A Cross-Layer Downlink Scheduling Scheme for Balancing QoS in IEEE 802.16 Broadband Wireless Access Systems

Hassen Hamouda
Department of computer science, CSHG
Majmaah University, Saudi Arabia
Email: h.hamouda@mu.edu.sa

Mohamed Ouwais Kabaou
National School of Engineers of Gabes
University of Gabes, Tunisia
Email: ouwayess@yahoo.fr

Med Salim Bouhlel
Higher Institute of Biotechnology of Sfax
University of Sfax, Tunisia
Email: medsalim.bouhlel@enis.rnu.tn

***Abstract*—IEEE 802.16 OFDMA wireless network is expected to be the dominant system used by operators in these last decade due to its promising solutions for providing high data rate. However, IEEE 802.16 packet scheduling and resources assigning among Mobile Stations (MSs) is still the main challenge due to unfairness and low performance which occur when allocating resources to MSs. Furthermore, MaxSNR scheduler is often recognized by the scientific community as the most effective in OFDMA wireless networks. Thanks to its opportunistic operation, it takes into account changes in the states of links, maximizes system throughput and so can accommodate a very large number of users in the network. However, this does not come cheap: MSs far from the access point are systematically penalized compared to closer ones; this is because of their little favorable transmission conditions (caused by path loss). In this paper, we propose an improvement to MaxSNR scheduler called Dynamic and Fair MaxSNR (DFMaxSNR). The DFMaxSNR consists not only of exploiting the concept of opportunistic scheduler in order to maximize system throughput, but also of correcting the unequal spectral efficiencies induced by path loss attenuations in order to minimize the delay and the packet loss rate for real time applications. Performance evaluation shows that this well-balanced resource allocation scheme outperforms other existing schedulers (MaxSNR, PF and RR) and demonstrate that choosing between high system capacity and high fairness is not required.**

***Keywords*-multiuser diversity, opportunistic scheduling, throughput maximization, Fairness hierarchy, QoS, QoE.**

I. Introduction

The main purpose of a resource allocation scheme is to maximize the system throughput while ensuring fairness, delay and packet loss rate within QoS requirements to satisfy end-MSs QoE. In the research literature, numerous cross-layer (physical (PHY) and medium-access control layers) fairness-aware scheduling schemes for IEEE 802.16 systems have been reported [1], [2], [3]. Major of theses schemes are based on three well-known schedulers: Round Robin (RR), Maximum Signal to Noise Ratio (MaxSNR) and Proportional Fair (PF). First of all, RR is a very simple scheduler, which can be easily implemented and used regularly as a reference to measure the gain of the performances of the modern schedulers. The RR assigns resources to each MS in a round robin fashion, one after the other [4], [5]. Each MS is therefore sure not only to have the same amount of resources but also to reach the medium regularly which allows making this allocation strategy fair. However, because of its non-opportunistic process, RR does not take into account radio conditions and consequently it affects both system throughput and multiuser diversity. Many studies have, therefore, sought to address this critical issue for current and future networks. They have concluded that an opportunistic approach is a paramount solution to achieve an optimal allocation of radio resources. Primarily, two well known algorithms have emerged: the Maximum Signal-to-Noise Ratio (MaxSNR)(also known as Maximum Carrier to Interference ratio (Max C/I)) and the Proportional Fair (PF).

Taking advantage of the multiuser and frequency diversity, MaxSNR scheduling constantly allocates the radio resource to the MS who has the best spectral efficiency (highest SNR) [6], [7], [8], [9]. Based on a lower path loss and therefore a greater SNR strategy, nearby MSs will often be, if not always, selected before remote MSs which will be allocated only the remainder. For this reason MaxSNR is currently considered as the most efficient scheduler from the perspective of maximizing the throughput but also the least fair.

