Characterization of Wireless Fair Service Extensions for Packet Scheduling in OFDMA

Carle Lengoumbi, Philippe Martins, Philippe Godlewski
Telecom-Paris, Ecole Nationale Supérieure des Télécommunications
LTCI-UMR 5141 CNRS, 46 rue Barrault 75013 Paris, France
Email:{carle.lengoumbi, philippe.martins, philippe.godlewski }@enst.fr

***Abstract*—We investigate packet scheduling and subchannel allocation in the downlink of OFDMA systems. The Wireless Fair Service (WFS) has been designed to provide fair scheduling and handle various traffic simultaneously. In this paper, WFS is extended for OFDMA i.e. to support multiple subchannels and link adaptation schemes. Enhanced and opportunist extensions (resp. EWFS and OWFS) are defined to take account of subchannels' state. Both schedulers are characterized regarding the key features of WFS.**

***Index terms*—EWFS, OFDMA, OWFS, Packet Scheduling, WFS.**

I. INTRODUCTION

Next broadband wireless communication systems are likely to use OFDMA (Orthogonal Frequency Division Multiple Access) as multiple access technique. In OFDMA, distinct subset of subcarriers, called subchannels, are allocated to users based on their channel's response. OFDMA exploits multi-user diversity since a subchannel in deep fade for a user can be good for another. Once each subchannel is assigned, AMC (Adaptive Modulation and Coding) is assumed to enhance spectral efficiency. Other OFDMA advantages are inherent to the well-known multicarrier transmission scheme OFDM (Orthogonal Frequency Division Multiplexing), namely high data rate transmissions and immunity to inter-symbol interference.

Packetized General Processor Sharing (PGPS) which provides fair packet scheduling, is described in [1]. It belongs to the class of wireline algorithms ([1]-[3]) which approximate the GPS (General Processor Sharing) policy and try to maintain proportional fairness during packet scheduling. The service of flow i, in GPS policy, is relative to its bandwidth share r_i (also called rate weight). The Wireless Fair Service (WFS) is introduced ([4]) to support various kind of traffic and provide fair packet scheduling in an error prone channel. During the error free service, flow i is scheduled according to a rate weight r_i and a delay weight Φ_i to decouple bandwidth and delay. The channel state can be either good (the flow transmits with a single rate) or bad (the flow can not transmit at all). When a flow is scheduled and has a bad channel state, another flow can transmit. A flow is called leading (resp. lagging) as it transmits more (resp. less) than the corresponding flow without any channel errors. The compensation model decides how lagging flows catch up their lag from leading flows. WFS provides interesting properties. One is graceful service degradation (GSD) to leading flows. It avoids leading flows starving until lagging flows catch up their lag. GSD contributes to short term fairness. Delay bounds and long-term throughput are provided to lagging flows. Besides, WFS supports both error sensitive and delay sensitive flows.

The main limitation of WFS is the channel error model. Considering the channel either good or bad is too pessimistic. Link adaptation schemes allow to change both modulation and coding rate according to the channel quality and maintain a fixed BER (Bit Error Rate). Some algorithms ([5]-[6]) have been proposed to take a single multirate channel into account. In [5], the CS-WFQ (Channel State Independent Wireless Fair Queing) modifies the rate weight of a flow according to its current goodput. A flow with a low goodput has a large rate weight and thus receives more transmission opportunities to catch its lag. However, the channel state of such a flow is not checked before transmission. In [6], the MWFS (Multirate Wireless Fair Scheduling) includes each flow transmission rate into the scheduling decision. Unlike WFS, MWFS only uses one rate weight. In [7], multi-user OFDM systems are considered and the CPLD-PGPS (Channel-Condition Packet Length Dependent Packetized General Processor Sharing) is proposed. PGPS scheduling with channel gain considerations is followed by resource allocation. The latter is realized through an optimization algorithm. However, link adaptation schemes are not considered in [7].

In this paper, fair packet scheduling is adapted to the context of OFDMA To benefit from (i) bandwidth and delay decoupling and (ii) AMC spectral efficiency, we define two WFS extensions called Enhanced Wireless Fair Service (EWFS) and

Opportunist Wireless Fair Service (OWFS, [8]). Here, we do not use complex optimization algorithms for resource allocation. Once scheduled, a user transmits on the best available subchannel. For simplicity, one user is associated with one flow. In [4], specific exemples are presented to illustrate WFS fundamental properties. In this paper, we characterize OWFS and EWFS regarding these properties. We focus on the decoupling of rate and delay, graceful service degradation, support of error and delay sensitive flows and influence of delay weight.

