s.

Recall that since the noise on a SISO sub-channel sc is a Gaussian noise, the expression of precoder coefficient of each service s, is defined as following:

$$\left[\omega^2_{k,n,sc}\right]^s = \frac{2(M_{k,n,sc}-1)N_0}{3\sigma^2_{k,n,sc}} \times \left[erf^{-1}\left(1 - \frac{bert^s_{sc}\sqrt{M_{k,n,sc}}\log_2(M_{k,n,sc})}{2\left(\sqrt{M_{k,n,sc}}-1\right)}\right)\right]^2 \tag{13}$$

where N_0 is the spectral density of noise power and $bert^s_{sc}$ is the target BER per sub-channel required before channel coding. It is determined using the ECC rate and sub-channel BER ber^s_{sc} required after channel decoding.

The proposed cross-layer parameterization scheme consists in doing a non-exhaustive research of available configurations ($[\omega^2_{k,n,sc}]^s$, $M_{k,n,sc}$ and $r_{k,n,sc}$), in order to reduce the complexity of the proposed scheme. Indeed, an exhaustive research considerably increases the number of configurations for which we can't compute the precoding coefficients ($[\omega^2_{k,n,sc}]^s$). Actually, a needed power to reach the BER required by a service is computed, according to the SNR value, the spectral efficiency and the ECC rate (Eq. 13). When this needed power exceeds the maximum power, the configuration gets unavailable. Therefore, a low complex method is used to only determine the available configurations, which respect the power and BER constraints. The idea consists to compute the precoding coefficients $[\omega^2_{k,n,sc}]^s$ in a hierarchical way. It was shown in [18], when the modulation order and the ECC capacity increase, the needed power to respect the required BER increases as well. For this reason, we research the available configurations in the increasing order, and stop it when the required power is upper to the maximum power. Afterward, the optimal configuration retained, will be the one that provides a throughput close to the threshold rate of service.

The constraints of the cross-layer parameterization scheme are summarized below:

- PHY layer constraints: power (E_T) and BER (ber^s)
- MAC sublayer constraint: threshold rate per service D^s_{th}

Without loss of optimality, an efficient and low-complex power allocation is established to find the optimal configuration which respects the constraints above.

Thus, the threshold rate per service can be determined in oder to minimize the packet loss rate $PLR^s_k(\mu T)$. We propose to calculate these threshold rates per service by approximating the packet loss rate to a packet loss probability $PLP^s_k(\mu T)$. The authors in [36] define the relationship between the bits errors rate and the packet loss probability. In fact, we suggested to

compute the packet loss probability $PLP^s_k(\mu T)$ according to the target bit error rate constraint $bert^s$ as follows:

$$PLR^s_k(\mu T) \approx PLP^s_k = 1 - (1 - bert^s)nb^s_p \tag{14}$$

By identification, we can deduce the threshold data rate of each service by:

$$D^s_{th} \approx nb^s_p\left(Q^s_k(\mu T) + m^s_k(\mu T) - L^s - PLR^s_k(\mu T).E\left[m^s_k(\mu T)\right]\right) \tag{15}$$

2.3 PHY/MAC cross-layer optimization

The purpose of the cross-layer scheme is to determine the optimal precoding coefficients according to total power transmission and throughput thresholds per service. This cross-layer parameterization is described in the algorithm 1. According to the channel conditions, we configured the system to ensure an optimal transmission power for each user in order to obtain a good transmission robustness and also improve each user data rate. The optimization problem for a service s, can be described as follows:

$$\left[\hat{\omega}^2_{k,n,sc}\right]^s = \arg\max_{\left[\omega^2_{k,n,sc}\right]^s} \frac{B}{N}\sum_{k=1}^{K}\sum_{n\in\Omega_k}\sum_{sc=1}^{b}\log_2\left(1+(\omega^2_{k,n,sc})^s.\sigma^2_{k,n,sc}\right) \tag{16}$$

Subject to:

First constraint **C1**: $\{\Omega_k\}_{k=1}^{K}$ sets are disjoint and form a partition of $\{1,\ldots,N\}$;

Second constraint **C2**: $\sum_{sc=1}^{b}[\omega^2_{k,n,sc}]^s \leq E_T \ \forall \ k \in \{1,\ldots,K\}, n \in \{1,\ldots,N\}$ and $s \in \{rtps; nrtps; be\}$;

Third constraint **C3**: $D^s_{th} \leq D^s_k \ \forall \ k \in \{1,\ldots,K\}$, $s \in \{rtps; nrtps; be\}$, and D^s_k must be the nearest value of D^s_{th};

Subsequently, the best user must be selected for each sub-carrier, thanks to his QoS requirements and the channel conditions equally, in order to optimize the use of both the power transmission and the total bandwidth. Therefore, a resource allocation scheme based on PHY/MAC cross layer QoS-guaranteed scheduling, should propose how to define the order of the users and various traffics to be served and to distribute the sub-carriers according to respect the users expectations.

3 Resource allocation based on cross layer QoS-guaranteed scheduling

In this section, we propose a radio resource allocation scheme based on PHY/MAC cross layer approach and QoS-guaranteed scheduling algorithm. The idea consists in choosing the optimal scenarios in terms of: *priority service* $\hat{s}$; *priority user* $\hat{k}$ and *appropriate sub-carrier* $\hat{n}$ which assure a

---

Algorithm 1 - Joint adaptation algorithm (PHY/MAC adaptation)

Input :
  K : users number, N : sub-carriers number and b : sub-channels number ;
  $\sigma^2_{k,n,sc}$: singular values of the MIMO channel ;
  $bert^s$: target bit error rate and $E_T = 1$: maximum transmission power ;
  $M^s_{k,n,sc} \in \{M_1, ..., M_{max}\}$: modulation order
  $r^s_{k,n,sc} \in \{r_1, ..., r_{max}\}$, : error code corrector rate ;

Operation :
1. **for each service** $s \in \{rtps, nrtps, be\}$ **do**
2.   **step 1 :** Compute service threshold throughput : D^s_{th} as in Eq. (15)
3.   **for each user** $k \in K$ **do**
4.   **for each sub-carrier** $n \in N$ **do**
5.   **for each sub-channel** $sc \in b$ **do**
6.   **for each modulation order** $M^s_{k,n,sc} \in \{M_1, ..., M_{max}\}$ **do**
7.   **for each error code corrector rate** $r^s_{k,n,sc} \in \{r_1, ..., r_{max}\}$ **do**
8.   **step 2 :** Compute the target BER ($bert^s_{sc}$)
                              before channel coding according to ber^s and $r^s_{k,n,sc}$
9.   **end**
10.   **step 3 :** Compute the precoding coefficient $(\omega^2_{k,n,sc})^s$ according to
                            Eq. (13), under power (E_T), target BER ($bert^s$) and service
                            threshold throughput (D^s_{th}) constraints
11.   **end**
12.   **step 4 :** Compute the gain of channel : $\gamma^s_{k,n} = [\omega^2_{k,n,sc}]^s \cdot \sigma^2_{k,n,sc}$
13.   **end**
14.   **end**
15.   **end**
16. **end**

---

good trade-off between resource allocation and users QoS requirements. Firstly, we define the service scheduling priority functions for all users on all sub-carriers. To determine the user $\hat{k}$ with the highest priority, requesting the priority service $\hat{s}$, we proceed to the computation of these priority functions and then, the pair $\{s; k\}$, who has the greater priority function value, is selected. Afterward, we allocate to this priority pair, the appropriate sub-carriers $\hat{n}$, in order to achieve his expectations. Indeed, we exhibit in the Sect. 2.3, a joint adaptation algorithm, which allows to rank the available sub-carriers according to their gain value.

Moreover, the process of sub-carriers distribution will be performed according to two steps: one with sub-carrier partitioning and other one without sub-carrier partitioning. The first approach only considers the threshold throughput constraints of services, but don't taking into account in the proportionality of active connections. In this case, the system automatically decides to the partition of resources for each service according to their priorities. Concerning the second method, the parameters α^s are introduced to adjust the proportion of bandwidth of each service. The idea is to avoid that one service, which is always a priority service, don't uses the total bandwidth. Depending on our assumption, which supposes a simultaneous request of three services, a excellent performance in our case, can be assimilated to the reception of the three services. As shown in [5], a good performance is achieved by maximizing the users satisfaction rates and ensuring the fairness among services as well as users.

3.1 Scheduling priority

We defined the priority service $\hat{s}$, requested by the priority user $\hat{k}$, according to the priority functions $P^s_{k,n}(t)$ established in [9] and [3]. In our work, we used these priority functions to classify the pairs $(s; k)$.

– Let us define $\beta^s (s \in \{rtps; nrtps; be\})$ the priority settings of the three services.

• For RTPS connection, we can define the priority function of the user k on the sub-carrier n as follows:
$P^{rtps}_{k,n}(\mu T)$

$$= \begin{cases} \beta^{rtps} . \dfrac{D^{rtps}_{k,n}(\mu T)}{\overline{R}_k(\mu T)} . \dfrac{Q^{rtps}_k(\mu T).B}{\overline{r}_k(\mu T)} . \dfrac{\omega_k}{\tau_k - 2T}, & \omega_k < \tau_k - 2T \\[2ex] \beta^{rtps} . \dfrac{D^{rtps}_{k,n}(\mu T)}{\overline{R}_k(\mu T)} . \dfrac{Q^{rtps}_k(\mu T).B}{\overline{r}_k(\mu T)}, & \omega_k \geq \tau_k - 2T \end{cases}$$

(17)

where B is the bandwidth, $D^{rtps}_{k,n}(\mu T)$ corresponds to the instantaneous data rate of each user k on the sub-carrier n, using the RTPS service, $\overline{R}_k(\mu T)$ is the average data arrival rate of user k, $\overline{r}_k(\mu T)$ is the average data service rate since the service setup, ω_k denotes the longest packet waiting time, τ_k corresponds to the maximum latency which is negotiated when the connection is established.

The term $\frac{\omega_k}{\tau_k - 2T}$ is introduced to process the high latency requirement of the RTPS services. This therm indicates that

the packet should be transmitted immediately if it deadline expires before the next symbol is fully served.