To handle the constraints of high throughput and fairness, the Proportional Fair (PF) scheme was developed. PF uses a strategy based on the trade-off between maximum achievable average throughput and fairness. The proportional fair scheme (PF) has been widely used in mobile networks with the objective of maximizing the long-term fairness and throughput [10], [11]. However, it has no mechanism to guarantee the QoS for real-time services.

In this work, we proposed a dynamic an fair scheduling algorithm that combines opportunistic concept and buffer occupancy status. The originality of our scheduler resides on considering a new parameter called Fairness Factor (FF) which balance the QoS amongst all MSs. The principle of our proposal as well as its performance analysis will be discussed in the next sections.

The rest of this paper is organized as follows. Section II, introduces the system description. In section III, we present the design of our proposed DFMaxSNR scheduler. Performance evaluation and results discussion is given in Section IV. We finally draw the conclusion of our work in Section V.

978-1-5090-5935-5/17/$31.00

II. System Model

The paper focuses on the radio resources allocation among the set of MSs situated in the coverage zone of an access point. Furthermore, we focuses on a centralized approach which consists of scheduling the downlink transmissions after buffering the packets originating from the backhaul network in the access node. At the physical layer, we assume the existence of a structure of frames as that of the IEEE 802.16-2004 system [12]. The duration of a frame is fixed at 2 ms, a value less than the channel coherence time. With these assumptions, the transmission on each subcarrier is subject to constant attenuation over frame duration and therefore, the state of the links of each of these can be considered as static on its duration. In the following, an elementary Resource Unit (RU) is defined as the couple "subcarrier, time interval". Moreover, each RU can be allocated according to the criteria of the system scheduler to any MS belonging to the coverage area of the access point. Furthermore, transmissions on these RUs is carried out at a specific modulation order of Quadrature Amplitude Modulation(QAM) determined using the Channel State Information (CSI), which depends on the state of the link associated to the sub-band of the RU frequency considered and selected for the MS. Perfect knowledge of the channel state is assumed to be available at the receiver. The current channel attenuation on each subcarrier and for each MS is estimated by the access point based on the SNR of the signal sent by each MS during the uplink subframe.

III. Dynamic and Fair MaxSNR Algorithm

The MaxSNR scheduler allocates the RU to the flow whose $m_{k,n}$ parameter is the greatest. This strategy maximizes the spectral efficiency of the system and therefore allows efficient use of the bandwidth. However, it gives a high priority to MSs close to the access point over the remote ones causing a serious unfairness in the network. To solve this problem, we propose an improvement to MaxSNR called Dynamic and Fair Maximum Signal-to-Noise Ratio (DFMaxSNR).

A. *Principle of DFMaxSNR algorithm*

The main purpose of our proposal consists of maintaining the ability of the original algorithm in maximizing throughput while correcting the unfairness it induces. The principle of this correction is to achieve an opportunistic allocation while assigning more RUs to remote MSs so that all network MSs get the same throughput. Therefore, The maximum number of bits $m_{k,n}$, which can be transmitted in a time slot of the subcarrier n if this RU is allocated to the flow of the MS k, is :

$$m_{k,n} = \max\{q \in S, q \leq q_{k,n}\} \tag{1}$$

Where the maximum number of bits $q_{k,n}$ that may be transmitted to the flow of the MS k in a time slot on the BER_{target} can then be expressed as follow:

$$q_{k,n} = \left\lfloor \log_2 \left(1 - \frac{3 * P_{\max} * T_s * a_k \times a_{k,n}^2}{2N_0 \left[erfc^{-1}\left(\frac{BER_{\text{target}}}{2}\right)\right]^2} \right) \right\rfloor \tag{2}$$

In order to compensate the harmful influence of path loss on fairness, a new parameter called "Fairness Factor" (FF) is introduced. The FF parameter allows turning the MaxSNR into a fair algorithm without affecting the overall system capacity.

