Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2009, Article ID 726495, 14 pages
doi:10.1155/2009/726495

Research Article

A Fair Opportunistic Access Scheme for Multiuser OFDM Wireless Networks

Cédric Gueguen and Sébastien Baey

Université Pierre et Marie Curie (UPMC) - Paris 6, Laboratoire d'Informatique de Paris 6 (LIP6),
104 avenue du Président Kennedy, 75016 Paris, France

Correspondence should be addressed to Cédric Gueguen, cedric.gueguen@lip6.fr

Received 2 July 2008; Revised 13 December 2008; Accepted 1 February 2009

Recommended by Zhu Han

We propose a new access scheme for efficient support of multimedia services in OFDM wireless networks, both in the uplink and in the downlink. This scheme further increases the benefits of opportunistic scheduling by extending this cross-layer technique to higher layers. Access to the medium is granted based on a system of weights that dynamically accounts for both the experienced QoS and the transmission conditions. This new approach enables the full support of multimedia services with the adequate traffic and QoS differentiation while maximizing the system capacity and keeping a special attention on fairness. Performance evaluation shows that the proposed access technique outperforms existing wireless access schemes and demonstrates that choosing between high fairness and high system throughput is no more required.

1. Introduction

Providing mobile multimedia transmission services with an adequate QoS is very challenging. In contrast with wired communications, wireless transmissions are subject to many channel impairments such as path loss, shadowing, and multipath fading [1–4]. These phenomena severely affect the transmission capabilities and in turn the QoS experienced by applications, in terms of data integrity but also in terms of the supplementary delays or packet losses which appear when the effective bit rate at the physical layer is low. The past decades have witnessed intense research efforts on wireless digital communications. Among all the studied transmission techniques, IOrthogonal Frequency Division Multiplexing (OFDM) has clearly emerged for future broadband wireless multimedia networks (4G systems) and is already widely implemented in most recent wireless systems like 802.11a/g or 802.16. The basic principle of OFDM for fighting the effects of multipath propagation is to subdivide the available channel bandwidth in subfrequency bands of width inferior to the coherence bandwidth of the channel (inverse of the delay spread). The transmission of a high-speed signal on a broadband frequency selective channel is then substituted with the transmission on multiple subcarriers of slow speed signals which are very resistant to intersymbol interference and subject to flat fading. This subdivision of the overall bandwidth in multiple channels provides frequency diversity which added to time, and multiuser diversity may result in a very spectrally efficient system subject to an adequate scheduling.

MAC protocols currently used in wireless local area networks were originally and primarily designed in the wired local area network context. However, conventional access methods like Round Robin (RR) and Random Access (RA) are not well adapted to the wireless environment and provide poor throughput. Much interest has recently been given to the design of scheduling algorithms that maximize the performance of multiuser OFDM systems. Opportunistic scheduling techniques take advantage of multiuser diversity by preferably allocating the resources to the active mobile(s) with the most favourable channel conditions at a given time. This technique was explored first in single carrier communications [5]. More recently, opportunistic scheduling has been exploited in multicarrier systems [6, 7]. These schemes are derived from the maximum signal-to-noise ratio (MaxSNR), also known as maximum carrier

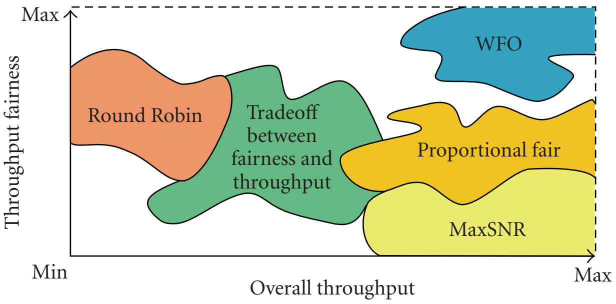


Figure 1: Tradeoff between overall throughput and fairness.

to interference ratio (MaxC/I), technique which allocates the resource at a given time to the active mobile with the greatest SNR. Dynamically adapting the modulation and coding allows then to always make the most efficient use of the radio resource and come closer to the Shannon limit. This maximizes the system capacity of an information theory point of view. However, it assumes that the user with the most favourable transmission conditions has information to transmit at the considered time instant. It does not take into account the variability of the traffic and the queuing aspects.

Pure opportunistic scheduling does not take into account the delay constraints of the flows to convey and suffer of a lack of fairness. References [8, 9] introduce opportunistic schemes coupled with a system of quota. This improves fairness but reduces the efficiency of utilization of the multiuser diversity with prejudice on system throughput. Proportional fair (PF) algorithms have recently been proposed to incorporate a certain level of fairness while keeping the benefits of multiuser diversity [10–14]. The basic principle is to allocate resources to a user, when its channel conditions are the most favourable with respect to its time average. In these schemes, fairness consists in guaranteeing an equal share of the total available bandwidth to each mobile, whatever its position or channel conditions.

However, performance analysis of PF-based protocols has shown that fairness issues persist since these algorithms do not ensure an equal throughput [15, 16]. The main issues are fairness considering mobiles with unequal spatial positioning, different traffic types, or different QoS targets. PF scheduling does not take into account the delay constraints and is not well adapted to multimedia services which introduce heterogeneous users, new traffic patterns with highly variable bit rates and stringent QoS requirements in terms of delay, and packet loss. Recently, [17] proposed the multimedia adaptive OFDM proportional fair (MAOPF) algorithm, an evolution of the classical PF that considers multimedia services. The principle of the MAOPF is to share the total available bandwidth among users proportionally to their bit rate requirement. Although this enables the coexistence of applications with unequal bit rates, heterogeneous QoS requirements are still not well supported. Moreover, the MAOPF allocates all OFDM subcarriers to the same mobile. This does not fully take advantage of the multiuser diversity and has a negative impact on the system capacity.

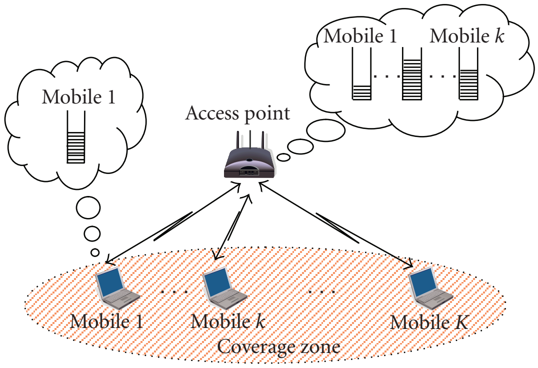


Figure 2: Allocation of radio resources among the set of mobiles situated in the coverage zone of an access point.

This paper proposes a new MAC protocol for efficient support of multimedia services in multiuser OFDM wireless networks. This protocol, which we call the "Weighted Fair Opportunistic (WFO)" protocol, applies cross-layer design concepts taking into account both the OFDM physical layer specificities (transmission conditions) and the higher layer constraints (traffic patterns, QoS constraints). Physical layer information are used in order to take advantage of the time, frequency, and multiuser diversity and maximize the system capacity. Higher layer information are exploited in a weighted system that introduces dynamic priorities between flows for ensuring the same QoS level to all mobiles. This result in an efficient scheme which guarantees the differentiated QoS constraints (data integrity and delay targets) of heterogeneous traffic flows like those generated by multimedia applications. Additionally, this bandwidth management avoids trading capacity for fairness as illustrated in Figure 1.

The paper is organized as follows. Section 2 provides a detailed description of the system under study. Section 3 introduces the QoS management principle embodied in the proposed protocol. Section 4 describes the integrated scheduling algorithm. In Section 5, we present a detailed performance evaluation through a simulation study. Section 6 concludes the paper.

2. System Description

We focus on the proper allocation of radio resources among the set of mobiles situated in the coverage zone of an access point (see Figure 2). We consider a centralized approach. The packets originating from the backhaul network are buffered in the access point which schedules the downlink transmissions. In the uplink, the mobiles signal their traffic backlog to the access point which builds the uplink resource mapping.

We assume that the physical layer is operated using the structure described in Figure 3 which ensures a good compatibility with the OFDM-based transmission mode of the IEEE 802.16-2004 [18, 19]. The total available bandwidth is divided in subfrequency bands or subcarriers. The radio resource is further divided in the time domain in frames.

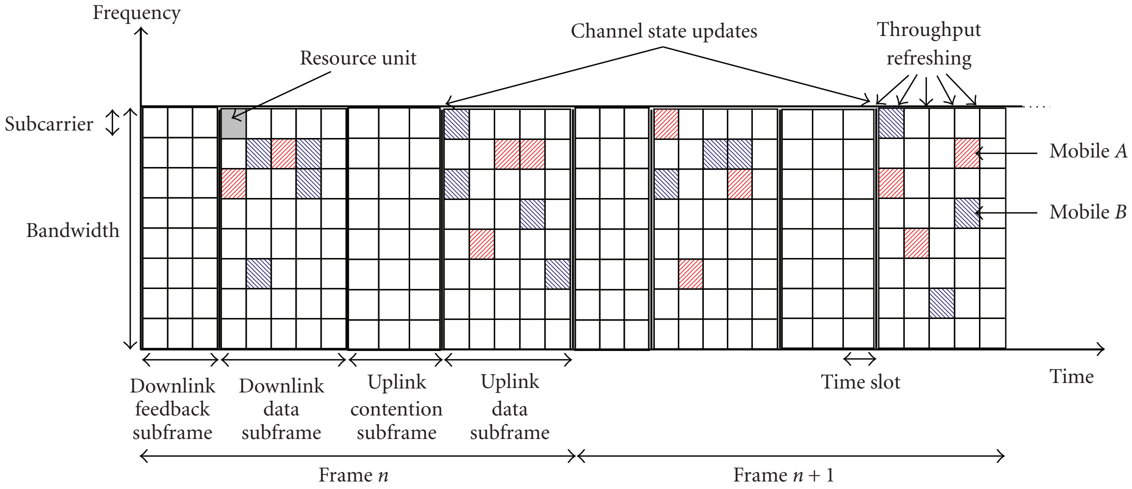


Figure 3: WFO frame structure in TDD mode.