In practice, the user's averaged data rate $\overline{R}_k(\mu T)$ and the average data rate of the service $\overline{r}_k(\mu T)$ can be updated as follows:

$$\overline{R}_k((\mu+1)T)$$
$$= \begin{cases} (1-\theta).\overline{R}_k(\mu T) + \theta.R_k((\mu+1)T), & \overline{R}_k((\mu+1)T) > 0 \\ (1-\theta).\overline{R}_k(\mu T) & \overline{R}_k((\mu+1)T) = 0 \end{cases}$$
(18)

$$\overline{r}_k((\mu+1)T)$$
$$= \begin{cases} (1-\theta).\overline{r}_k(\mu T) + \theta.r_k((\mu+1)T), & \overline{r}_k((\mu+1)T) > 0 \\ (1-\theta).\overline{r}_k(\mu T) & \overline{r}_k((\mu+1)T) = 0 \end{cases}$$
(19)

where $\theta \in [0, 1]$ is the forgetting factor, $r_k(\mu T)$ is the data service rate and $R_k(\mu T) = \sum_{n \in \Omega_k} R_{k,n}(\mu T)$ is the served data rate in the time frame μT, ($R_{k,n}(\mu T)$ corresponds to the data rate on sub-carrier n of the user k; it is equaled to 0, if sub-carrier n isn't assigned to user k).

– For NRTPS connection, we can define the priority function of the user k on the sub-carrier n as follows:

$$P_{k,n}^{nrtps}(\mu T)$$
$$= \begin{cases} \beta^{nrtps}.\dfrac{D_{k,n}^{nrtps}(\mu T)}{\overline{R}_k(t)}.\dfrac{Q_k^{nrtps}(\mu T).B}{\overline{r}_k(\mu T)}.\dfrac{Q_k^{nrtps}(\mu T)}{\psi_k}, & Q_k(\mu T) \geq \psi_k \\[2mm] \beta^{nrtps}.\dfrac{D_{k,n}^{nrtps}(\mu T)}{\overline{R}_k(\mu T)}.\dfrac{Q_k^{nrtps}(\mu T).B}{\overline{r}_k(\mu T)}, & Q_k^{nrtps}(\mu T) < \psi_k \end{cases}$$
(20)

where $D_{k,n}^{nrtps}(\mu T)$ corresponds to the instantaneous data rate of each user k on the sub-carrier n, using the NRTPS service, ψ_k is the maximum number of packets in the correspondent queue.

The term $\frac{Q_k(\mu T)}{\psi_k}$ is introduced to manage the data rate requirement of NRTPS services. Large value of this term indicates that the buffer is full, so the packets of the user k should be given high priority and served earlier.

• For BE connection, we can define the priority function of the user k on the sub-carrier n as follows:

$$P_{k,n}^{be}(\mu T) = \beta^{be}.\dfrac{D_{k,n}^{be}(\mu T)}{\overline{R}_k(\mu T)}.\dfrac{Q_k^{be}(\mu T).B}{\overline{r}_k(\mu T)}$$
(21)

where $D_{k,n}^{be}(\mu T)$ is to the instantaneous data rate of each user k on the sub-carrier n, using the BE service.

For this class of service, no real-time constraint and no bandwidth guarantee is specified. However, it is necessary to consider data loss due to the overflow of the buffer. We chose a value of β^{be}, which gives a minimum priority to BE connections.

For each sub-carrier n, there are k instantaneous data rates, which can be used by the users. However, a sub-carrier n is assigned to only one user during the frame time μT. According to the computation of the priority function of each service s and for each user k, we can define the priority pair per sub-carrier $(\hat{s}; \hat{k})_n$ at the frame time μT as follows:

$$(\hat{s}; \hat{k})_n = arg \max_{s \in \{rtps, nrtps, be\}; k \in \{1, \ldots, K\}} (P_{k,n}^s(\mu T))$$
(22)

3.2 Sub-carriers allocation without moderated compartmentalization

Once the priority pair $(\hat{s}; \hat{k})_n$ is found, we assign the appropriate sub-carrier through the assignment parameter $\delta_{\hat{k},n}^{\hat{s}}(\mu T)$ defined at each frame time μT by:

$$\delta_{\hat{k},n}^{\hat{s}}(\mu T) = \begin{cases} 1, & \text{if assignment} \\ 0, & \text{else} \end{cases}$$
(23)

Our principle of assignment aims to give a preference to the appropriate sub-carrier in terms of robustness to the priority pair $(\hat{s}; \hat{k})_n$. Indeed, with the joint parametrization, we set the system to obtain the best robust channel gains, meeting the required constraints defined in Sect. 2.2.1. Afterward, for each frame time μT, we assign to the priority pair $(\hat{s}; \hat{k})_n$, the sub-carrier maximizes the gain of channel after precoding such as:

$$\hat{n} = arg \max_{n \in \{1, \ldots, N\}} \left(\delta_{\hat{k},n}^{\hat{s}}(\mu T).\sum_{sc=1}^{b} \left(\gamma_{\hat{k},n,sc}^{\hat{s}} \right) \right)$$
(24)

After that, we can compute the instantaneous data rate $D_{\hat{k},\hat{n}}^{\hat{s}}(\mu T)$ that can be received by the priority pair $(\hat{s}; \hat{k})_{\hat{n}}$ on the appropriate sub-carrier $\hat{n}$ as follows:

$$D_{\hat{k},\hat{n}}^{\hat{s}}(\mu T) = \delta_{\hat{k},\hat{n}}^{\hat{s}}(\mu T) \sum_{sc=1}^{b} \log_2 \left(1 + \gamma_{\hat{k},\hat{n},sc}^{\hat{s}} \right).\dfrac{B}{N}$$
(25)

As long as the data rate $D_{\hat{k}}^{\hat{s}}(\mu T)$ of the priority pair $(\hat{s}; \hat{k})$ is inferior than the threshold rate constraint $D_{th}^{\hat{s}}$, the obtained throughput together with another priority sub-carrier will be added to the new priority pair throughput, until this new data rate is superior or equal to the threshold rate constraint. Thus, a set of appropriate sub-carriers, denoted $\Omega_{\hat{k}}^{\hat{s}}$, will be allocated to the priority pair in order to respect the threshold rate constraint.

$$D_{th}^{\hat{s}} \leq D_{\hat{k}}^{\hat{s}}(\mu T) \quad \text{with} \quad D_{\hat{k}}^{\hat{s}}(\mu T) = \sum_{\hat{n} \in \Omega_{\hat{k}}^{\hat{s}}} D_{\hat{k},\hat{n}}^{\hat{s}}(\mu T)$$
(26)

In the Algorithm 2, we describe our radio resource allocation without sub-carrier partitioning, which is based on the cross layer QoS-guaranteed scheduling scheme.

The main idea of this algorithm, is to classify a lot of sub-carriers, according at first to their gain value, and secondly, allocating the best sub-carriers $\hat{n}$ to the priority pair $(\hat{s}; \hat{k})_{\hat{n}}$ together with the use of the priority functions.

The challenge of this strategy is to find the best number of sub-carriers used by each service, to allow the guarantee of QoS users. Thus, we introduce the partitioning parameter α^s, which represents the proportion of available service s sub-carriers. Thereby, we subdivide the bandwidth in

---

Algorithm 2 - Resource allocation without sub-carriers partitioning

---

Input :
- $N_{free} = N$: N_{free} is the number of unallocated sub-carriers ;
- $\delta_{k,n}^s(\mu T) = 0$: assignment parameter ;
- $(\omega_{k,n,sc}^2)^s$: precoder coefficient as in Eq. (13) ;
- D_{th}^s : service threshold rate constraint as in Eq. (15) ;
- $Pr_{k,n}^s(\mu T) = 0$: priority function of service s ;

Operation :
1.   **for each service** $s \in \{rtps; nrtps; be\}$ **do**
2.   **for each user** $k \in K$ **do**
3.   **step 1 :** Use the dMMPP model to simulate the users queue $Q_k^s(\mu T)$;
4.   **for each sub-carrier** $n \in N$ **do**
5.   **step 2 :** Compute the gain of channel : $\gamma_{k,n}^s = \sum_{sc=1}^{b} [(\omega_{k,n,sc}^2)^s . \sigma_{k,n,sc}^2]$;
6.   **step 3 :** Compute the instantaneous data rate : $D_{k,n}^s(\mu T) = \frac{B}{N}.\log_2(1 + \gamma_{k,n}^s)$;
7.   **step 4 :** Compute the priority function : $Pr_{k,n}^s(\mu T)$ according to
                       equations (17), (20) and (21);
8.   **end**
9.   **end**
10. **end**
11. **while** ($N_{free} \neq 0$) **do**
12.   **step 5 :** Select the priority pair per sub-carrier $(\hat{s}; \hat{k})_n$ by Eq. (22);
13.   **step 6 :** Select the appropriate sub-carrier $\hat{n}$ by Eq. (24) and update $\delta_{\hat{k},\hat{n}}^{\hat{s}}(\mu T) = 1$;
14.   **step 7 :** Compute the instantaneous throughput received $D_{\hat{k},\hat{n}}^{\hat{s}}(\mu T)$ by Eq. (25);
15.   **step 8 :** Compute user throughput $D_{\hat{k}}^{\hat{s}}(\mu T) = D_{\hat{k}}^{\hat{s}}(\mu T) + D_{\hat{k},\hat{n}}^{\hat{s}}(\mu T)$;
16.   **step 9 :** Update : $\Omega_{\hat{k}}^{\hat{s}} = \Omega_{\hat{k}}^{\hat{s}} \cup \{\hat{n}\};$ $N_{free} = N_{free} - 1;$
                    $$Q_{\hat{k}}^{\hat{s}}((\mu+1)T) = Q_{\hat{k}}^{\hat{s}}(\mu T) + m_{\hat{k}}^{\hat{s}}(\mu T) - \frac{D_{\hat{k}}^{\hat{s}}(\mu T)}{nb^{\hat{s}}};$$
17.   **While** ($D_{th}^s \leq D_{\hat{k}}^{\hat{s}}(\mu T)$) **do**
18.   **step 10 :** Search free sub-carriers $\hat{n}_{free}$ and update $\delta_{\hat{k},\hat{n}_{free}}^{\hat{s}}(\mu T) = 1$;
19.   **step 11 :** Compute the instantaneous throughput received $D_{\hat{k},\hat{n}_{free}}^{\hat{s}}(\mu T)$ by Eq. (25);
20.   **step 12 :** Update user throughput $D_{\hat{k}}^{\hat{s}}(\mu T) = D_{\hat{k}}^{\hat{s}}(\mu T) + D_{\hat{k},n_{free}}^{\hat{s}}(\mu T)$;
21.   **step 13 :** Update : $\Omega_{\hat{k}}^{\hat{s}} = \Omega_{\hat{k}}^{\hat{s}} \cup \{\hat{n}_{free}\};$ $N_{free} = N_{free} - 1;$
                    $$Q_{\hat{k}}^{\hat{s}}((\mu+1)T) = Q_{\hat{k}}^{\hat{s}}(\mu T) + m_{\hat{k}}^{\hat{s}}(\mu T) - \frac{D_{\hat{k}}^{\hat{s}}(\mu T)}{nb^{\hat{s}}};$$
22.   **end**
23. **end**

3.3 Sub-carriers allocation with moderated compartmentalization

In this section, we subdivide the total bandwidth, in order to establish some parts of sub-carriers which are associated to the considered service type. The objective is to carry out a moderated repartition of sub-carriers between services, in order to avoid that one service, which is always a priority service, doesn't use all available resources.

three parts, then, we process to the same sub-carrier allocation scheme, proposed in Sect. 3.2.