$$FF_k = \frac{PN_k^{\alpha}}{P^{\alpha}} \tag{3}$$

Where PN_k^{α} is the number of packets existing in the buffer of the MS k during the scheduling phase α and P^{α} is the number of packets received until the scheduling phase α. For each time slot, the DFMaxSNR scheduler refresh values of $m_{k,n}$ and FF_k and allocates a time interval of the subcarrier n to the MS j which has the most favorable conditions for transmission (best $m_{k,n}$) in relation to its FF_k value:

$$j = argmax_{\text{k}} \left\{ \begin{matrix} m_{k,n} & if\ FF_k == 1 \\ FF_k & if\ 0 <= FF_k < 1 \end{matrix} \right\},\ k = 1,..,K \tag{4}$$

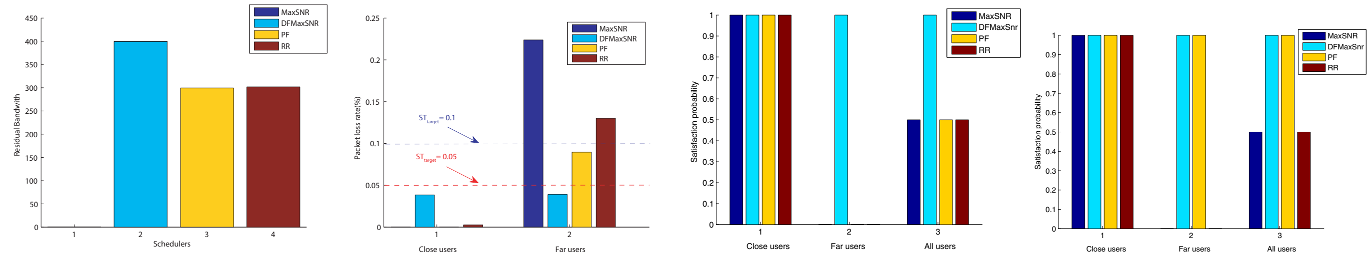
where K is the number of active MSs.

Thanks to FF_k parameter, the negative impact of path loss on the fairness is corrected. In fact, this parameter compensates the low throughput per RU obtained by remote MSs by proportionally increasing the priority of their flows. Moreover, taking into consideration the $m_{k,n}$ in the priority allocation process, the selected flows are those offering good transmission capacity. Indeed, our dynamic scheduler takes decisions based not only on the state of links but also on the state of the buffer occupancy of all MSs, contrarily to MaxSNR which is based only on the state of links.

B. *DFMaxSNR Description*

The DFMaxSNR scheduling scheme is performed subcarrier by subcarrier for each time slot and the flow of each active MS can be determined by their $m_{k,n}$ and FF_k values. The detailed description of our approach is given below:

- **Step 0**: The scheduler refreshes the $m_{k,n}$, CC_{max} (the maximum number of subcarriers that can be allocated to the user k) and NSC_{max} (the maximum number of subcarriers available in the system) values. Integers CC_k (the current number of subcarriers allocated to the user k), NSC (the current number of subcarriers processed by the scheduler), and t (time slot) are initialized to 0, 1 and 1 respectively.
- **Step 1**: For each time slot, the scheduler refreshes BO_k, PN_k^{α} and P^{α} values and calculates FF_k value.
- **Step 2**: For each subcarrier, If FF_k is equal to 1, the scheduler selects the flow of the user k who has the highest $m_{k,n}$. Else the scheduler selects the flow of the user k who has the highest FF_k. If multiple users have the same FF_k value, the scheduler selects the user with the highest BO_k. If multiple users have the same FF_k and BO_k values, the scheduler selects the user with the lowest CC_k.
  - **Sub-Step 2-1**: The scheduler allocates the RU (time slot t and subcarrier n) to the flow of the user k with



(a) Residual Bandwidth

(b) Packet loss rate

(c) Satisfaction probability per group when ST_{target}=0.05

(d) Satisfaction probability per group when ST_{target}=0.1

Fig. 1. QoE and system capacity performances when users are in static location

a capacity of $m_{k,n}$ bits, subtracts $m_{k,n}$ bits from the BO_k value. If the current packet is served, the scheduler decrement NP_k (number of packets of the user k). The scheduler proceeds to step 2-2.