Each frame is itself divided in time slots of constant duration. The time slot duration is an integer multiple of the OFDM symbol duration. The number of subcarriers is chosen so that the width of each subfrequency band is inferior to the coherence bandwidth of the channel. Moreover, the frame duration is fixed to a value much smaller than the coherence time (inverse of the Doppler spread) of the channel. With these assumptions, the transmission on each subcarrier is subject to flat fading with a channel state that can be considered static during each frame.

The elementary resource unit (RU) is defined as any (subcarrier, time slot) pair. Each of these RUs may be allocated to any mobile with a specific modulation order. Transmissions performed on different RUs by different mobiles have independent channel state variations [20]. On each RU, the modulation scheme is QAM with a modulation order adapted to the channel state between the access point and the mobile to which it is allocated. This provides the flexible resource allocation framework required for opportunistic scheduling.

The system is operated using time division duplexing with four subframes: the *downlink feedback subframe*, the *downlink data subframe*, the *uplink contention subframe*, and the *uplink data subframe*. The uplink and downlink data subframes are used for transmission of user data. In the downlink feedback subframe, the access point sends control information towards its mobiles. This control information is used for signalling to each mobile the RU(s) which have been allocated in the next uplink and downlink data subframes, the modulation order selected for each of these RUs and the recommended emission power in the uplink. In the uplink contention subframe, the active mobiles send their current traffic backlog and information elements such as QoS measures and transmit power. The uplink contention subframe is also used by the mobiles for establishing their connections. This frame structure supposes a perfect time and frequency synchronization between the mobiles and the access point as described in [21]. Therefore, each frame starts with a preamble used for synchronization purposes. Additional preambles may also be used in the frame.

3. The WFO Protocol QoS Management Principle

The crucial objective of the WFO protocol is to fully support multimedia transmission services, including the widest range of services: VoIP, videoconference, email, and file transfer. This requires the coexistence of delay sensitive flows as well as non-real-time traffic with looser delay constraints but with tight data integrity targets. In order to deal with the various and heterogeneous QoS requirements of multimedia services, the WFO protocol relies on a generic approach of QoS management.

We define a *service flow* as a traffic stream and its QoS profile, in a given transmission direction. A mobile may have multiple service flows both in the uplink and the downlink. An application may also use several service flows enabling for instance the implementation of Unequal Error Protection schemes in the physical layer. Each service flow possesses its own transmission buffer. In the following, index k is used to designate a given service flow among the set of service flows to be scheduled in a given transmission direction.

The QoS profile is defined as the set of parameters that characterizes the QoS requirements of a service flow mainly in terms of data integrity and delay. In the following, data integrity requirements are specified by a bit error rate (BER) target, which we denote by $BER_{target,k}$ for service flow k. Delay requirements are specified at the packet level. We assume traffic streams are organized at the MAC level in blocks of bits of constant size that we call packets. The packet delay is defined as the time between the arrival of the packet in the transmission buffer and the time of its reception by the mobile or the access point. This delay is roughly equal to the packet waiting time in the service flow transmission buffer neglecting the transmission and propagation delays.

Adequately specifying the delay requirements is challenging. We believe that the meaningful constraint is a limitation of the occurrences of large delay values. By analogy with the concept of outage used in system coverage planning, we define the concept of *delay outage*. A service flow k is in delay outage when its packets experience a delay greater than a given application specific threshold denoted T_k. We define the packet delay outage ratio ($PDOR_k$) experienced by each service flow k as the percentage of packets that do not meet the delay threshold T_k in the total number of packets transmitted. The experienced PDOR value is tracked all along the lifetime of the service flows; at each transmission of a packet of service flow k, the total number of packets whose delay exceeded the delay threshold T_k divided by the total number of packets transmitted since the beginning of the connection is computed. Additionally, we define the packet delay outage ratio target, denoted $PDOR_{target}$, as the maximum ratio of packets that may be delivered after the delay threshold. This characterizes the delay requirements of any service flow in a generic approach. Figure 4 illustrates an example cumulative distribution of the packet delay of service flow k at a given time instant. The objective of the WFO protocol is to regulate the experienced PDOR along the lifetime of the service flow such as its value stays below the PDOR target. This ensures the satisfaction of the delay requirements at a short-time scale.

In the WFO protocol, QoS management is organized in two parts: data integrity management and delay management. Data integrity is guaranteed by the physical layer mainly by adapting the modulation scheme and the transmit power to the mobile specific channel state. This is achieved considering each service flow independently. Delay management is performed considering all service flows jointly and scheduling the packets according to their distance to the PDOR target. Fairness is provided by guaranteeing the same level of satisfaction of delay constraints to all service flows, that is, guaranteeing the same PDOR to all service flows. The joint satisfaction of the delay constraints relies on the dynamics of the traffic streams that are multiplexed. Data integrity and delay management are integrated using the WFO scheduling algorithm.

4. The WFO Scheduling Algorithm

The core of the WFO protocol is its scheduling algorithm. This scheduling is performed during the uplink data transmission phase. The scheduler, located in the access node, grants RUs to each service flow as a function of

(i) its QoS profile (BER target, delay threshold, and PDOR target),

(ii) its currently experienced QoS (BER and PDOR),

(iii) its traffic backlog,

(iv) its channel state.

The QoS profile is signaled in the connection establishment phase. In the uplink, the currently experienced PDOR and the traffic backlog (buffer occupancy) are signaled by the mobile in the contention subframe. The experienced BER is tracked directly by the access node. Reciprocally, in the downlink, the currently experienced PDOR and the traffic backlog are calculated by the access node, and the experienced BER is signaled.

Additionally, knowledge of the channel state is supposed to be available at the receiver [22]. The current channel attenuation on each subcarrier and for each mobile is estimated by the access node based on the SNR of the signal sent by each mobile during the uplink contention subframe. Assuming that the channel state is stable on a scale of 50 milliseconds [23], and using a frame duration of 2 milliseconds, the mobiles will transmit their control information alternatively on each subcarrier so that the access node may refresh the channel state information once every 25 frames.

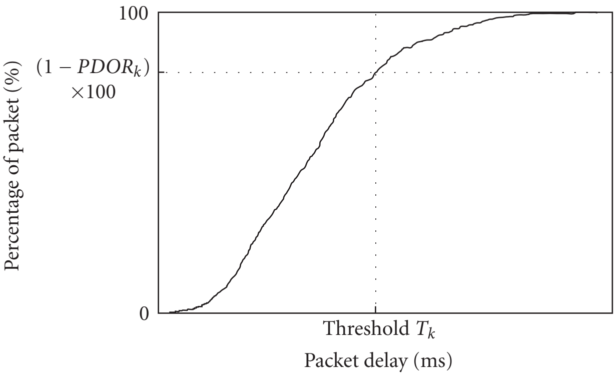


FIGURE 4: An example of packet delay CDF and experienced PDOR.

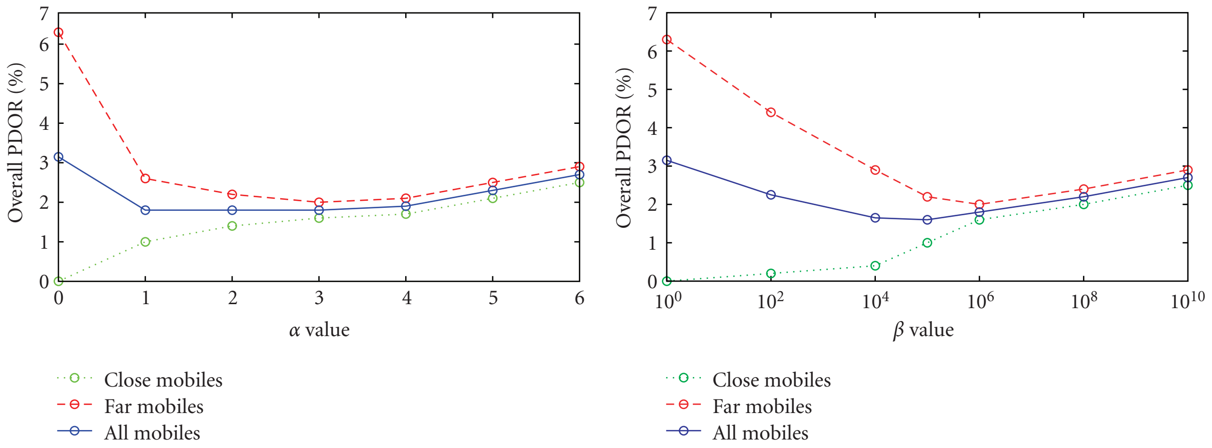
The WFO scheduling algorithm relies on weights that set the dynamic priorities for allocating the resource. These weights are built in order to satisfy two major objectives: system throughput maximization and fairness as explained below.