Let us define N^{rtps}, N^{nrtps} et N^{be}, respectively corresponding to the maximum number of sub-carrier of RTPS, NRTPS and BE service. The sub-carrier number of a service s would be defined as follows:

$$N^{rtps} = \alpha^{rtps}N, \quad N^{nrtps} = \alpha^{nrtps}N \quad \text{and} \quad N^{be} = \alpha^{be}N \qquad (27)$$

In the Algorithm 3, we describe our radio resource allocation with sub-carrier partitioning. At first, we

compute the total number of sub-carriers for each class of service. After that, we use the same steps in the algorithm 2 to allocate the sub-carriers to the users.

NRTPS and BE services. This MS MU MIMO–OFDMA system has been set to work on 5 MHz bandwidth with 512 FFT sub-carriers. A frequency selective Rayleigh fading

---

Algorithm 3 - Resource allocation with sub-carriers partitioning

Input :
- α^s : is the percentage of sub-carriers reserved to the service s ;
- N^s : is the number of sub-carriers per service s ;
- $N_{free} = N$: is the total number of unallocated sub-carriers ;
- N_{free}^s : is the number of unallocated sub-carriers in the service s proportion ;
- $\delta_{k,n}^s(\mu T) = 0$: assignment parameter ;
- $(\omega_{k,n,sc}^2)^s$: precoder coefficient, as in Eq (13) ;
- D_{th}^s : service threshold rate constraint, as in Eq (15) ;
- $Pr_{k,n}^s(\mu T) = 0$: priority function of service s ;

Operation :
1.   **for each service** $s \in \{rtps; nrtps; be\}$ **do**
2.   **for each user** $k \in K$ **do**
3.   **step 1 :** Compute the sub-carrier number per service N^s, as in Eq (27) and
                 update $N_{free}^s = N^s$;
4.   **step 2 :** Use the dMMPP model to simulate the users queue $Q_k^s(\mu T)$
5.   **for each sub-carrier** $n \in N$ **do**
6.   **step 3 :** Compute the gain of channel : $\gamma_{k,n}^s = \sum_{sc=1}^b [(\omega_{k,n,sc}^2)^s . \sigma_{k,n,sc}^2]$
7.   **step 4 :** Compute the instantaneous data rate : $D_{k,n}^s(\mu T) = \frac{B}{N} . \log_2(1 + \gamma_{k,n}^s)$;
8.   **step 5 :** Compute the priority function : $Pr_{k,n}^s(\mu T)$ according to
                    Equations (17), (20) and (21);
9.   **end**
10.   **end**
11.   **end**
12.   **while** ($N_{free} \neq 0$) **do**
13.   **step 6 :** Select the priority pair per sub-carrier $(\hat{s}; \hat{k})_n$ by Eq. (22)
14.   **step 7 :** Select the appropriate sub-carrier $\hat{n}$ by Eq. (24) and update $\delta_{k,\hat{n}}^{\hat{s}}(\mu T) = 1$;
15.   **step 8 :** Compute the instantaneous throughput received $D_{\hat{k},\hat{n}}^{\hat{s}}(\mu T)$ by Eq. (25);
16.   **step 9 :** Compute user throughput $D_{\hat{k}}^{\hat{s}}(\mu T) = D_{\hat{k}}^{\hat{s}}(\mu T) + D_{\hat{k},\hat{n}}^{\hat{s}}(\mu T)$
17.   **step 10 :** Update : $\Omega_{\hat{k}}^{\hat{s}} = \Omega_{\hat{k}}^{\hat{s}} \cup \{\hat{n}\}$; $N_{free}^{\hat{s}} = N_{free}^{\hat{s}} - 1$;
                 $$Q_{\hat{k}}^{\hat{s}}((\mu + 1)T) = Q_{\hat{k}}^{\hat{s}}(\mu T) + m_{\hat{k}}^{\hat{s}}(\mu T) - \frac{D_{\hat{k}}^{\hat{s}}(\mu T)}{nb^{\hat{s}}};$$
18.   **If** ($N_{free}^{\hat{s}} > 0$)
19.   **While** ($D_{th}^s \leq D_{\hat{k}}^{\hat{s}}(\mu T)$) **do**
20.   **Step 11 :** Search free sub-carriers $\hat{n}_{free}$ and update $\delta_{k,\hat{n}_{free}}^{\hat{s}}(\mu T) = 1$;
21.   **step 12 :** Compute the instantaneous throughput received $D_{\hat{k},\hat{n}_{free}}^{\hat{s}}(\mu T)$ by Eq. (25);
22.   **step 13 :** Update user throughput $D_{\hat{k}}^{\hat{s}}(\mu T) = D_{\hat{k}}^{\hat{s}}(\mu T) + D_{\hat{k},n_{free}}^{\hat{s}}(\mu T)$;
23.   **step 14 :** Update : $\Omega_{\hat{k}}^{\hat{s}} = \Omega_{\hat{k}}^{\hat{s}} \cup \{\hat{n}_{free}\}$;
                 $$N_{free}^{\hat{s}} = N_{free}^{\hat{s}} - 1; \quad Q_{\hat{k}}^{\hat{s}}((\mu+1)T) = Q_{\hat{k}}^{\hat{s}}(\mu T) + m_{\hat{k}}^{\hat{s}}(\mu T) - \frac{D_{\hat{k}}^{\hat{s}}(\mu T)}{nb^{\hat{s}}};$$
24.   **end**
25.   **end**
26.   **end**

---

4 Numerical results and discussions

4.1 Simulation context

We consider a MIMO–OFDMA system with a single BS and multiples users, requesting at the same time RTPS,

propagation model was used to simulate a MIMO (4×4) channel (with $M_T = 4$ and $M_R = 4$). We assume that, all users have independent time-varying fading MIMO channel characteristics. The chosen carrier frequency is 2.5 GHz and the signal-to-noise ratio has been set to 5 dB. This SNR value assumes a low gain of the channel which

Table 1 Services parameters

Services	ber^s/D_{th}^s (bps)	nb_p^s (bits)	L^s (packets)
RTPS	$10^{-6}/20{,}459$	1024	200
NRTPS	$10^{-9}/81{,}919$	2048	200
BE	$10^{-12}/122{,}870$	4096	500

improves the resource allocation process. Indeed, the lower the gain of channel, the more necessary it is to develop an optimal resource allocation scheme. The radio channel is assumed to be stationary, during the frame time $T = 5$ ms. We use a punctured RCPC code with five coding rates $r \in \{\frac{4}{5} \frac{2}{3} \frac{1}{2} \frac{1}{3} \frac{1}{4}\}$ and a M-QAM modulation with three spectral efficiency levels $M \in \{4, 16, 64\}$. To evaluate the performance of various QoS service, we consider 100 RTPS users, 100 NRTPS users and 100 BE users. For each user, a two-phase dMMPP model is considered for packets arrival with parameters set as matrix $U = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}$ and matrix $\lambda = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$. This setting requires a number of packets for users which is inferior than the buffer size, in order to avoid packet loss. Other system simulation parameters are given in the Table 1. The whole of the proposed system is simulated by SCILAB simulation tool.

4.2 Performance evaluation

The proposed radio resource allocation scheme is performed according to two distinguish approaches. The first one is based on a sub-carriers allocation without partitioning of the bandwidth and the second one consists to subdivide the total bandwidth in three parts. Each part is associated to a given class of service among the RTPS, NRTPS or BE service. Afterwards, these two approaches are simulated according to two different configurations, denoted by config(A) and config(B). In the config(A), we only consider for the PHY/MAC cross layer design, the power allocation strategy developed in [32] and [14], in which the main objective is to decrease the total transmit power effectively. We adapt this power allocation scheme with a precoding design and in a multi-service context. For our proposed scheme (config(B)), a link adaptation scheme is considered. In fact, we adopt a joint parametrization algorithm based on adaptive UPA and UEP techniques. These two configurations (config(A) and config(B)) are associated with our QoS-guaranteed scheduling algorithm in order to distribute the available sub-carriers to the users who simultaneously demand the three services. Then, we compare the performance of the proposed QoS-guaranteed scheduling algorithm with another conventional scheduling

algorithm. This scheduler, proposed for OFDMA systems, is the maximum SNR (Max-SNR). In this technique, users are only selected over each sub-channel according to their CSI. In fact, the priority functions only depend on the users data rate.

The different approaches are evaluated according to three principal criteria. The first criterion corresponds to the ratio between the number of served user by a given service and the total number of users requesting this service. This first criterion is called *User Satisfaction Rate*, and is denoted by *USR*.

Let us define USR^s, the users satisfaction rate per service s; $Kser^s$, the number of served users by the service s and K^s, the total number of users which request the service s. The users satisfaction rate per service USR^s can be formulated as follows:

$$USR^s = \frac{Kser^s * 100\,\%}{K^s} \tag{28}$$

Further, the parameters $Kser_{max}^s$ and $Kser_{min}^s$ are respectively identified such as the maximum and the minimum number of users, which can be served with a service s. Considering this equation, when all users are served (that means that: $Kser_{max}^s = K^s$), the term $\frac{Kser_{max}^s}{K^s}$ can be simplified, and the users satisfaction rate USR^s is equal to 100 %. Otherwise, when there are no served users ($Kser_{min}^s = 0$), the term $\frac{Kser_{min}^s}{K^s}$ is equal to 0 and the users satisfaction rate value is 0 %.