The remainder of the paper is organized as follows. Section II presents error free service models of WFS, MWFS, EWFS and OWFS whereas section III describes their compensation models. In Section IV, simulation results illustrate the fundamental properties of WFS and chararcterize EWFS and OWFS. Section V concludes the paper.

II. Error free service models

In WFS ([4]), each flow i is assigned two weights (r_i, Φ_i) , the first for rate and the second for delay. It provides rate and delay decoupling as illustrated in IV. When a packet p^k_i arrives, at virtual time $V(A(p^k_i))$, a virtual start tag and a virtual finish tag are respectively computed as:

$$S(p^k_i) = \max\ (V(A(p^k_i)), S(p_i^{k-1}) + L_i^{k-1}/r_i) \quad (1)$$

$$F(p^k_i) = S(p^k_i) + L_i^k/\Phi_i \quad (2)$$

Regarding the virtual time, $dV/dt = C(t) / \Sigma_{i\in B(t)}\, r_i$, $C(t)$ is the server capacity at time t and $B(t)$ is the set of flows that have packets to send. WFS schedules the flow with the minimum finish tag among flows whose start tag is less than $V(t)+\rho$; the parameter ρ is called the lookahead.

MWFS ([6]) considers a multirate system with M possible transmission modes on a single channel. MWFS schedules the flow with the minimum start tag, a flow i receives service in proportion to its transmission mode m_i. In (3)/(4), MWFS only considers one weight and can not provide rate and delay decoupling. In [6], how virtual time is updated is not explicitly mentioned.

$$S(p^k_i) = \max\ (V(A(p^k_i)), F(p_i^{k-1})) \quad (3)$$

$$F(p^k_i) = S(p^k_i) + L_i^k/\ (r_i\ m_i) \quad (4)$$

In this paper, we consider an OFDMA system with S subchannels and M transmission modes. Unlike MWFS, EWFS and OWFS consider multiple subchannels. Let $m_{i,best}$ be the transmission mode on flow i best available subchannel and let $m_{i,max}$ be flow i highest transmission mode on all subchannels. As MWFS, EWFS modifies the finish tag according to the channel state of the flow. However, EWFS uses the same start tag than WFS to provide rate and delay decoupling.

$$S(p^k_i) = \max\ (V(A(p^k_i)), S(p_i^{k-1}) + L_i^{k-1}/r_i) \quad (5)$$

$$F(p^{HOL}_i) = S(p^{HOL}_i) + L_i^{HOL}/\ (\Phi_i\ m_{i,best}) \quad (6)$$

The finish tag is only computed for head of line (HOL) packets. EWFS schedules the flow with the minimum finish tag among flows whose start tag is less than $V(t)+\rho$. EWFS schedules flows until all subchannels are allocated. The finish tag is expressed depending on the best subchannel available. Once a flow is scheduled, it receives the best subchannel available.

Compared to WFS, OWFS ([8]) modifies both start and finish tags. Equation (6) still holds but in OWFS, (5) becomes (7). When start tag is computed (at packet arrival), subchannels are not yet allocated; each flow highest mode is used in (7) to favour ones with good channel quality.

$$S(p^k_i) = \max\ (V(A(p^k_i)), S(p_i^{k-1})+L_i^{k-1}/(r_i\ m_{i,max})) \quad (7)$$

In OWFS, flows with good channel quality are favoured as soon as packet arrives whereas in EWFS, the channel state impacts flows only at the scheduling instant. Perfect channel state information (CSI) is assumed in the paper. In practice, feedback channels can be used to report channel quality to the base station (BS). The capacity of the system is approximated as the average capacity $C_{avg}(t)$ of a user (on all subchannels); it is used to update virtual time ([9]). In equation (4) (resp. (6)/(7)), the mode m_i (resp. $m_{i,best}$ / $m_{i,max}$) used in practice is max (m_i,1) (resp. max ($m_{i,best}$,1) / max($m_{i,max}$,1)) so that the finish or start tags can not be infinite (when m_i=0 / $m_{i,best}$=0). If a flow is scheduled and can not transmit then compensation will be launched.