4.1. System Throughput Maximization. The WFO maximizes the system throughput in a MAC/PHY opportunistic approach. Data integrity requirements of the service flows are enforced considering each service flow independently adapting the modulation scheme and the transmit power to the mobile specific channel state. At each scheduling epoch, the scheduler computes the maximum number of bits $m_{k,n}$ that can be transmitted in a time slot of subcarrier n if assigned to service flow k, for all k and all n. This number of bits is limited by two main factors: the data integrity requirement and the supported modulation orders.

The bit error probability is upper bounded by the symbol error probability [6], and the time slot duration is assumed equal to the duration T_s of an OFDM symbol. The required received power $P_r(q,k)$ for transmitting q bits in an RU while keeping below the data integrity requirement $\text{BER}_{\text{target},k}$ of service flow k is a function of the modulation type, its order, and the single-sided power spectral density of noise N_0. For QAM and a modulation order M on a flat fading channel [1],

$$P_r(q,k) = \frac{2N_0}{3T_s}\left[\operatorname{erfc}^{-1}\left(\frac{\text{BER}_{\text{target},k}}{2}\right)\right]^2 (M-1), \quad (1)$$

where $M = 2^q$, and erfc is the complementary error function. $P_r(q,k)$ may also be determined in practice based on BER history and updated according to information collected on experienced BER.



(a) Exponent parameter calibration

(b) Normalization parameter calibration

Figure 5: α and β calibration.

The transmit power $P_{k,n}$ of service flow k on subcarrier n is upper bounded to a value $P_{\max}$ which complies with the transmit power spectral density regulation:

$$P_{k,n} \leq P_{\max}. \tag{2}$$

Given the channel gain $a_{k,n}$ experienced by service flow k on subcarrier n (including path loss and Rayleigh fading),

$$P_r(q,k) \leq a_{k,n} P_{\max}. \tag{3}$$

Hence, the maximum number of bits $q_{k,n}$ of service flow k which can be transmitted on a time slot of subcarrier n while keeping below its BER target is

$$q_{k,n} \leq \left\lfloor \log_2 \left(1 + \frac{3P_{\max} \times T_s \times a_{k,n}}{2N_0 \left[\operatorname{erfc}^{-1}\left(\mathrm{BER}_{\mathrm{target},k}/2\right)\right]^2} \right) \right\rfloor. \tag{4}$$

We further assume that the supported QAM modulation orders are limited such as q belongs to the set $S = \{0, 2, 4, \ldots, q_{\max}\}$. Hence, the maximum number of bits $m_{k,n}$ that will be transmitted on a time slot of subcarrier n if this RU is allocated to the service flow k is

$$m_{k,n} = \max \{ q \in S,\ q \leq q_{k,n} \}. \tag{5}$$

MaxSNR-based schemes allocate the resources to the flows which have the greatest $m_{k,n}$ values. This bandwidth allocation strategy maximizes the bandwidth usage efficiency but suffers of a significant lack of fairness. In order to provide fairness while preserving the system throughput maximization, a new parameter is introduced which modulates this pure opportunistic resource allocation.

4.2. Fairness Support. The second major objective of the WFO is to provide fairness, that is, guaranteeing the same PDOR to all service flows as explained in Section 3. This is achieved by extending the above cross-layer design to higher layers. A new weighted fair (WF) parameter is introduced based on the current estimation of the PDOR of service flow k:

$$\mathrm{WF}_k = f(\mathrm{PDOR}_k), \tag{6}$$

where f is a strictly positive and monotonically increasing function. The WFO scheduling principle is then to allocate a time slot of subcarrier n to the mobile k which has the greatest WFO parameter value $\mathrm{WFO}_{k,n}$ with

$$\mathrm{WFO}_{k,n} = \mathrm{WF}_k \times m_{k,n}. \tag{7}$$

Based on the PDOR, the WF parameters directly account for the level of satisfaction of the delay constraints for an efficient QoS management. The PDOR is more relevant and simpler to use than the service flow throughput, the buffer occupancy, or the waiting time of each packet to schedule which would introduce a great complexity in the scheduling algorithm. The WFO parameters introduce dynamic priorities that delay the flows which currently easily respect their delay threshold to the benefit of others which go through a critical period.

Our studies on the algorithm performance have shown that a polynomial function f suits well

$$f(x) = 1 + \beta x^{\alpha}. \tag{8}$$

The exponent parameter α allows being more sensitive and reactive to PDOR fluctuations which guarantees fairness at a short-time scale. β is a normalization parameter that ensures that WF_k and $m_{k,n}$ are in the same order of magnitude. Given that PDOR_k has an order of magnitude 10^{-2}, β should be set to $10^{2\alpha}$. With this choice, WF_k is always in the same order of magnitude as $m_{k,n}$ and allows to manage both fairness and system throughput maximization.

By extensive simulations, we analyzed the influence of the value of the pair (α, β) on the performances of the WFO scheduling scheme and adequately tuned $f(x)$. Figures 5(a) and 5(b) illustrate the calibration study. Here, half mobiles are close to the access point and the second half, twice

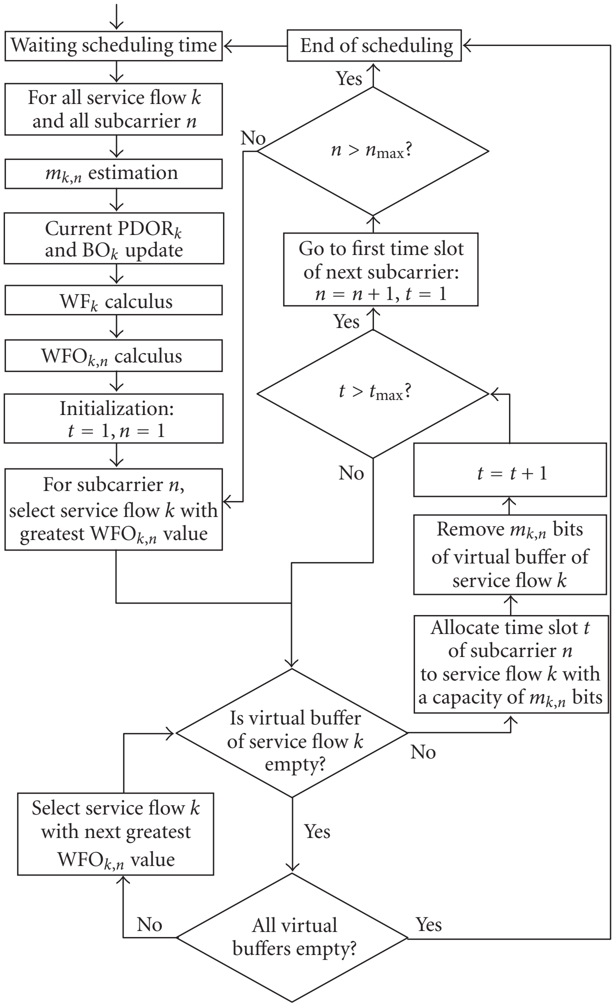


Figure 6: WFO scheduling algorithm flow chart.

other farther. All mobiles run a same application with same delay and BER requirements as described in Section 5.1. Figure 5(a) represents the overall PDOR (computed on all transmitted packets) obtained for different values of α coupled with a β value of $10^{2\alpha}$ as defined above. A cubic exponent suits well offering sufficient reactivity to PDOR fluctuations. Hence, in the following α is assumed to be equal to 3. Figure 5(b) shows the WFO performances obtained for each β value when α is set to 3. It confirms that when β is too small, the weighted parameter has no influence and fairness is lost. On the contrary, if β is too high, $m_{k,n}$ looses weight in the scheduling, and the system throughput maximization decrease. Good values for β range between 10^5 and 10^6. In the following, β is taken equal to 10^6.

Additionally, Figures 5(a) and 5(b) show the potential of the WFO. Indeed, when α or β equals zero, the function f is constant and $m_{k,n}$ only has influence in the scheduling. With this setting, the WFO behaves as the MaxSNR yielding unfair performances. In contrast, the adequate tuning of α and β brings the wanted fairness.

The dynamic priorities introduced by the WFO algorithm evolve as a function of the specific channel conditions and currently experienced QoS of each service flow in a cross-layer higher layers/MAC/PHY approach. This result in a well-balanced resource allocation which keeps a maximum number of service flows active across time but with continuously low traffic backlogs. Preserving this multiuser diversity allows to continuously take a maximal benefit of opportunistic scheduling and thus maximize the bandwidth usage efficiency. Additionally, this also achieves a time uniform fair allocation of the RUs to the service flows ensuring the required short term fairness [24, 25].