The second criterion is the *Average Throughput* of the served users, which represents the average quantity of informations, received by users. This second criterion gives us some indications about the global performances of the studied scheme.

The third criterion represents the effective bit error rate per service $eBER^s$. Indeed, the proposed system is jointly parameterized in order to ensure a service QoS requirement in term of target bit error rate (ber^s). This criterion allows to verify if the service QoS constraints are achieved.

Furthermore, we also evaluate the impact of the service parameter β^s over our algorithms.

4.2.1 Results without sub-carriers compartmentalization

Remember that the proposed QoS-guaranteed scheduling algorithm uses the priority functions to classify the pairs (service s; user k), and these functions are characterized by the priority setting β^s. This parameter is very interesting, because it allows to adjust the priority between services. Generally, a high priority is assigned to real time services and low priority to the best effort services. So, β^{rtps} is usually equal to 1 (the maximum value), and the other parameters, which are β^{nrtps} and β^{BE} are less than 1. However, no rules are

defined for the choice of these parameters. Therefore, we proposed to vary the parameters β^{nrtps} in order to evaluate its impact in the proposed scheme.

Table 2 represents the rate of sub-carrier resources for each service, according to β^{nrtps} values. We varied the parameter β^{nrtps} from 0.1 to 0.9. This parameter must always be lower than the parameter β^{rtps} and upper than the parameter β^{be}. For that reason, we set the parameters $\beta^{rtps} = 1$ and $\beta^{be} = 0.1$. The number of users K is equal to 95. This number of users assumes a maximum rate satisfaction of the service RTPS in the two configurations. In fact, 85 % of RTPS connections, 68 % of NRTPS connections and 0 % of BE connections are satisfied in config Ⓐ, and in config Ⓑ, 100 % of RTPS connections 100 % of NRTPS connections and 0 % of BE connection are ensured.

We can notice as well, in config Ⓐ, the greater the value of β^{nrtps}, the higher the NRTPS service proportion. Moreover, for the great values of β^{nrtps} (upper to 0.5), the rate of RTPS connections considerably decreases, while the rate of NRTPS connections increases. In this case, we see that resources are not being efficiency distribution because the priority service (RTPS) isn't any more favored. In config Ⓑ, the proportion of each service remains almost constant for all β^{nrtps} values. The proposed scheme ensures a good distribution of the available resources in order to avoid the excessive use of these resources. We also note that, no BE connection is accepted. The resource isn't sufficient to satisfy all demands of connection. Thus, the system is overloaded. Therefore, we can reduce the number of users (new value of K is equal to 80), in order to evaluate the influence of the parameter β^{nrtps}, when congestion occurs.

Table 3 shows the rate of resources for each service, according to β^{nrtps} values. We pick the same configurations as previously, but only reduce the total number of users ($K = 80$). For config Ⓐ, the NRTPS connection rate grows with the values of the parameter β^{nrtps}, and for the large value (upper to 0.5), the rates are constant. Likewise, the RTPS connection rate decreases with this parameter β^{nrtps},

and when β^{nrtps} is upper to 0.5, the rates are constant. Furthermore, there are still no BE connections for all β^{nrtps} values. However, we can see the available BE connections, when we apply our proposed solution (config Ⓑ). In our method, the rates of RTPS and NRTPS connections are decreased in favor to the BE connections.

Due to the earlier observations, we select for the next simulations, a value of $\beta^{nrtps} = 0.6$. Afterward, we investigate the behavior of the parameter β^{be}.

Table 4 shows the rate of resources for service, according to β^{be}. We set this parameter from 0.1 to 0.6. That is always less than the paramters β^{nrtps} and β^{rtps} ($\beta^{nrtps} = 0.6$ and $\beta^{rtps} = 1$). We can observe that in the two configurations (config Ⓐ and config Ⓑ) the rate of resources remains almost constant. All β^{be} values really ensure a minimum priority of BE services. It confirms that, BE connections will be served only if all RTPS and NRTPS connections have no packets to transmit. The same behaviors is true for the number of users equal to 100 and 80(cf. to the Table 5). Thus, to ensure a minimum priority of service BE, we select the parameter $\beta^{be} = 0.1$.

Figure 4 represents the user satisfaction rate per service USR^s for the two configurations (config Ⓐ and config Ⓑ), according to a total number of users between 30 and 100. The maximum number of served users of each service $Kser^s_{max}$ is very clearly pointed out. This Fig. 4 allows us to compare the performance between config Ⓐ and config Ⓑ. Considering the three services (RTPS, NRTPS and BE service), we can notice that, at most 35 users have a satisfaction rate equal to 100 % with config Ⓐ (dotted lines). While in config Ⓑ (full lines), 100 % of satisfaction rate is obtained for at most 40 users. Furthermore, the user satisfaction per service in config Ⓐ, is generally still lower than in case of config Ⓑ. Indeed, there is at most, 35 users who have received BE services, 75 users who are served by RTPS services and 80 users who have received the NRTPS connections, in config Ⓐ. Concerning config Ⓑ, we can observed at most 40 users for BE connections, 95 users for RTPS connections and 100 users for NRTPS connections. Depending on our assumption which supposes a

Table 2 Impact of service parameters β^{nrtps} with $\beta^{rtps} = 1, \beta^{be} = 0.1$ and $K = 95$

β^{nrtps}	0.1 (%)	0.2 (%)	0.3 (%)	0.4 (%)	0.5 (%)	0.6 (%)	0.7 (%)	0.8 (%)	0.9 (%)
Config Ⓐ									
RTPS	34.96	33	31.24	29.88	7.42	**7.42**	7.42	7.42	7.42
NRTPS	65.03	66.99	68.55	70.11	92.57	**92.57**	92.57	92.57	92.57
BE	0.00	0.00	0.00	0.00	0.00	**0.00**	0.00	0.00	0.00
Config Ⓑ									
RTPS	26.95	25	24.02	22.85	23.43	**24.02**	24.02	24.02	24.02
NRTPS	73.04	75	75.97	77.14	76.55	**75.97**	75.97	75.97	75.97
BE	0.00	0.00	0.00	0.00	0.00	**0.00**	0.00	0.00	0.00

Table 3 Impact of service parameters β^{nrtps} with $\beta^{rtps} = 1$, $\beta^{be} = 0.1$ and $K = 80$

β^{nrtps}	0.1 (%)	0.2 (%)	0.3 (%)	0.4 (%)	0.5 (%)	0.6 (%)	0.7 (%)	0.8 (%)	0.9 (%)
Config Ⓐ									
RTPS	29.10	27.53	25.58	24.41	22.07	**22.07**	22.07	22.07	22.07
NRTPS	70.89	72.46	74.41	75.58	77.92	**77.92**	77.92	77.92	77.92
BE	0.00	0.00	0.00	0.00	0.00	**0.00**	0.00	0.00	0.00
Config Ⓑ									
RTPS	16.01	16.79	16.99	16.79	17.57	**21.09**	21.09	21.09	21.09
NRTPS	65.03	64.64	64.25	64.64	64.64	**63.08**	63.08	63.08	63.08
BE	18.94	18.55	18.75	18.55	17.77	**15.82**	15.82	15.82	15.82

Table 4 Impact of service parameters β^{be} with $\beta^{rtps} = 1$, $\beta^{nrtps} = 0.6$ and $K = 95$

β^{be}	0.1 (%)	0.2 (%)	0.3 (%)	0.4 (%)	0.5 (%)	0.6 (%)
Config Ⓐ						
RTPS	7.42	7.42	7.42	7.42	7.42	7.42
NRTPS	92.57	92.57	92.57	92.57	92.57	92.57
BE	0.00	0.00	0.00	0.00	0.00	0.00
Config Ⓑ						
RTPS	24.02	24.02	24.02	24.02	24.02	24.02
NRTPS	75.97	75.97	75.97	75.97	75.97	75.97
BE	0.00	0.00	0.00	0.00	0.00	0.00

Table 5 Impact of service parameters β^{be} with $\beta^{rtps} = 1$, $\beta^{nrtps} = 0.6$ and $K = 80$

β^{be}	0.1 (%)	0.2 (%)	0.3 (%)	0.4 (%)	0.5 (%)	0.6 (%)
Config Ⓐ						
RTPS	22.07	22.07	22.07	22.07	22.07	22.07
NRTPS	79.29	79.29	79.29	79.29	79.29	79.29
BE	0.00	0.00	0.00	0.00	0.00	0.00
Config Ⓑ						
RTPS	21.09	21.09	21.09	21.09	21.09	21.09
NRTPS	63.08	63.08	63.08	63.08	63.08	63.08
BE	15.82	15.82	15.82	15.82	15.82	15.82

simultaneous demand of three services, the proposed scheme (config Ⓑ) allows to satisfy 5 additional served users in comparison with the config Ⓐ.

Afterward, we were interested in the average throughput of served users, who have a satisfaction rate equal to 100 %. For that reason, the Fig. 5 shows the average data rate per service of users, who together demand the RTPS, NRTPS and BE services. We can notice that, the user QoS requirements are always satisfied for the two configurations. Indeed, in config Ⓐ (in green), the average throughput per service of the 35 served users, is upper than the service threshold rate (in blue), and in config Ⓑ (in red), the average data rate per service of the 40 served users, is also upper than the service threshold rate. However, config Ⓐ uses more data rate than necessary, which leads to a waste of resources. On the other hand, the proposed scheme is able to improve the use of resources, by assigning to the served users, the throughput close to their needs. Thereby, the excess recovered data rate allows to accept other connection demands, which improve the user satisfaction rate. That is achieved thanks to the implementation of the PHY/MAC cross layer design.

In order to bring out better our proposed contribution, the performance results may also show the per-user and per-service throughput. Thus, the Fig. 6 depicts the per-user and per-service throughput of config Ⓐ and config Ⓑ. However, the representation of the per-user and per-service throughput of all served users on the same figure, would be completely illegible. For this reason, we only consider the per-user and per-service throughput of 10 served users, which are randomly selected, following a uniformly distribution.