- **Sub-Step 2-2**: The scheduler increments the value of CC_k with one unit and increments the value of NSC by one. If CC_k value is greater than CC_{max} value, the scheduler resets CC_k values to 0. If NSC is greater than NCS_{max} then NSC is reinitialized to 1. The scheduler proceeds to step 2-3.
- **Sub-Step 2-3**: The scheduler increments t by one unit. If t value is greater than t_{max}, the scheduler proceeds to step 3. Otherwise, it returns to step 1.

- **Step 3**: The allocation process is completed.

IV. Performance evaluation

In this section we compares the proposed dynamic and fair opportunistic scheduler with the MaxSNR, the PF and the RR. In the simulations, we assume 1024 subcarriers and 500 time slots. Furthermore, we chose a WiMax network with QAM (2,4,6,8) where the order can be adapted according to the radio conditions, subcarrier by subcarrier. In addition, we chose to test our scheme with a very demanding type of application. In the simulations presented below, all MSs run a videoconference application. Each MS has a traffic composed of an MPEG-4 video stream [13] multiplexed with an AMR voice stream [14]. Moreover, the Urgency Waiting Time (UWT) tolerated for each streaming packet was fixed to 200 ms, the average bit rate was equal to 400 kbit/s and the Satisfaction Threshold target (ST_{target}) was initially fixed to 5% and then to 10%.

A. Static location scenario

In this scenario we took as assumption that all MSs were stationary and situated in a fixed distance from the access point. In addition, in order to study the influence of distance on the behavior of each scheduler, we considered that a first half of MSs (Group 1) are positioned close to the base station at a d_{ref}. However, the second half of MSs (Group 2) was located twice as far. Thus, the number of bits per RU (values of $m_{k,n}$) for group 1 and group 2 was equal to 8 and 4, respectively.

1) System capacity analysis: We intend by this section to analyze the effect of all allocation schemes on the system capacity.

Figure 1.a corroborates that the DFMaxSNR keeps more residual bandwidth available for hosting other potential MSs. The result clearly show that the DFMaxSNR scheduler provides the best bandwidth usage compared to MaxSNR, PF and RR.

2) Quality of Service (QoS) and Quality of Experience (QoE): In order to evaluate the satisfaction degree of end MSs, we have proposed an optimization problem [15] called Satisfaction Probability ($P_{satisfaction}$) which consists of maximizing the number of satisfied users. This optimization problem is formulated as follows:

$$\begin{cases} max(P_{satisfaction}) \quad Such\ as \quad PLR_k < ST_{target} \\ \\ P_{satisfaction} = \frac{Number\ of\ satisfied\ users}{Total\ number\ of\ users} \end{cases} \tag{5}$$

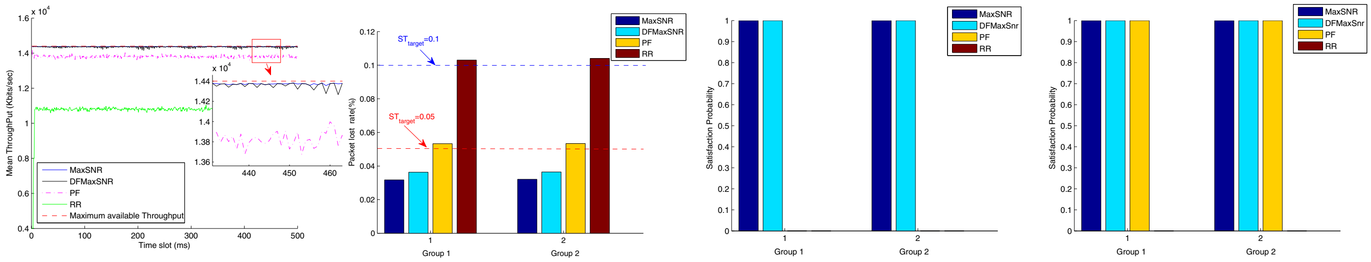
where PLR_k is the packet loss rate of user k.