4.3. Global WFO Scheduling Algorithm Description. The WFO scheduling algorithm is detailed in Figure 6. The scheduling is run subcarrier by subcarrier and on a time slot basis for improved granularity. In the allocation process of a given time slot, the priority of a service flow with respect to another is determined by the magnitude of its WFO parameter. All service flows are scheduled simultaneously in a single run of the algorithm, whatever their QoS profile is. QoS differentiation is achieved by means of the WFO parameters. Service flows with low delay constraints like best effort traffic are qualified with a quite high delay threshold. As a result, their PDOR is always very small compared to other low latency traffic whose priority increases dramatically as soon their smaller delay threshold is not respected. In the following, we describe the proposed scheduling algorithm step by step.

Step 1. The scheduler refreshes the current PDOR_k and buffer occupancy BO_k values of each service flow k and computes the $m_{k,n}$, WF_k, and $\mathrm{WFO}_{k,n}$ parameters for each service flow and each subcarrier. Then, n and t are initialized to 1.

Step 2. For subcarrier n, the scheduler selects the service flow k with the greatest $\mathrm{WFO}_{k,n}$ value.

Substep 2.1. If the virtual buffer occupancy (we define the virtual buffer occupancy as the current buffer occupancy of service flow k minus the number of bits already allocated to this service flow) of service flow k is positive, the schedulers go to Substep 2.2. Else, if all virtual buffers are null or negative, the scheduler goes to Step 3. Otherwise, the scheduler selects the next service flow k with the greatest $\mathrm{WFO}_{k,n}$ value and restarts Substep 2.1.

Substep 2.2. The scheduler allocates time slot t of subcarrier n to service flow k with a capacity $m_{k,n}$ bits, removes $m_{k,n}$ bits of its virtual buffer, and increments the value of t. If t is smaller than the maximum number $t_{\max}$ of time slots by subcarrier, go to Substep 2.1 for allocating the next time slot. Else, go to next substep.

Substep 2.3. Increment the value of n. If n is smaller than the maximum number $n_{\max}$ of subcarriers, go to Step 2 for allocating the time slots of the next subcarrier. Otherwise, go to Step 3.

Step 3. All virtual buffers are empty; or all time slots of all subcarriers are allocated and the scheduling ends.

5. Performance Evaluation

In this section, we compare the proposed weighted fair opportunistic scheduling with the Round Robin (RR), MaxSNR, PF, and MAOPF schemes implemented with subcarrier by subcarrier allocation. Performance evaluation results are obtained using OPNET discrete event simulations.

In the simulations, we assume 128 subcarriers and 5 time slots in a frame. The channel gain model on each subcarrier considers free space path loss and multipath Rayleigh fading [4]. We introduce a reference distance d_{ref} for which the free space attenuation equals a_{ref}. As a result the channel gain is given by

$$a_{k,n} = a_{\text{ref}} \times \left(\frac{d_{\text{ref}}}{d_k}\right)^{3.5} \times \alpha_{k,n}^2, \qquad (9)$$

where d_k is the distance to the access point of the mobile owning the service flow k, and $\alpha_{k,n}^2$ represents the flat fading experienced by this service flow k if transmitted on subcarrier n. In the following, $\alpha_{k,n}$ is Rayleigh distributed with an expectancy equal to unity.

The maximum transmit power satisfies

$$10\log_{10}\left(\frac{P_{\max} T_s}{N_0} \times a_{\text{ref}}\right) = 31\text{ dB}. \qquad (10)$$

The BER target is taken equal to 10^{-3}. With this setting, the value of $m_{k,n}$ for the mobiles situated at the reference distance is 6 bits when $\alpha_{k,n}^2$ equals unity.

We assume all mobiles run the same videoconference application. This demanding type of application generates a high volume of data with high sporadicity and requires tight delay constraints which substantially complicates the task of the scheduler. Each mobile has only one service flow with a traffic composed of an MPEG-4 video stream [26] and an AMR voice stream [27].

The problem we are studying is quite different with the sum-rate maximization with water filling for instance. The purpose of the scheduler proposed in this paper is to maximize the traffic load that can be admitted in the wireless access network while fulfilling delay constraints. This is achieved by both taking into account the radio conditions but also the variations in the incoming traffic. In this context, we cannot for instance assume that each mobile has some traffic to send at each scheduling epoch. Traffic overload is not realistic in a wireless access network because it corresponds to situations where the excess traffic experiences an unbounded delay. This is why, in all our simulations, the traffic load (offered traffic) does not exceed the system capacity. In these conditions, the offered traffic is strictly equal to the traffic carried over the wireless interface and all mobiles get served sooner or later. The bit rate sent by each mobile is equal to its incoming traffic. Fairness in terms of bit rate sent by each mobile is rigorously achieved. The purpose of the scheduler is to dynamically assign the resource units to the mobiles at the best time in order to meet the traffic delay constraints. This is why we adopted the PDOR as a measure of the fairness in terms of QoS level obtained by each mobile.

Table 1: First scenario setup.

Group	Distance d_k	Delay threshold T_k	Data rate
1	2 d_{ref}	80 ms	80 Kbps
2	3 d_{ref}	80 ms	80 Kbps

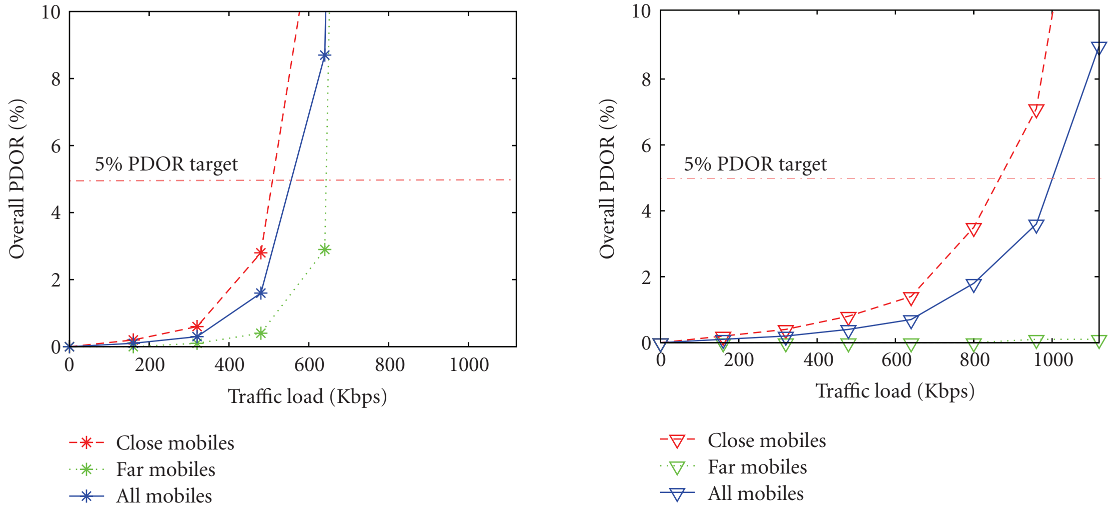
Four simulation scenarii were used in the performance evaluation. In the first scenario, we analyzed the behavior of the schedulers when mobiles occupy different geographical positions. The second scenario examines the performance of the schedulers when mobiles have heterogeneous bit rate requirements. QoS differentiation is evaluated in the third scenario. The fourth simulation scenario considers mobiles with both heterogeneous geographical positions, bit rate, and QoS requirements.

5.1. First Scenario: Influence of the Distance on the Schedulers Performances

5.1. First Scenario: Influence of the Distance on the Schedulers Performances. In wireless networks, it is well known that the closest mobiles to the access point generally obtain better QoS than mobiles more distant thanks to their higher spectral efficiency. In order to study the influence of the distance on the scheduling performances, a first half of mobiles are situated close to the access point and a second half 1.5 farther. The other parameters are identical for all the mobiles as described in Table 1. The total number of mobiles sets the traffic load.