In order to evaluate the contribution of the integration of our proposed QoS-guaranteed scheduling algorithm, we compare it to the Max-SNR scheduling algorithm. Thus,

Fig. 4 Comparison config(A) versus config(B). User satisfaction rate per service USR^s
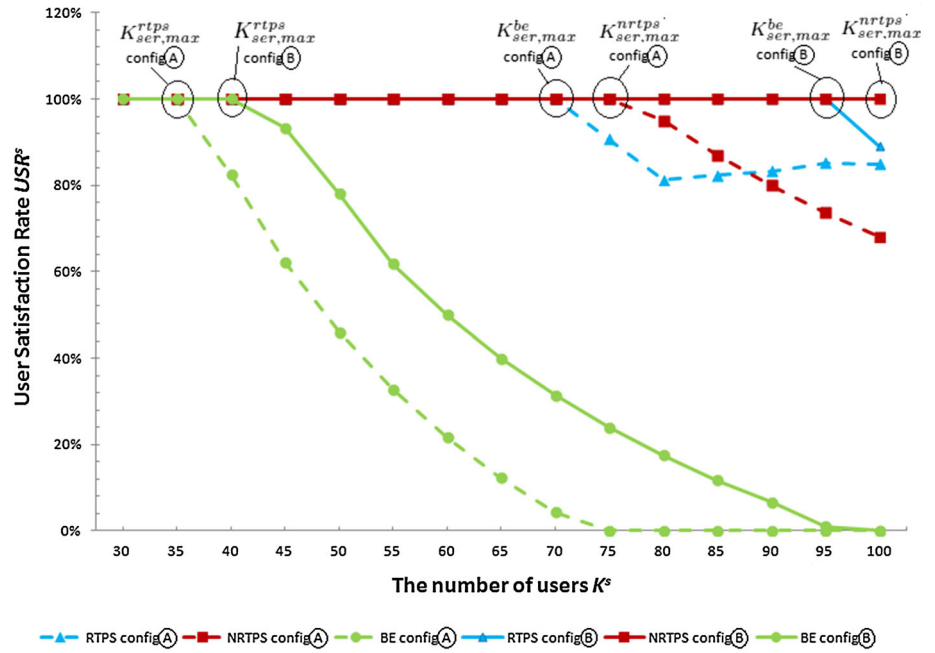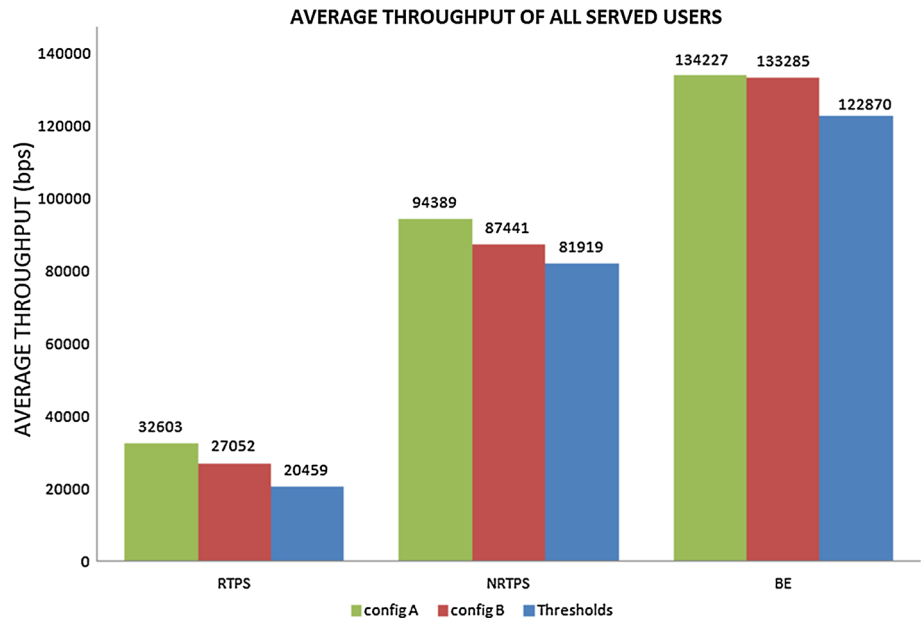

Fig. 5 Average throughput of all served users

the Fig. 7 depicts the users satisfaction rate per service USR^s versus the total number of users for the conventional Max-SNR scheduling scheme (dotted lines), and very clearly points out the maximum number of served users of each service $Kser^s_{max}$. We can observe that, with this conventional scheme at most 15 users have a satisfaction rate equal to 100 %. Recall that, in our proposed scheme (config(B)) (full lines) at most 40 users had a satisfaction rate equal to 100 %. Therefore, config(B) allows to satisfy 25 additional served users in comparison with Max-SNR scheduling algorithm. In addition, the satisfaction rates of

the three services are very close for a number of users between 15 and 100. This behavior clearly shows that, the Max-SNR scheduling scheme isn't adapted to the multi-service context. In fact, this scheduling scheme don't takes into account the service specificities. The Max-SNR scheduling algorithm only considers the CSI. It can be concluded that our proposed QoS-guaranteed scheduling algorithm works well for a mixture of RTPS, NRTPS and BE services.

Figure 8 represents the average throughput of the served users, according to the three services for Max-SNR

Fig. 6 Comparison config(A) versus config(B). The per-user and per service throughput for served users, which are randomly selected
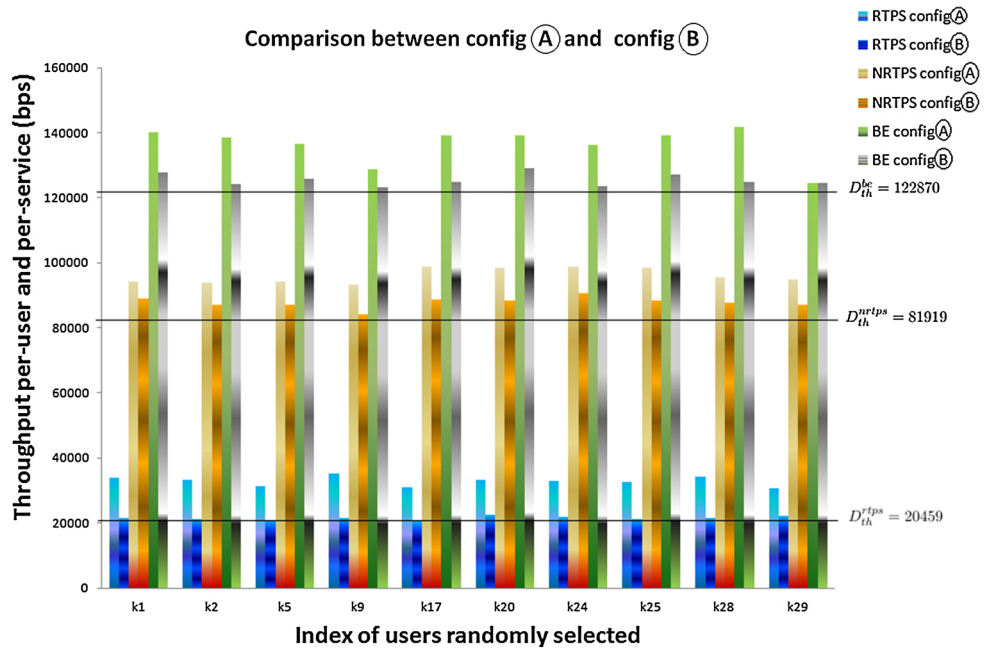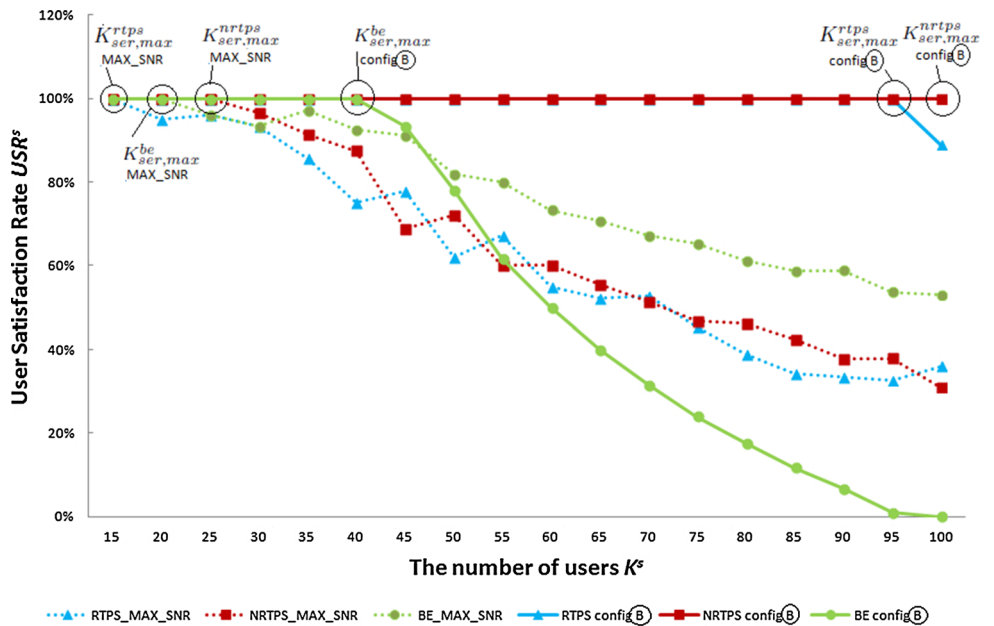

Fig. 7 Comparison Max-SNR scheduler versus config(B). User satisfaction rate per service USR^s

scheduling and our QoS-Guaranteed scheduling algorithms. With these two scheduling algorithms, the user QoS requirements are always satisfied. Indeed, the average throughput of the Max-SNR scheduler (in yellow), is upper than the service threshold rates (in blue), and the average throughput of the proposed QoS-guaranteed scheduler (in red), is also upper than the service threshold rates. However, the average throughput of the Max-SNR method achieves the highest value. This behavior can be explained as follows: in Max-SNR method, the scheduler only selects a connection with the best channel transmission condition.