$$PLR_k = \frac{DP_k}{NP} \tag{6}$$

where DP_k is the number of dropped packets of the user k and NP is the total number of packets. The Packet Loss Rate (PLR_k) experienced by each user is tracked all along the lifetime of its connection. At each time slot the total number of packets whose delay exceeded the Urgency Waiting Time (UWT) divided by the total number of packets transmitted since the beginning of the connection is computed. A user is considered satisfied if its PLR_k is less than ST_{target}.

Moreover, before starting our analyze, we must define rigorously diverse levels of fairness hierarchy. Thereby, fairness can be classified as follows:

- Level 1 fairness hierarchy: Fairness can be considered in terms of access to the resource. All MSs will have the same number of subcarriers.
- Level 2 fairness hierarchy: However, given the disparity of radio conditions between MSs, allocating the same amount of RUs to each MS will not guarantee them the same rate. Based on this point of view, the highest level



(a) System throughput (b) Packet loss rate per group (c) Satisfaction probability per group when ST_{target}=0.05 (d) Satisfaction probability per group when ST_{target}=0.1

Fig. 2. QoE and system capacity performances when users are in dynamic location

of fairness consist of ensuring an equivalent rate for all MSs.

- Level 3 fairness hierarchy: However, in wireless networks, offering the same rate to two MSs when one has low needs (in terms of throughput, maximum delay, jitter, BER...) and the other has more severe constraints seems unfair. Giving "too much" to the former, he will always be greatly satisfied, while the latter will necessarily be much less. In fact, the optimal fairness is not to provide the same rate to everyone but rather to secure the same degree of satisfaction.

Our study focuses now on the ability of schedulers to ensure fairness. Being fair consists of ensuring a high QoS and consequently a same satisfaction degree (QoE) to all MSs. Figure 1.b shows the PLR per group, Figure 1.c and Figure 1.d shows the probability of satisfaction ($P_{satisfaction}$) measured by each scheduler when ST_{target} is equal to 0.05 and then to 0.1. These results corroborates that both the opportunistic MaxSNR scheduler and the classical RR scheduler provides the worst performance and are unable to ensure the same satisfaction degree for MSs of different groups. it appears that only close users respect their QoS target with a final PLR_k smaller than ST_{target} (both ST_{target} equal to 0.05 and ST_{target} equal to 0.1). Consequently, 50% of users (users of group 1) are satisfied and 50% of users (users of group 2) are unsatisfied. In fact, the RR shares RUs equitably among all MSs but never considers the fact that the farthest MSs have a lower spectral efficiency than the nearest ones. This results in an inequality in instantaneous throughput and thus provides an unfairness. Furthermore, the MaxSNR scheduler shows large disparities in the treatment of the two groups. Nearby users see that their flow easily meet the delay constraints ($PLR_k < ST_{target}$) while the farther ones see much of their packets arrive "late" ($PLR_k >= ST_{target}$). The more the load increases, the more this difference in treatment is marked, leading ultimately to an early degradation of the Qos and consequently of the QoE of the remote MSs. However, the PF is fairer and ensures more priority to remote MSs providing to everyone in the medium/long term the same number of RUs. Indeed, results provided by PF scheme represents an improvement compared to the ones of MaxSNR. it appears that when ST_{target} is equal to 0.05, only close MSs respect their QoS target with a final PLR_k smaller than ST_{target}. Consequently 100% of MSs of the first group are satisfied and 100% of MSs of the second group are unsatisfied. These results are explained by the fact that remote MSs do not have the same spectral efficiency as the closest MSs due to path loss. Despite the equal sharing of bandwidth between MSs, different rates are achieved inducing disparities in the packet transfer time, QoS and QoE levels. However, when ST_{target} is equal to 0.1, 100% of the close MSs as well as 100% of the far MSs respect their QoS target with a final PLR_k smaller than ST_{target}. Consequently 100% of MSs of the first group as well as 100% of MSs of the second group are satisfied. In contrast, the DFMaxSNR provides a high level of QoS for all MSs. Thereby, in order to satisfy the QoS constraints, a precise share of bandwidth is allocated to each MS regardless of its position. In fact, the new prioritization system based on FF_k parameter significantly reduces the PLR of remote MSs without causing a real "shortfall" to nearby MSs. Consequently, 100% of the close MSs as well as 100% of the far MSs respect their QoS target with a final PLR smaller than ST_{target}. Compared to MaxSNR, PF and RR, DFMaxSNR still offers a real improvement in terms of QoS and QoE (Level 3 fairness hierarchy).