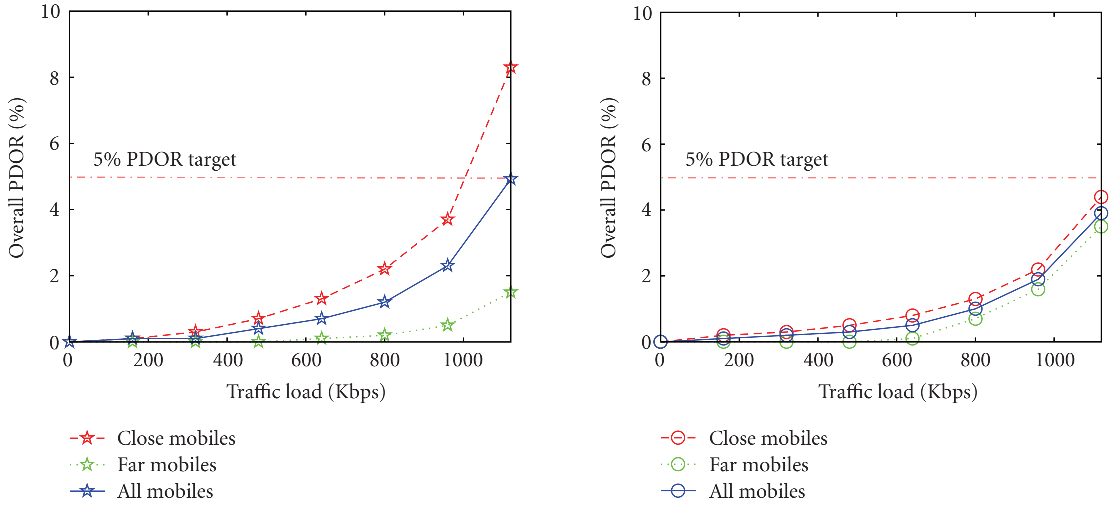
First we focus on the fairness provided by each scheduler. Figures 7(a), 7(b), 7(c), and 7(d) display the overall PDOR for different traffic loads considering the influence of the distance on the scheduling. The classical RR fails to ensure the same PDOR to all mobiles. Actually, the RR fairly allocates the RUs to the mobiles without taking in consideration that far mobiles have a much lower spectral efficiency than closer ones. Moreover, the RR does not take benefit of multiuser diversity which results in a bad utilization of the bandwidth and in turn, poor system throughput. Consequently, an acceptable PDOR target of 5% is exceeded even with relatively low traffic loads. Based on opportunistic scheduling, the three other schemes globally show better QoS performances supporting a higher traffic load. However, MaxSNR, PF, and MAOPF still show severe fairness deficiencies (in this context where all mobiles have an equal source bit rate, the MAOPF and PF perform the same scheduling). Close mobiles easily respect their delay requirement while far mobiles experience much higher delays and go past the 5% PDOR target when the traffic load increases. In contrast, the WFO provides the same QoS level to all mobiles whatever their respective position. The WFO is the only one to guarantee a totally fair allocation. This allows to reach higher traffic loads with an acceptable PDOR for all mobiles. Additionally, looking at the overall PDOR for all mobiles at different traffic loads shows that, besides fairness, the WFO provides a better overall QoS level as well.

Observing the mean buffer occupancy in Figure 8(a), the WFO clearly limits the buffer occupancy to a same and reasonable value whatever the position of the mobile. This allows to stay under the PDOR target for any traffic load. With its system of weights, the WFO dynamically adjusts the



(a) With RR

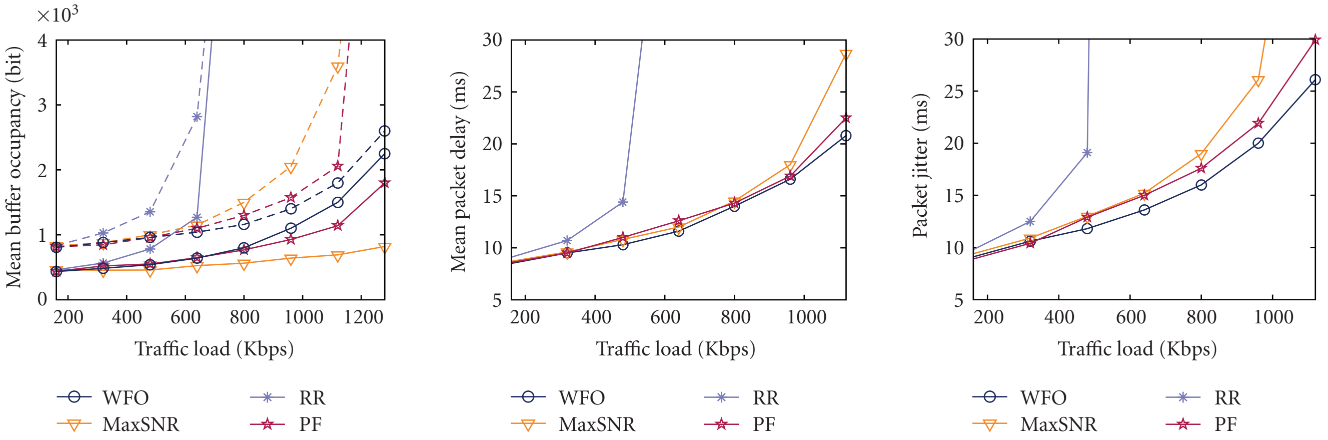
(b) With MaxSNR



(c) With PF (and MAOPF)

(d) With WFO

Figure 7: Measured QoS with respect to distance.

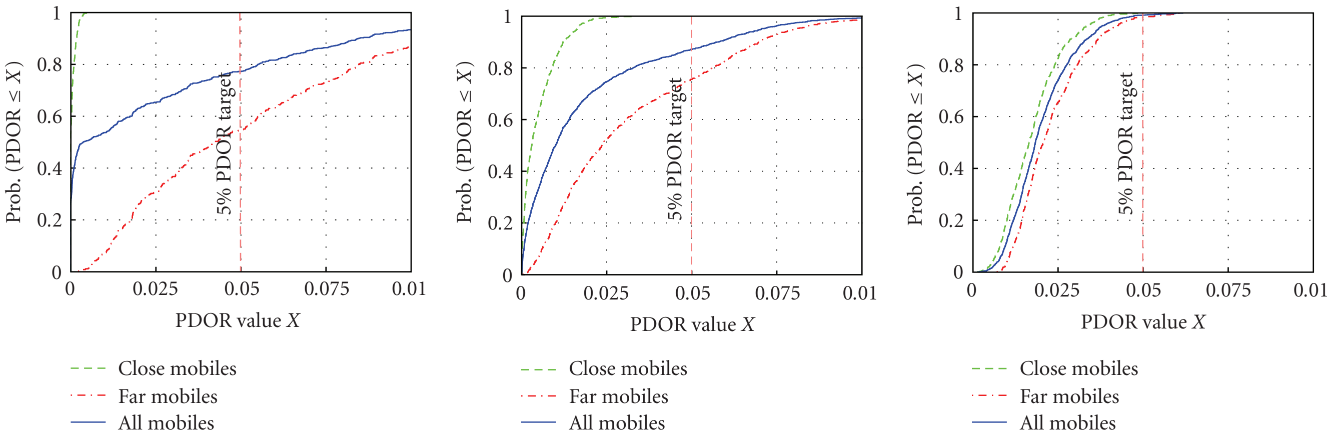


(a) Mean buffer occupancy for close mobiles (solid lines) and far mobiles (dashed lines)

(b) Mean packet delay

(c) Packet jitter

Figure 8: Buffer occupancy, delay and jitter.



(a) CDF of end cycle PDOR with MaxSNR

(b) CDF of end cycle PDOR with PF

(c) CDF of end cycle PDOR with WFO

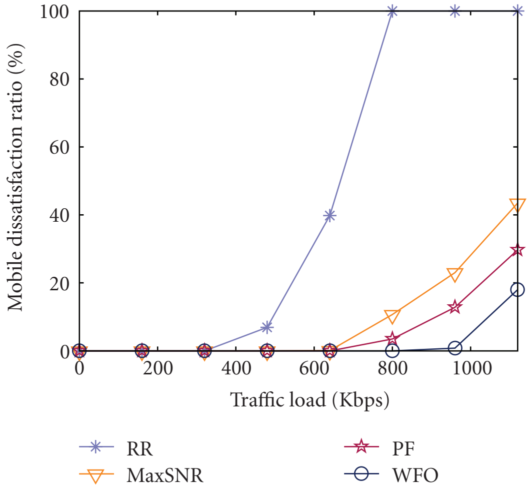
FIGURE 9: Perceived QoS with different allocation schemes.

relative priority of the flows according to their experienced delay. With this approach, sparingly delaying the closer mobiles, the WFO builds on the breathing space offered by the easy respect of the delay constraints of the closer mobiles (with better spectral efficiency) for helping the farther ones. The WFO interesting performance results are corroborated in Figures 8(b) and 8(c), where the overall values of the mean packet delay and jitter obtained using the WFO are smaller.

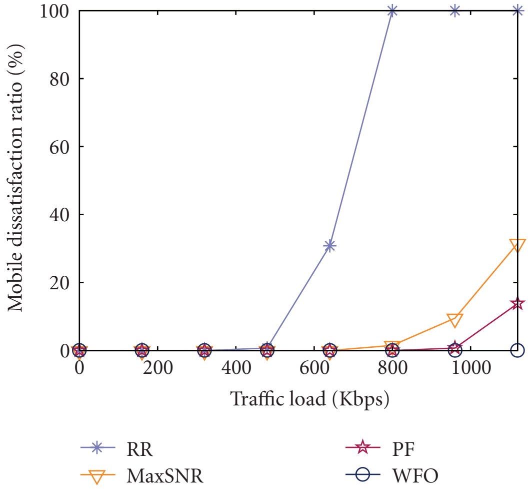
We then had a look at the QoS satisfaction level that each mobile perceives across the lifetime of a connection. We divided the connection of each mobile in cycles of five minutes and measured the PDOR at the end of each cycle. Figure 9 shows the CDF of end cycle PDOR values for a traffic load of 960 Kbps, using, respectively, the MaxSNR, the PF, and the WFO schemes (RR performances are not presented here since they are not able to support this high traffic load). We also estimated the mobile dissatisfaction ratio. We checked if at the end of each cycle the delay constraint is met or not. We then computed the mobile dissatisfaction ratio defined as the number of times that the mobiles are not satisfied (experienced PDOR$\geq$ PDOR$_{\text{target}}$) divided by the total number of cycles (cf. Figure 10).