This scheme don't take into account the QoS constraints, and thereby leads to a significant waste of resources. However, our proposed QoS-guaranteed scheduling algorithm, takes into account the channel conditions and QoS constraints, in order to guarantee the required service QoS (i.e achieve a minimum threshold throughput for each service).

Furthermore, we compare the per-user and per-service throughput between config(B) and Max-SNR scheduler for 10 served users (cf Fig. 9). The results clearly show that the service throughput of each user is upper than each

Fig. 8 Average throughput comparison between Max-SNR scheduler and our QoS-guaranteed scheduler
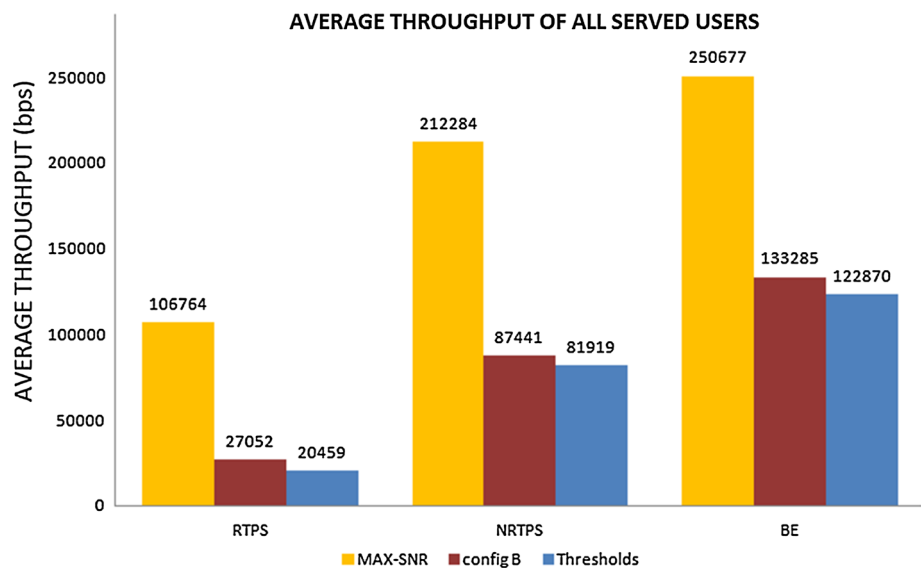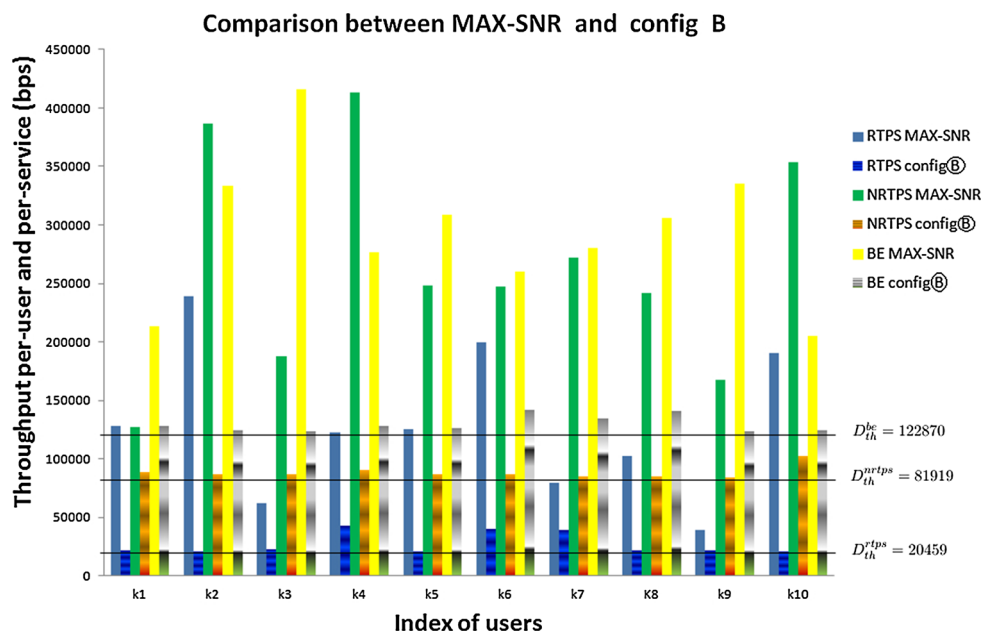

Fig. 9 Comparison Max-SNR scheduler versus config(B). The per-user and per service throughput for ten served users

service threshold rate, and Max-SNR scheduler design uses to many throughput than necessary.

Figures 10 and 11 respectively represent the effective bit error rate per service $eBER^s$ for config(A) and config(B). Indeed, the proposed system is jointly parameterized in order to ensure a QoS constraint for each service, in term of target bit error rate (ber^s). The basic idea consists in assigning the better sub-carrier $\hat{n}$ to the appropriate user $\hat{k}$, which requests the service $\hat{s}$. For this reason, we evaluate the effective bit error rate per service for 5 served users over the both configurations. The results clearly show that service constraints are respected. The effective bit error rate of each service $eBER^s$ is under the target bit error rate ber^s. However, our method (config(B)) improves the using

of resources and satisfied 5 additional served users in comparison to the reference configuration (config(A)), while guaranteeing the same QoS. Likewise, when we consider the effective bit error rate per service $eBER^s$ with Max-SNR configuration, the same conclusions are available. The service QoS constraints are always respected, but this last configuration uses more throughput than necessary. That leads to a waste of resources

4.2.2 Results with sub-carriers compartmentalization

The second approach of our study consists in subdivide the bandwidth in three parts. The idea is to ensure a moderated repartition of sub-carriers between services. Depending on

Fig. 10 Effective bit error rate per service for five users $BERd^s$ in in configⒶ
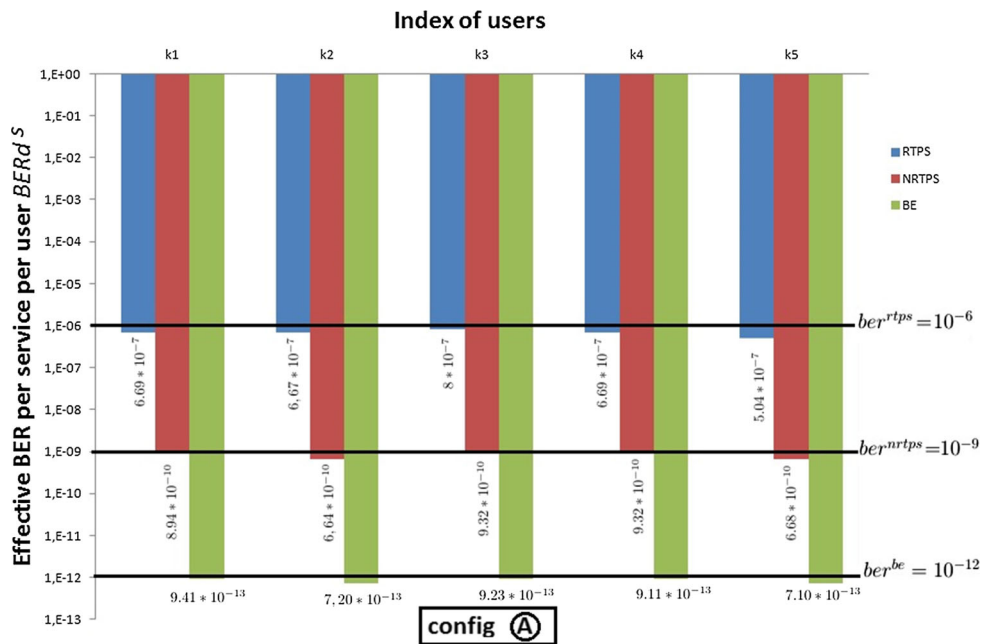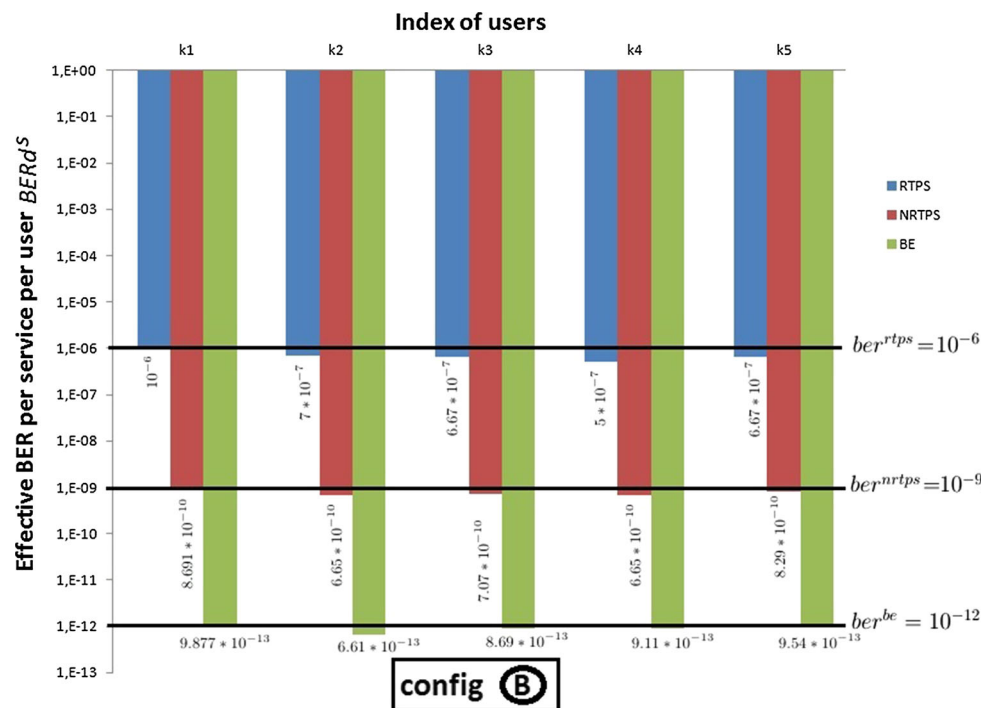

Fig. 11 Effective bit error rate per service for five users $BERd^s$ in in configⒷ

the parameters α^s, we can allocate a maximum number of sub-carriers per service. We select the parameters to a logical way, which consists to achieve the same performance in terms of user satisfaction, in comparison with the approach without sub-carriers compartmentalization. Thus, we can compare the two approaches in term of sub-carriers repartition. The parameters α^s and N^s, which lead to the same performances with the approach without sub-carriers compartmentalization, are given in Table 6.