III. Compensation models

In WFS, a flow that can not transmit due to channel errors, can be replaced by another flow. WFS defines *logical slots* (refered to as slots in the sequel) which represent transmission opportunities. In fact, when a packet arrives, a slot is created; the latter is tagged instead of the packet. If the BS has to destroy a flow's packet for timer expiration or excess of retransmissions, the slot is not destroyed so the flow's transmission opportunity is saved. Each flow has two positive counter E and G. The lead counter $E(j)$ of flow j is incremented when it receives the slot of another flow i and flow j can transmit. In this case, the lag counter $G(i)$ of flow i

is incremented if the lead counter is zero, otherwise the lead counter is decreased. A flow i is leading if $E(i)>0$; it is lagging if $G(i)>0$ else the flow is said *in sync*. Leading flows must release a proportion $E(i)/E_{\max}(i)$ of their slots where $E_{\max}(i)$ is flow i leading bound. Slots released, by a flow that can not transmit or by a leading flow, are given in priority to lagging flows in a WRR (weighted round robin) fashion (the different weights are given by lag counters). After lagging flows, priority to receive a released slot is given to leading flows satisfying $E(i)<E_{\max}(i)$, followed by *in sync* flows and then by leading flows with $E(i)=E_{\max}(i)$. If any of these flows has a good channel, the transmission opportunity is lost.

MWFS does not use the same compensation model than WFS. A flow has one counter E, if positive the flow is leading otherwise it is lagging. A lagging flows unable to transmit gives the opportunity to the flow with the smallest start tag and a good channel. The authors define a compensation tag given by $L(p_i^{HOL})/(E(i)\, m_i)$. The lagging flow with the smallest compensation tag receives a slot released by a leading flow.

EWFS and OWFS have the same compensation model. A flow has one counter E, if positive the flow is leading, if null the flow is *in sync*, otherwise it is lagging. Main WFS rules are kept; the differences with WFS come from multiple subchannels considerations. The priority order to receive a released slot (lagging, leading and in sync) is unchanged. Lagging flows do not receive compensation in WRR manner. In each category (lagging, leading or in sync), the flow with the highest mode on the best subchannel available is chosen. The way to break ties among lagging flows depends on the kind of traffic. When there are only NRT lagging flows, the one with the highest lag is chosen to avoid TCP retransmissions. When there are only RT lagging flows, the flows with the lowest lag is chosen to maintain quality of RT connexions. Indeed, RT connexions with high lag are likely to be cut. Besides, RT traffic flows have high delay weight to achieve time constraints. When there are both RT and NRT lagging flows, if the highest lag is NRT, it is chosen otherwise the lowest lag RT is chosen.

IV. SIMULATION RESULTS

In this section, we illustrate WFS decoupling of rate and delay and graceful service degradation. EWFS and OWFS are examined regarding these properties. Then, we compare EWFS and OWFS regarding the support of RT and NRT flows and the influence of delay weight.

A. Rate and delay decoupling

As in [4] (example 1b), we consider 3 Poisson sources with error free channels. Source 1 has average arrival rate 0.11, sources 2 and 3 have average rate of 0.44 each. We set r_1=0.11, r_2=0.44, r_3=0.44, Φ_1=0.9, Φ_2=0.09, Φ_3=0.009 and $\rho=\infty$. Performances mesures are W: ratio between the number of transmitted packets of the flow and the total number of transmitted packets; D_{max} and D_{avg}: maximum and average delay of successfully transmitted packets; σ_D: standard deviation of delay; d^{nq}: maximum new queue delay (i.e. delay of a packet arrived in an empty buffer). Performances are measured over an entire run of 50000 time units averaged over 25 simulation runs (I) and over short time windows (II), 5 small windows of 200 time units per simulation averaged over 5 simulations.

In Table I, though source 1 has the smallest rate weight and because it has the largest delay weight, source 1 experiences the smallest delay. Source 3, despite a large rate weight, has a large delay. This shows that rate and delay are decoupled in WFS. EWFS and OWFS behave exactly the same when there is one single channel and one transmission mode (M=1, S=1). Table II (resp. Table III) shows EWFS (resp. OWFS) results when there are multiple channels and transmission modes (M=5, S=2). In both algorithms, rate and delay are still decoupled. Maximum delays and delay standard deviation are higher with EWFS than with OWFS.

Table I: WFS, EWFS (M=1, S=1), OWFS (M=1, S=1)

	Src	W	D_{avg}	D_{max}	d^{nq}	σ_D
I	1	0.11	0.6	9.3	1.2	0.5
	2	0.44	1.1	8.8	4.5	1
	3	0.44	10.1	77.7	28.7	10.7
II	1	0.11	0.6	1.2	1	0.3
	2	0.45	1	3.7	2.2	0.8
	3	0.43	8.4	21	6.6	8.8

Table II: EWFS (M=5, S=2)

	Src	W	D_{avg}	D_{max}	d^{nq}	σ_D
I	1	0.11	1.3	36.5	7.1	2.7
	2	0.44	2.7	28.1	6.7	3.2
	3	0.44	12	52.1	12.6	7.4
II	1	0.11	1.3	6.5	2.8	2.3
	2	0.45	2.5	8.4	2.7	2.4
	3	0.43	10.5	20.9	4.2	6.6