Highly unfair, MaxSNR fully satisfies the required QoS of close mobiles at the expense of the satisfaction of far mobiles. Indeed, only 54.5 percents of these latter experience a final PDOR inferior to a PDOR target of 5% (cf. Figure 9(a)). Unnecessary priorities are given to close mobiles which easily respect their QoS constraints while more attention should be given to the farther. These inadequate priority management dramatically increases the global mobile dissatisfaction which reaches 23% as shown in Figures 9(a) and 10(a).

PF brings more fairness and allocates more priority to far mobiles. Compared to MaxSNR, PF offers a QoS support improvement with only 12.8% of dissatisfied mobiles (cf. Figures 9(b) and 10(a)). Fairness is still not total since the farther mobiles have a lower spectral efficiency than the closer ones due to path loss. All mobiles do not all benefit of an equal average throughput despite they all obtain an equal share of bandwidth. This induces heterogeneous delays

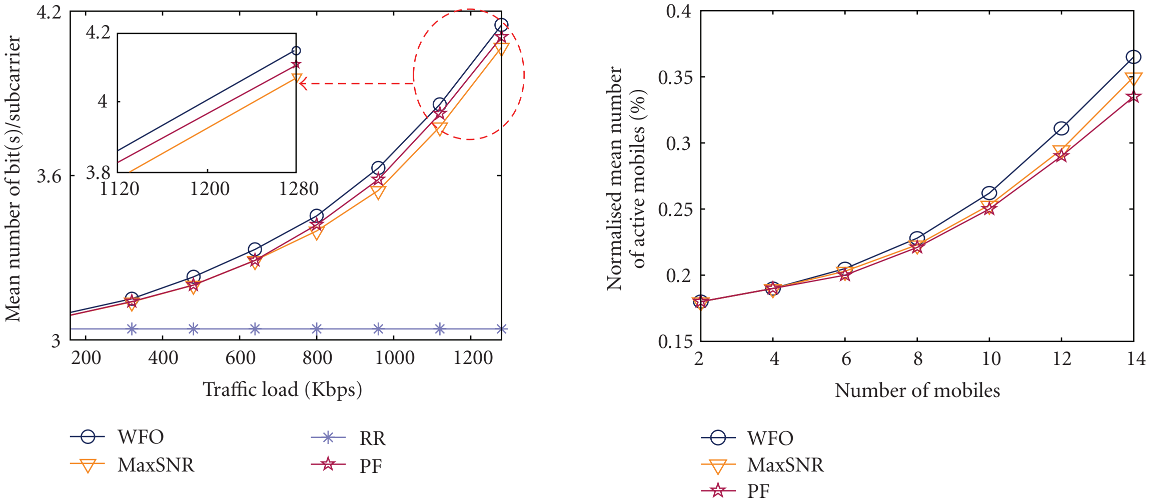


(a) Mobile dissatisfaction—PDOR$_{\text{target}}$ = 5%



(b) Mobile dissatisfaction—PDOR$_{\text{target}}$ = 10%

FIGURE 10: Analysis of the respect of QoS constraints for different targeted QoS.



(a) Spectral efficiency

(b) Multiuser diversity

Figure 11: Bandwidth usage efficiency.

and unequal QoS. This fairness improvement compared to MaxSNR indicates however that some flows can be slightly delayed to the benefit of others without significantly affecting their QoS.

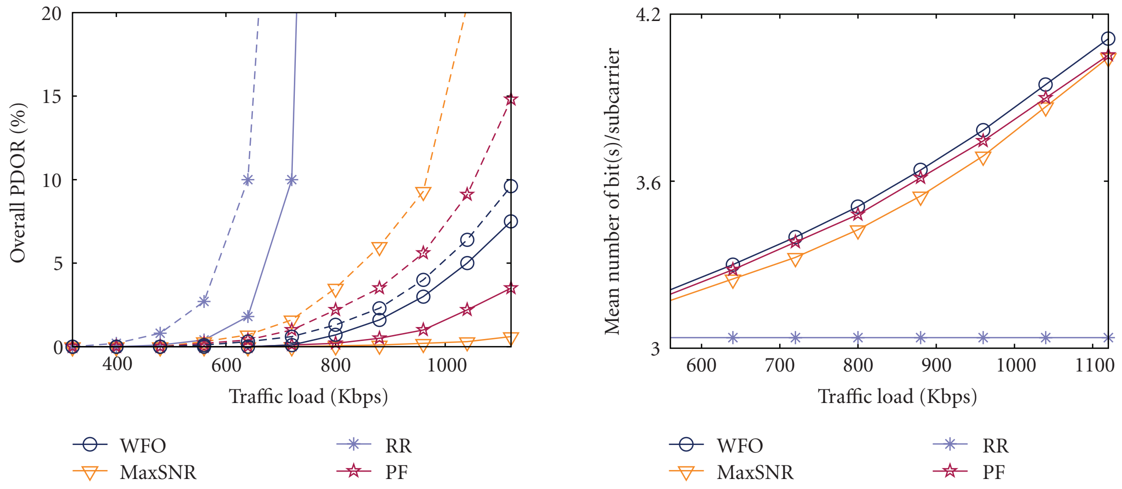
The WFO was built on this idea. The easy satisfaction of close mobiles (with better spectral efficiency) offers a degree of freedom which ideally should be exploited in order to help the farther ones. WFO allocates to each mobile the accurate share of bandwidth required for the satisfaction of its QoS constraints, whatever its position is. With WFO, only 0.8 percents of the mobiles are dissatisfied (cf. Figures 9(c) and 10(a)). Additionally, compared to Figures 9(a), 9(b), and 9(c) exhibits superimposed curves which proves the WFO high fairness, included at short term.

Figure 10 shows that the WFO brings the largest level of satisfaction. Indeed, for a tight PDOR target of 5% (see Figure 10(a)), the dissatisfaction ratio with a high traffic load of 1120 Kbps is equal to 18% with the WFO versus 29.7% with the best of the other scheduling schemes. If we set the PDOR target to 10%, the dissatisfaction ratio with a high traffic load of 1120 Kbps is 0% with the WFO versus 13.8% with the best of the other scheduling schemes (PF).

We finally studied the system capacity offered by the four scheduling algorithms. Figure 11(a) shows the average number of bits carried on a used subcarrier by each tested scheduler under various traffic loads. As expected, the nonopportunistic Round Robin scheduling provides a constant spectral efficiency, that is, an equal bit rate per subcarrier whatever the traffic load since it does not take advantage of the multiuser diversity. The three other tested schedulers show better results. In contrast with RR, with the opportunistic schedulers (MaxSNR, PF, WFO), we observe an interesting inflection of the spectral efficiency curve when the traffic load increases. The join analysis of Figures 11(a) and 11(b) shows that the spectral efficiency of opportunistic scheduling is an increasing function of the number of active mobiles, thanks to the exploitation of this supplementary multiuser diversity. Consequently, MaxSNR, PF, and WFO increase their spectral efficiency with the traffic load, and the system capacity is highly extended compared to networks which use classical scheduling algorithms. With these three schedulers, all mobiles are served even at the highest traffic load of 1280 Kbps.

The performance of the four schedulers can be further qualified by computing the theoretical maximal system throughput. Considering the Rayleigh distribution, it can be noticed that $\alpha_{k,n}^2$ is greater or equal to 8 with a probability of only 0.002. In these ideal situations, close mobiles can transmit/receive 6 bits per RU while far mobiles may transmit/receive 4 bits per RU. If the scheduler always allocated the RUs to the mobiles in these ideal situations, an overall efficiency of 5 bits per RU would be obtained which yields a theoretical maximal system throughput of 1600 Kbps. Comparing this value to the highest traffic load in Figure 11(a) (1280 Kbps) further demonstrates the good efficiency obtained with the opportunistic schedulers that nearly always serve the mobiles when their channel conditions are very good. This result also shows that the WFO scheduling has slightly better performances than the two other opportunistic schedulers. Keeping more mobiles active (cf. Figure 11(b)) but with a relatively lower traffic backlog (cf. Figure 8(a)), the WFO scheme preserves multiuser diversity and takes more advantage of it obtaining a slightly higher bit rate per subcarrier (cf. Figure 11(a)).

In the results described above, the traffic load was varied by increasing or decreasing the number of mobiles in the system, which modified the multiuser diversity. This exhibited the opportunistic behavior of the schedulers and especially their ability to take advantage of the multiuser diversity brought with the increase of the number of mobiles. We also studied below the ability of each scheduler to take profit of the multiuser diversity brought by a given number of users. In Figure 12, we provide complementary results obtained in a context where the traffic load variation is done



(a) Measured QoS for close mobiles (solid lines) and far mobiles (dashed lines)

(b) Spectral efficiency

FIGURE 12: Performances of schedulers with fixed multiuser diversity.

TABLE 2: Second scenario setup.

Group	Number of mobiles	Distance	Delay threshold	Data rate
1	9	1.6 d_{ref}	80 ms	80 Kbps
2	3	1.6 d_{ref}	80 ms	240 Kbps

TABLE 3: Third scenario setup.