Figure 12 represents the use satisfaction rate per service USR^s for performance comparison between configⒶ and configⒷ, with bandwidth partitioning. We very clearly indicate the maximum number of served users of each service $Kser^s_{max}$. The results show that in configⒶ (dotted lines) at most 35 users have a satisfaction rate equal to 100 % and in configⒷ (full lines) at most 40 served users are fully satisfied. We really obtained the same performances in term of user satisfaction comparatively to the

Table 6 Values of the parameters α^s and N^s

Services	RTPS	NRTPS	BE
α^s	10 %	34 %	56 %
N^s	≈51 sub-carriers	≈174 sub-carriers	≈287 sub-carriers

approach without sub-carriers compartmentalization. However, we can observe a moderated proportionality of the sub-carriers per service. Indeed, the user satisfactions per service are close. This behavior is more prominent with config(B). This new approach allows to achieve a excellent

Fig. 12 Comparison config(A) versus config(B). User satisfaction rate per service USR^s with $\alpha^{rtps} = 10\%$, $\alpha^{nrtps} = 34\%$, and $\alpha^{be} = 56\%$
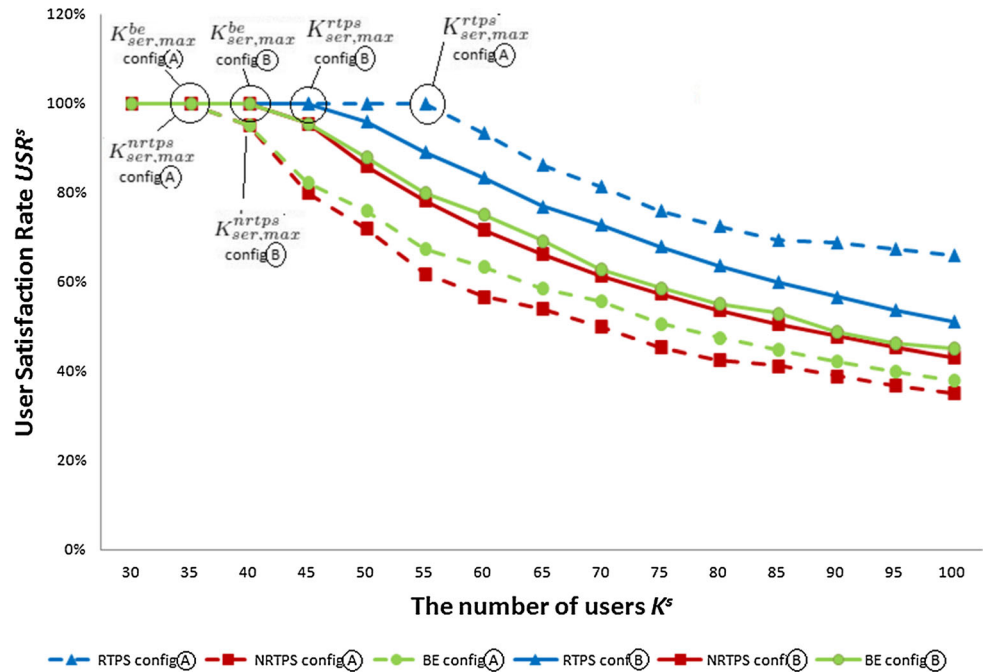

Fig. 13 Average throughput of users with $\alpha^{rtps} = 10\%$, $\alpha^{nrtps} = 34\%$ and $\alpha^{be} = 56\%$
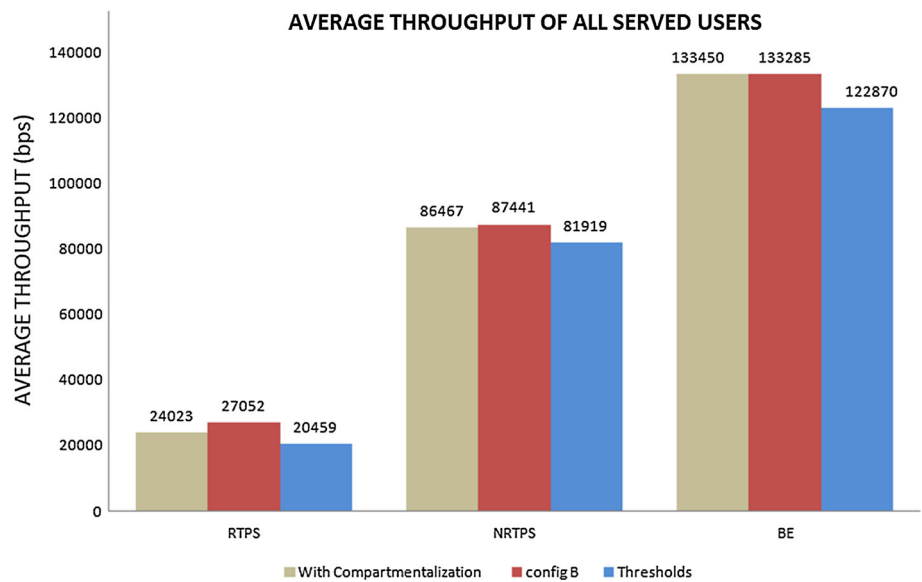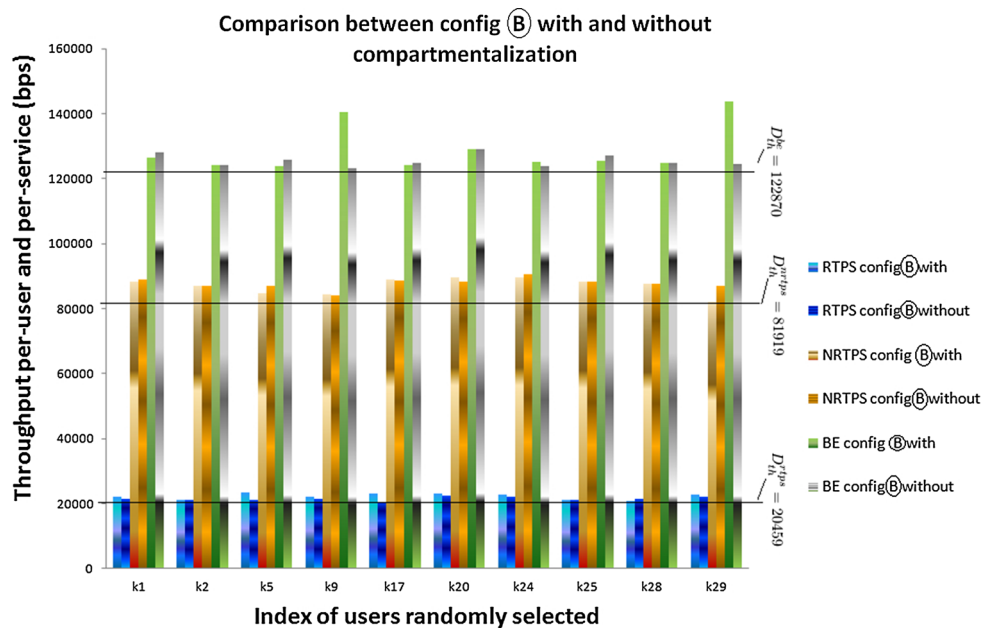
QoS because the user's satisfaction is maximized and a moderated fairness among services is ensured.

Afterward, we presents the results in terms of users average throughput with the parameters $\alpha^{rtps} = 10\,\%$, $\alpha^{nrtps} = 34\,\%$ and $\alpha^{be} = 56\,\%$. In fact, the Figs. 13 and 14 show that the method with compartmentalization, give the same performances in comparison with a method without compartmentalization. However, this moderated repartition of the sub-carriers allows to control the use of resources and therefore, adapt the priority of the services, according to the traffic. In fact, when the congestion of the real-time service occurs, some sub-carriers reserved to the non-real-time service can be borrowed, in order to increase the system performances.

5 Conclusion

In this paper, we have developed a radio resource allocation strategy for multiple connections with different QoS requirements, which can be used in a multi-user MIMO–OFDMA system. First, we have used an unequal power allocation method by the implementation of the precoding scheme. Then we were interested in a joint parameterization of the system elements. Thus, we have establish an efficient and low-complex power allocation to find the optimal joint parameterization. Afterward, we also proposed a QoS-guaranteed scheduling, ensuring a good trade-off between the resource allocation and the respect for QoS users. The priority of each user is based on his channel quality and the priority of his connection. We scheduled the user which has firstly, the highest priority, and allocated

the sub-channel to the connection with the highest QoS priority of that user. The simulation results show clearly an increase of the number of served users, between the proposed PHY/MAC cross-layer scheme and only power allocation technique. Moreover, our proposed QoS-guaranteed scheduling algorithm works well in a multi-service context comparatively to a conventional Max-SNR scheduling algorithm. Indeed, the proposed system optimizes the users data rates and increases the user satisfaction. In order to avoid that a service uses all available resources, we proposed to subdivide to the total bandwidth among classes of service. We have set a logical and manual choice of the partitioning parameters, according to the characteristic of the service s. This last approach provides the same performances as with our proposed system without compartmentalization. However, it allows to control the use of resources.

References

1. Abot, J., Olivier, C., Perrine, C., & Pousset, Y. (2012). A link adaptation scheme optimized for wireless jpeg 2000 transmission over realistic MIMO systems. *Signal Processing: Image Communication, 27*(10), 1066–1078.
2. Abot, J., Perrine, C., Pousset, Y., & Olivier, C. (2013). An unequal power allocation designed for MIMO systems using content characteristics. *International Journal of research and Reviews in Applied Sciences, 16*(1), 16.
3. Afif, M., Sohtsinda, H., Perrine, C., Pousset, Y., & Olivier, C. (2013). A novel cross layer policy based priority management for radio resource allocation in 4G MIMO–OFDMA system (IEEE 802.16m), Trento, Italia. In *IEEE Global Information Infrastructure Symposium* (pp. 1–3).

4. Ahmadi, S. (2010). *Mobile WiMAX: A systems approach to understanding IEEE 802.16 m radio access technology*. London: Academic Press.