B. Dynamic location context

In the previous scenario, the groups of users were static and placed in two different positions of the access point. The objective was to compare the behavior of schedulers and to study their ability to ensure fairness. Indeed, in order to provide more analysis, we adopt now a more general context including an intracellular mobility. In this scenario, we supposed that both users groups move within the coverage zone with a velocity equal to 3km/h.

1) System Capacity: The performance of these algorithms can be qualified by computing the system throughput guaranteed by each scheduler. Figure 2.a and Figure 2.a shows that even though MaxSNR is always the best scheme in ensuring the best system throughput because it serves only users with the highest $m_{k,n}$, DFMaxSNR scheduler brings an important improvement close to MaxSNR in terms of system throughput and system capacity. However, RR scheduler is always the worst scheduler that ensures the worst system throughput as well as the worst system capacity.

2) Quality of Service (QoS) and Quality of Experience (QoE) : The performance of these algorithms can be qualified by computing the PLR, the number of satisfied and unsatisfied users and the satisfaction probability provided by each scheduler.

Figure 2.b shows the PLR per group and Figure 2.c and Figure 2.d shows the satisfaction probability per group when ST_{target} is set at 0.05 and ST_{target} is set at 0.1 respectively.

The results clearly show that the classical RR scheduler is the worst performer because it is incapable to ensure the same satisfaction degree for MSs of different groups. Consequently all MSs have not been able to respect their QoS target with a final PLR_k greater than ST_{target} (ST_{target}=0.05 and ST_{target}=0.1). Indeed, 100% of users are unsatisfied and the satisfaction probability ($P_{satisfaction}$) guaranteed by the RR scheduler is equal to 0. However, results provided by PF scheduler represents an improvement compared to the ones of RR. it appears that when ST_{target} is equal to 0.05, all users do not respect their QoS target with a final PLR_k greater than ST_{target}. Consequently 100% of users are unsatisfied and the satisfaction probability ($P_{satisfaction}$) guaranteed by PF scheduler is equal to 0. However, when ST_{target} is equal to 0.1, 100% of the close users as well as 100% of the far ones respect their QoS target with a final PLR_k smaller than ST_{target}. Consequently 100% of MSs of the first group as well as 100% of MSs of the second group are satisfied and the satisfaction probability ($P_{satisfaction}$) guaranteed by PF scheduler is equal to 1. In contrast, the performance of MaxSNR scheduler is better than that of PF. The result corroborates that in contrast with the static location scenario, the MaxSNR scheduler became an efficient scheduler when users are in dynamic location. Furthermore, DFMaxSNR scheduler obtains a good performance and guarantees a high QoS and high satisfaction degree to all MSs. it appears that for both values of ST_{target}, all MSs respect their QoS target with a final PLR_k smaller than ST_{target}. Consequently 100% of users are satisfied and the satisfaction probability ($P_{satisfaction}$) guaranteed by these schedulers (MaxSNR and DFMaxSNR) is equal to 1.