Group	Number of mobiles	Distance	Delay threshold	Data rate
1	7	2.7 d_{ref}	80 ms	80 Kbps
2	7	2.7 d_{ref}	250 ms	80 Kbps

through just increasing the mobile bit rate requirement and keeping a constant number of users (10 mobiles). The results in Figure 12(a) show that, like above, the WFO outperforms the other scheduling schemes. With its weighted algorithms, the WFO dynamically adjusts the mobiles priority and ensures a completely fair allocation. WFO is the only one which allows to reach higher traffic loads with an acceptable PDOR for all mobiles. Additionally, even if the traffic load increases without variation in the number of mobiles, the WFO keeps more mobiles active across the time than the other schemes and takes better advantage of the multiuser diversity. The analysis of Figure 12(b). confirms that WFO maximizes the average bit rate per subcarrier.

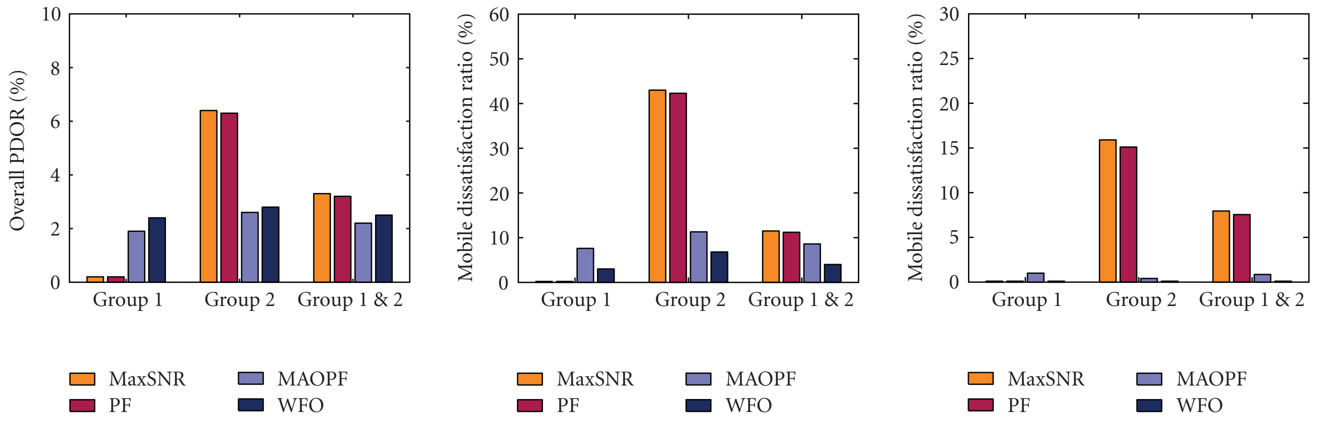
5.2. Second Scenario: Performance with Heterogeneous Bit Rate Sources. In this simulation scenario, mobiles are divided in two groups that differ only by their data rate as described in Table 2.

The four opportunistic scheduling strategies provide the same bandwidth usage ratio of 82% (RR performances are not reported here and in the following because its poor performances do not support the tested configurations). However, delay management considerably differs. Figure 13(a) shows the overall ratio of packets delivered after the threshold time, respectively, in Group 1, Group 2, and globally. The results show that the MaxSNR and the PF easily respect the delay constraints of low bit-rate mobiles but fail for the second group of mobiles. In contrast, the MAOPF and the WFO schemes provide fairness with an equal and moderate ratio of packets in delay outage whatever the source bit rate. The overall PDOR obtained with the MAOPF and the WFO is smaller than with the two other schemes. Here, the two multimedia oriented schedulers provide fair QoS management and better QoS support. Regarding the perceived QoS, Figures 13(b) and 13(c) show that the WFO outperforms the other schedulers including the MAOPF which do not directly manage the PDOR fluctuations.

5.3. Third Scenario: Performance with Heterogeneous Delay Constraints. We then studied the influence of heterogeneous delay requirements on the scheduling performances. In this simulation scenario, mobiles are divided in two groups that differ only by their delay requirements (cf. Table 3).

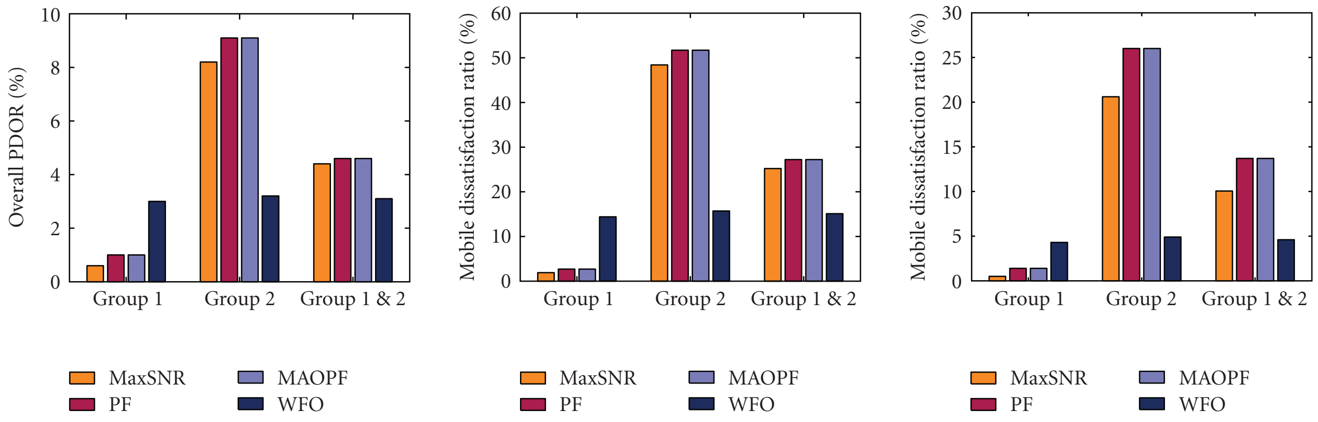
In this context where all mobiles have an equal source bit rate, the MAOPF and PF perform the same scheduling. Figure 14 clearly shows that the WFO outperforms the three other schemes ensuring fair QoS support and provides the largest QoS satisfaction level. This is processed with the WFO weighted system which dynamically controls the delay in a generic manner by monitoring the distance to the delay threshold thanks to a continuous and efficient regulation of the PDOR. This provides full QoS differentiation.

As explained above, the sum of incoming traffics of the mobiles is inferior to the system throughput. In this context, the traffic of each mobile is served sooner or later, and the bit rate sent by each mobile is equal to its incoming traffic. Fairness is absolute in terms of bit rate sent by each mobile. High-delay-sensitive mobiles are not served more often than



(a) Overall PDOR (b) Mobile dissatisfaction—$\text{PDOR}_{\text{target}} = 5\%$ (c) Mobile dissatisfaction—$\text{PDOR}_{\text{target}} = 10\%$

Figure 13: Measured QoS with heterogeneous sources in terms of bit rate.



(a) Overall PDOR (b) Mobile dissatisfaction—$\text{PDOR}_{\text{target}} = 5\%$ (c) Mobile dissatisfaction—$\text{PDOR}_{\text{target}} = 10\%$

Figure 14: Measured QoS with heterogeneous sources in terms of delay requirement.

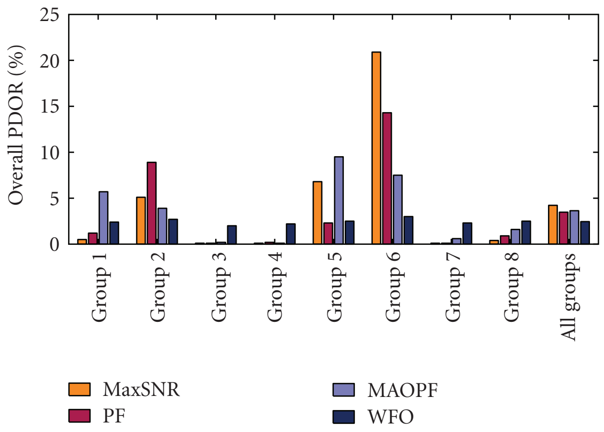


Figure 15: Overall PDOR.

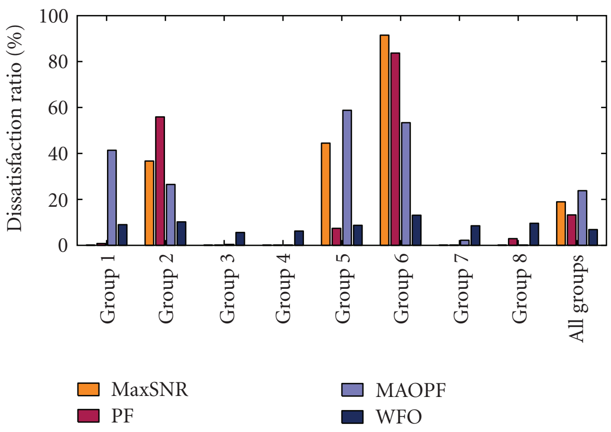


Figure 16: Mobile dissatisfaction when $\text{PDOR}_{\text{target}} = 5\%$.

other mobiles but earlier. It is only the time instant at which each high-delay-sensitive mobile and background mobile is served that differs. The purpose of the tested schedulers is to set dynamic priorities between the different types of traffics.