5. Ben Hassen, W., & Afif, M. (2012). A gain-computation enhancements resource allocation for heterogeneous service flows in IEEE 802.16 m mobile networks. *International Journal of Digital Multimedia Broadcasting, 2012*, 1–13.

6. Bocquet, W., Hayashi, K., & Sakai, H. (2006). A power allocation scheme for MIMO–OFDM systems. *IEEE international symposium on communications, control and signal processing (ISCSSP)*, Marrakech, Marocco.

7. Busche, H., Vanaev, A., & Rohling, H. (2009). Svd-based MIMO precoding and equalization schemes for realistic channel knowledge: Design criteria and performance evaluation. *Wireless Personal Communications, 48*(3), 347–359.

8. Chan, P. W., & Cheng, R. S.: Reduced-complexity power allocation in zeroforcing MIMO–OFDM downlink system with multiuser diversity. In *International symposium on information theory* (pp. 2320–2324), IEEE.

9. Chen, G., Wu, X., Zhang, W., & Zheng, X. (2011). A cross-layer resource allocation algorithm with dynamic buffer allocation mechanism. *Journal of Networks, 6*(1), 104–111.

10. Collin, L., Berder, O., Rostaing, P., & Burel, G. (2004). Optimal minimum distance-based precoder for MIMO spatial multiplexing systems. *IEEE Transactions on Signal Processing, 52*(3), 617–627.

11. C.R1002-0, G.C. (2004). cdma2000 evaluation methodology, v1.0.

12. Dahlman, E., Parkvall, S., & Skold, J. (2011). *4G: LTE/LTE-advanced for mobile broadband: LTE/LTE-advanced for mobile broadband*. London: Academic press.

13. Guan, Zj, Li, H., Xu, C. Q., Zhou, Xl, & Zhang, W. J. (2009). Adaptive subcarrier allocation for MIMO–OFDMA wireless systems using hungarian method. *Journal of Shanghai University, 13*, 146–149.

14. Habib, A., Baamrani, K. E., & Ouahman, A. (2011). A low-complexity power and bit allocation algorithm for multiuser MIMO–OFDM systems. *International Journal of Computer Science Issues (IJCSI), 8*(3), 596.

15. Hassan, N. U., & Assaad, M. (2009). Low complexity margin adaptive resource allocation in down-link MIMO–OFDMA system. *IEEE Transactions on Wireless Communications, 8*(7), 3365–3371.

16. Ho, W., & Liang, Y. C. (2009). Optimal resource allocation for multiuser MIMO–OFDM systems with user rate constraints. *IEEE Transactions on Vehicular Technology, 58*(3), 1190–1203.

17. Jiang, M., & Hanzo, L. (2007). Multiuser MIMO–OFDM for next-generation wireless systems. *Proceedings of the IEEE, 95*(7), 1430–1469.

18. Kambou, S., Perrine, C., Pousset, Y., Olivier, C., & Mhamdi, M. (2014). Low-Complexity and optimal resource allocation scheme for scalable video transmission over realistic noisy MIMO channels. In: *IEEE international conference on communication systems (ICCS)*, Macau, China (pp. 574–578).

19. Li, Q., Li, G., Lee, W., il Lee, M., Mazzarese, D., Clerckx, B., et al. (2010). MIMO techniques in WiMAX and LTE: A feature overview. *IEEE Communications Magazine, 48*(5), 86–92.

20. Lim, C., Yoo, T., Clerckx, B., Lee, B., & Shim, B. (2013). Recent trend of multiuser MIMO in LTE-advanced. *IEEE Communications Magazine, 51*(3), 127–135.

21. Lin, K. H., Mahmoud, S.S., & Hussain, Z. M. (2005). Adaptive modulation with space-time block coding for MIMO–OFDM systems. In *IEEE international conference on information technology and applications (ICITA)* (vol. 2, pp. 299–304), Sydney, Australia.

22. Liu, L., Chen, R., Geirhofer, S., Sayana, K., Shi, Z., & Zhou, Y. (2012). Downlink MIMO in LTE-advanced: SU-MIMO vs. MU-MIMO. *IEEE Communications Magazine, 50*(2), 140–147.

23. Moretti, M., & Perez-Neira, A. I. (2013). Efficient margin adaptive scheduling for MIMO–OFDMA systems. *IEEE Transactions on Wireless Communications, 12*(1), 278–287.

24. Moretti, M., Sanguinetti, L., & Wang, X. (2015). Resource allocation for power minimization in the downlink of thp-based spatial multiplexing MIMO–OFDMA systems. *IEEE Transactions on Vehicular Technology, 64*(1), 405–411.

25. Nikolaros, I., Zarakovitis, C., Skordoulis, D., Hadjinicolaou, M., & Ni, Q. (2010). Cross-layer design for multiuser OFDMA systems with cooperative game and MMPP queuing considerations. In *IEEE 10th international conference on computer and information technology (CIT)* (pp. 2649–2654), Bradford, UK.

26. Odhah, N. A., Hassan, E. S., Abdelnaby, M., Al-Hanafy, W. E., Dessouky, M. I., Alshebeili, S. A., et al. (2015). Adaptive resource allocation algorithms for multi-user MIMO–OFDM systems. *Wireless Personal Communications, 80*(1), 51–69.

27. Patteti, K., Kishan, R., & Anil, K. (2013). Novel study of scheduling strategies for multi-user MIMO–OFDM systems with limited feedback. *International Journal of Research in Computer and Communication Technology (IJRCCT), 2*(10), 936–942.

28. Prasad, S., Shukla, C., & Chisab, R. (2012). Performance analysis of ofdma in lte. In: *Third international conference on computing communication networking technologies (ICCCNT)* (pp. 1–7), Tamilnadu, India.

29. Sann Maw, M., & Sasase, I. (2008). Resource allocation scheme in MIMO–OFDMA system for user's different data throughput requirements. *IEICE Transactions on Communications, 91*(2), 494–504.

30. Sharma, V., & Lambotharan, S. (2006). Robust multiuser beamformers in MIMO–OFDM systems. In *IEEE international conference on communication systems* (pp. 1–5), Singapore.

31. Srikanth, S., Pandian, M., & Fernando, X. (2012). Orthogonal frequency division multiple access in wimax and lte: A comparison. *IEEE Communications Magazine, 50*(9), 153–161.

32. Sun, Q., Tian, H., Wang, S., Dong, K., & Zhang, P. (2009). Novel resource allocation algorithms for multiuser downlink MIMO–OFDM of FuTURE B3G systems. *Progress in Natural Science, 19*(9), 1141–1146.

33. Tarokh, V., Naguib, A., Seshadri, N., & Calderbank, A. R. (1999). Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths. *IEEE Transactions on Communications, 47*(2), 199–207.

34. Xia, P., Zhou, S., & Giannakis, G. B. (2004). Adaptive MIMO–OFDM based on partial channel state information. *IEEE Transactions on Signal Processing, 52*(1), 202–213.

35. Xiao, S., Xiao, X., Li, B., & Hu, Z. (2005). Adaptive subcarrier allocation for multiuser MIMO–OFDM systems in frequency selective fading channel. In *IEEE—International conference on wireless communications, networking and mobile computing (WiCOM)* (vol. 1, pp. 61–64), Wuhan, China.

36. Yang, F., Zhang, Q., Zhu, W., & Zhang, Y. Q. (2001). An efficient transport scheme for multimedia over wireless internet. In *Proceedings of the international conference on third generation wireless and beyond*, CA, USA.

37. Zhao, J., Guan, X., LI, X., & Zeng, L. (2014). Cross-layer in MIMO–OFDM system with adaptive modulation and coding: Design and analysis. *Chinese Journal of Electronics, 23*(2), 371–376.

38. Zhu, X., & Murch, R. (2002). Performance analysis of maximum likelihood detection in a MIMO antenna system. *IEEE Transactions on Communications, 50*(2), 187–191.

Samy Kambou received the M.S. degree in telecommunication from the University of Poitiers, in 2012. He is currently working toward the Ph.D. degree at the XLIM Laboratory. His research interests are in the areas of communications, signal processing, and networking, with emphasis on cross-layer design, optimal transmission of multimedia, quality of service provisionning, resource allocation and link adaptation.

Clency Perrine received his M.S. degree in telecommunications in 2002 and his Ph.D. in telecommunications and signal processing in 2005 from the University of Rennes 1. His thesis deals with an operational system for images transmission in HF band (3:30 MHz). In September 2006, he joined the XLIM Laboratory from the University of Poitiers (France) as an associate professor. His research interests are focused on Digital Communication Systems and joint source/channel coding schemes for images and video transmission through MIMO wireless channels.

Mériem Afif received the Ph.D., from Telecom ParisTech of Paris-France in Computer Networks and from Higher School of Communication of Tunis-Tunisia (Sup'com)-University of Carthage in information and communication technologies, in 2007. From 2001 to 2003 she was a research engineer in radio mobile networks at the Center of study and research in telecommunication in Tunis, and from 2003 to 2006 she was an engineering teacher in Higher School of Communication of Tunis-Tunisia (Sup'com).

Since 2009 she has served as an associate professor at the National Institute of Applied Sciences and Technology. She was a researcher, permanent member, in Mediatron as a Research lab on Radio Communication and Multimedia Networks. Her research interests include radio resource management, handovers in wireless networks, mobility and QoS management in heterogeneous radio access technologies and network cross-layer modeling.

Yannis Pousset was born in 1971. He received the Ph.D. degree in mobile radio communication from the University of Poitiers, in 1998. Since 2012, he is professor at the University of Poitiers in the department of electrical engineering. He develops its research activities in the XLIM Laboratory. His research interests include the study of adaptive links related to the optimal transmission of multimedia content over realistic spatio-temporal radio channel.

Christian Olivier received the degree in Mathematics (in Signal Processing) in 1979. His major interest was in stochastic process modelling. He was an Associate Professor at the Rouen University in France from 1987 to 1997, where his research area was mathematical tools for image segmentation, pattern analysis and recognition. Then he joined the University of Poitiers as a Professor. His research is always mathematical tools but for information representation and measurement, especially for image, video coding and wireless transmission in a joint coding strategy. At present, C. Olivier is an Emeritus Professor at Poitiers University.