V. Conclusion

In multiuser environments, it is important to design an efficient resource allocation algorithm. In this paper, we studied the RUs scheduling issue of IEEE 802.16 OFDMA systems. Our goal consists of improving the downlink capacity by reducing the delay and the packet loss rate of MSs. As an efficient scheme to the resource allocation, we designed a heuristic scheduling algorithm called Dynamic and Fair MaxSNR (DFMaxSNR). Based on a system of weights, this scheduler introduces dynamic priorities between the MSs according to both channel state and buffer status. DFMaxSNR consists not only of exploiting the concept of opportunistic scheduler but also of correcting the unequal spectral efficiencies induced by path loss attenuations for MSs with different geographical positions.

To show the effectiveness of our proposed scheme, we have compared the DFMaxSNR with the best well-known opportunistic schedulers (MaxSNR and PF) and the most common classic RR scheduler.

Results shows that the DFMaxSNR scheduler introduces remarkable multi-objective improvement of the QoS performance parameters, such us, packet loss rate, average throughput, system spectral efficiency and satisfaction probability ($P_{satisfaction}$).

References

[1] N. Varanese, J. L. Vicario, U. Spagnolini. On the asymptotic throughput of OFDMA systems with best-M CQI feedback. *IEEE Wireless Commun. Lett.*, vol. 1, pp. 145-148, June 2012.

[2] A.I. Sulyman. Throughput Gain using Threshold-based Multiuser Scheduling in WiMAX OFDMA. *EURASIP Journal on Wireless Communication Network*, 2011.

[3] C. L.Tan, Z.Zhu and W.Zhang. *A novel approach for bandwidth allocation among soft QoS traffic in wireless networks*, 2014.

[4] A. Kurne and A. Miettinen. Weighted round robin scheduling strategies in (e)gprs radio interface. *in Proc. IEEE Int. Vehicular Technology Conference(VTC)*,vol. 5, pp. 3155 3159, 2004.

[5] X. Yuan and Z. Duan. Fair round-robin : A low complexity packet scheduler with proportional and worst-case fairness. *IEEE Transactions on Computers*, vol.58, pp. 365 - 379, March, 2009.

[6] C.Gueguen and S.Baey. A fair maxsnr scheduling scheme for multiuser OFDM wireless systems. *In Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Tokyo, Japan, September,2009.

[7] C.Gueguen and S.Baey. Weighted fair opportunistic scheduling for multimedia qos support in multiuser ofdm wireless networks. *In Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Cannes, France, September, 2008.

[8] F. L.Tan, Z.Zhu and N.Xiong. maximization ressourse allocation in wireless networks: Methods and algorithms. *IEEE Transactions on System, Man and Cybernetics System*, 2015.

[9] A. Iyanda, I. Ahmad, H. Hassanein, S.A. Alshebeili. Capacity analysis of threshold-based SNR scheduler in LTE systems. *EURASIP Journal on Advances in Signal Processing*, 2013.

[10] C.Gueguen and S.Baey. Compensated proportional fair scheduling in multiuser OFDM wireless networks. *IEEE Int. Wireless and Mobile Computing, Networking and Communications (WIMOB)*, Avignon, France, octobre, 2008.

[11] 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.

[12] IEEEStd802.16-2004. *IEEE standard for local and metropolitan area networks, part 16: Air interface for fixed broadband wireless access systems*, IEEE Std 802.16-2004, Oct. 2004.

[13] S. Baey. Modeling mpeg4 video traffic based on a customization of the dbmap. *Int. Symposium on Performance Evaluation of Computer and Telecommunication System (SPECTS)*, July 2004.

[14] P. Brady. A model for generating on-off speech patterns in two-way conversation. *Bell System Technical Journal*, 1969.

[15] H. Hamouda, M.O. KABAOU and M.S. Bouhlel. A Well-balanced Resource Allocation Algorithm with Reducing Buffer Occupancy For Streaming Applications over OFDMA Systems. *The 31st IEEE International Conference on Advanced Information Networking and Applications (IEEE AINA)*, Taipei, Taiwan, 2017.