5.4. Fourth Scenario: Global Scheduling Performances Analysis. So far, we have analyzed the behavior of the schedulers in simple contexts considering one criterion at a time for better understanding its influence on the performances. In order to

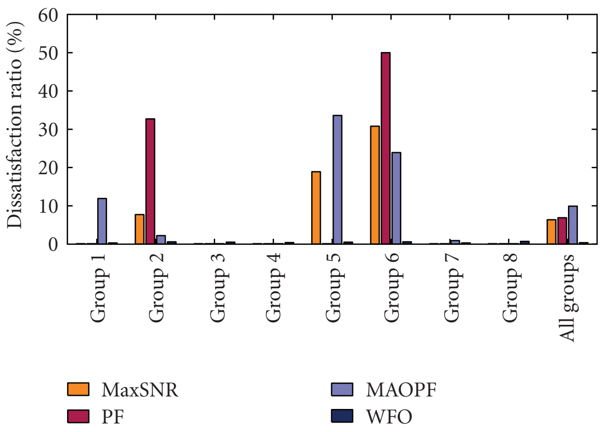


FIGURE 17: Mobile dissatisfaction when $\mathrm{PDOR}_{\mathrm{target}} = 10\%$.

TABLE 4: Fourth scenario setup.

Group	Number of mobiles	Distance	Delay threshold	Data rate
1	2	$2\,d_{\mathrm{ref}}$	80 ms	80 Kbps
2	1	$2\,d_{\mathrm{ref}}$	80 ms	160 Kbps
3	2	$2\,d_{\mathrm{ref}}$	250 ms	80 Kbps
4	1	$2\,d_{\mathrm{ref}}$	250 ms	160 Kbps
5	2	$2.4\,d_{\mathrm{ref}}$	80 ms	80 Kbps
6	1	$2.4\,d_{\mathrm{ref}}$	80 ms	160 Kbps
7	2	$2.4\,d_{\mathrm{ref}}$	250 ms	80 Kbps
8	1	$2.4\,d_{\mathrm{ref}}$	250 ms	160 Kbps

corroborate the good results of the WFO, we study in this section the performance of the tested protocols in a more general context. Eight groups of mobiles are considered here as described in Table 4.

Figures 15, 16, and 17, respectively, show the overall packet loss ratio and the dissatisfaction ratio with a PDOR target set to 5% and 10% for each group of mobiles and on the right, for all groups. MaxSNR provides a very poor QoS in groups 2, 5, and 6, that is, when delay requirements are stringent and the path loss or the source bit rate is high. This result confirms that MaxSNR severely lacks fairness in realistic scenarii. Mobile position has less consequences on fairness with PF. However, PF still shows deficiencies for mobiles with high data rate and tight delay threshold (groups 2 and 6). In comparison with PF, MAOPF brings more fairness between mobiles with heterogeneous data rate. Groups 2 and 6 experience less difficulties but at the expense of the satisfaction of groups 1 and 5. Globally, MaxSNR, PF, and MAOPF provide comparable performance results, each of them penalizing selectively some of the groups of mobiles. In contrast, WFO performs an efficient multiplexing and jointly manages all the mobiles so that they are all satisfied in a same proportion whatever their respective QoS constraints, positions, or data rate specificities. WFO allows to respect the delay thresholds in equity for all mobiles and satisfy the largest number.

6. Conclusion

In this paper, we propose a new MAC protocol for wireless multimedia networks, called "weighted fair opportunistic (WFO)" protocol. This access scheme operates on top of an OFDM-based physical layer and shows a good compatibility with the existing 802.16 standard. Full support of evolved multimedia services and QoS differentiation is enabled with the introduction of generic QoS attributes. Based on a system of weights, the WFO scheduling introduces dynamic priorities between the mobiles according to their transmission conditions and the delay they currently experience in a higher layers/MAC/PHY cross-layer approach. With its well-balanced resource allocation, the WFO scheme keeps a maximum number of service flows active across time but with relatively low traffic backlogs. Preserving the multiuser diversity, it takes a maximal benefit of the opportunistic scheduling technique for maximizing the system capacity. Simulation results show that the WFO outperforms other wireless OFDM-based scheduling schemes providing efficient QoS management. Fairness is ensured whatever the mobile position, the bit rate, or the delay constraints and without never sacrificing system capacity.

Acknowledgments

This work has been partially supported by the IST European Project WIP under Contract no. 27402. The authors also thanks the reviewers for their constructive comments.

References

[1] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.

[2] R. Steele and L. Hanzo, *Mobile Radio Communications*, IEEE Computer Society Press, Los Alamitos, Calif, USA, 2000.

[3] A. Goldsmith, *Wireless Communications*, Cambridge University Press, Cambridge, UK, 2005.

[4] J. D. Parsons, *The Mobile Radio Propagation Channel*, John Wiley & Sons, New York, NY, USA, 1992.

[5] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proceedings of IEEE International Conference on Communications (ICC '95)*, vol. 1, pp. 331–335, Seattle, Wash, USA, June 1995.

[6] C. Y. Wong, R. S. Cheng, K. Ben Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, 1999.

[7] X. Wang and W. Xiang, "An OFDM-TDMA/SA MAC protocol with QoS constraints for broadband wireless LANs," *Wireless Networks*, vol. 12, no. 2, pp. 159–170, 2006.

[8] Y. Fukui, N. Yamagaki, H. Tode, and K. Murakami, "Packet transfer scheduling scheme with throughput compensated considering wireless conditions," in *Proceedings of IEEE International Conference on Computer Communications and Networks (ICCCN '03)*, pp. 11–16, Dallas, Tex, USA, October 2003.

[9] H. Zhu and K. R. Liu, "Throughput maximization using adaptive modulation in wireless networks with fairness constraint," in *Proceedings of IEEE International Wireless Communications*

and Networking Conference (WCNC '03), vol. 1, pp. 243–246, New Orleans, La, USA, March 2003.

[10] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.

[11] H. Kim, K. Kim, Y. Han, and J. Lee, "An efficient scheduling algorithm for QoS in wireless packet data transmission," in *Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '02)*, vol. 5, pp. 2244–2248, Lisbon, Portugal, September 2002.

[12] W. Anchun, X. Liang, Z. Shidong, X. Xibin, and Y. Yan, "Dynamic resource management in the fourth generation wireless systems," in *Proceedings of the International Conference on Communication Technology (ICCT '03)*, vol. 2, pp. 1095–1098, Beijing, China, April 2003.

[13] P. Svedman, S. K. Wilson, and B. Ottersten, "A QoS-aware proportional fair scheduler for opportunistic OFDM," in *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC '04)*, vol. 1, pp. 558–562, Los Angeles, Calif, USA, September 2004.

[14] H. Kim, K. Kim, Y. Han, and S. Yun, "A proportional fair scheduling for multicarrier transmission systems," in *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC '04)*, vol. 1, pp. 409–413, Los Angeles, Calif, USA, September 2004.

[15] J.-G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 766–778, 2007.

[16] J. M. Holtzman, "Asymptotic analysis of proportional fair algorithm," in *Proceedings of the 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '01)*, vol. 2, pp. 33–37, San Diego, Calif, USA, September-October 2001.

[17] J. Z. Haiying and R. H. M. Hafez, "Novel scheduling algorithms for multimedia service in OFDM broadband wireless systems," in *Proceedings of IEEE International Conference on Communications (ICC '06)*, vol. 2, pp. 772–777, Istanbul, Turkey, July 2006.

[18] IEEEStd 802.16-2004, "IEEE standard for local and metropolitan area networks, part 16: air interface for fixed broadband wireless access systems," October 2004.

[19] C. Hoymann, "Analysis and performance evaluation of the OFDM-based metropolitan area network IEEE 802.16," *Computer Networks*, vol. 49, no. 3, pp. 341–363, 2005.

[20] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150–154, 2001.

[21] J.-J. van de Beek, P. O. Börjesson, M.-L. Boucheret, et al., "A time and frequency synchronization scheme for multiuser OFDM," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1900–1914, 1999.

[22] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 461–471, 1999.

[23] T. E. Truman and R. W. Brodersen, "A measurement-based characterization of the time variation of an indoor wireless channel," in *Proceedings of the 6th IEEE International Conference on Universal Personal Communications Record (ICUPC '97)*, vol. 1, pp. 25–32, San Diego, Calif, USA, October 1997.

[24] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 10, pp. 2053–2064, 2001.

[25] S. S. Kulkarni and C. Rosenberg, "Opportunistic scheduling: generalizations to include multiple constraints, multiple interfaces, and short term fairness," *Wireless Networks*, vol. 11, no. 5, pp. 557–569, 2005.

[26] S. Baey, "Modeling MPEG4 video traffic based on a customization of the DBMAP," in *Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS '04)*, pp. 705–714, San Jose, Calif, USA, July 2004.

[27] P. Brady, "A model for generating on-off speech patterns in two-way conversation," *The Bell System Technical Journal*, vol. 48, no. 9, pp. 2445–2472, 1969.