Table III: OWFS (M=5, S=2)

	Src	W	D_{avg}	D_{max}	d^{nq}	σ_D
I	1	0.11	1.2	15.5	10	1.4
	2	0.44	2	14.9	7	1.8
	3	0.44	9.2	29.5	14	4.7
II	1	0.11	1.07	5	3.8	1
	2	0.44	1.78	7	3.2	1.6
	3	0.44	8.65	18	5.5	4.6

B. *Graceful service degradation*

We consider 3 flows with identical rate and delay weight ($r_1 = r_2 = r_3 = 0.33$, $\Phi_1 = \Phi_2 = \Phi_3 = 0.33$). First M=1, S=1. Flow 1 is in error until t=100. Flow 2 and 3 are always error free. While flow 1 is in error, flow 3 receives extra slots and becomes leading. After t=100, leading flow 1 must release a proportion $E(1)/E_{max}(1)$ of slots. This provides an exponential service degradation as can be seen in Fig.1. In Fig.2, flow 1 is starved until t=100, then receives exponential compensation. At t=500, all flows have transmitted the same number of packets. Temporal fairness is, in this case, equivalent to throughput fairness. Fig.3 and Fig.4 illustrate the difference between temporal fairness and throughput fairness in a multirate context. Exponential service degradation is provided in EWFS and OWFS (results similar to Fig. 1 can be obtained). At high loads, altough temporal fairness is achieved (after t =600), throughput fairness can not be achieved in OWFS. Bad channel conditions increase the value of start tags and finish tags. At high loads, even if the channel becomes better, tags of new packets are penalized by important values of already buffered packets' tags.

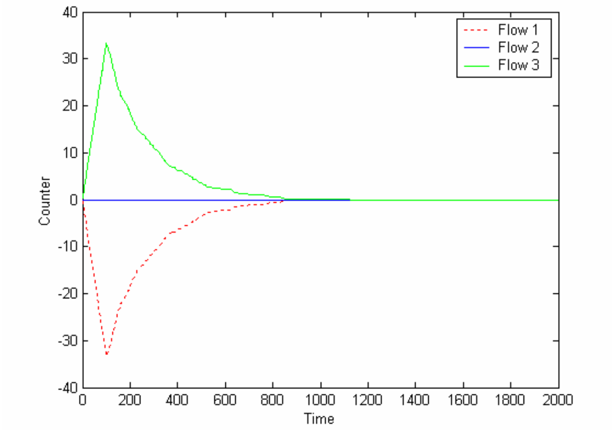

Figure 1: Counters evolution in WFS

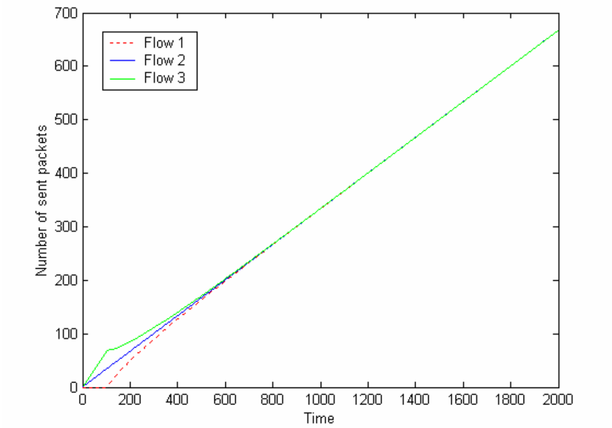

Figure 2: Throughput in WFS/ EWFS, OWFS (M=1, S=1)

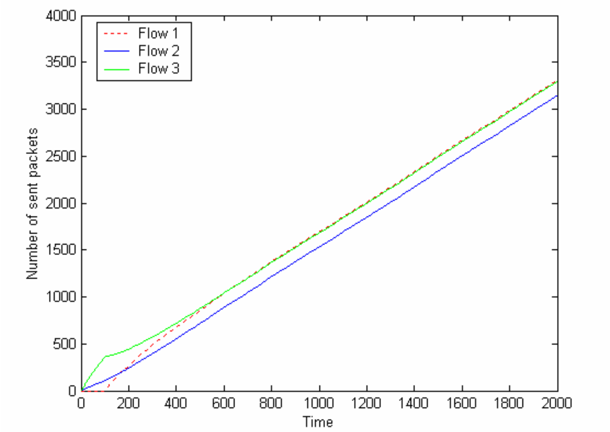

Figure 3: Throughput *in* EWFS (M=5, S=2)

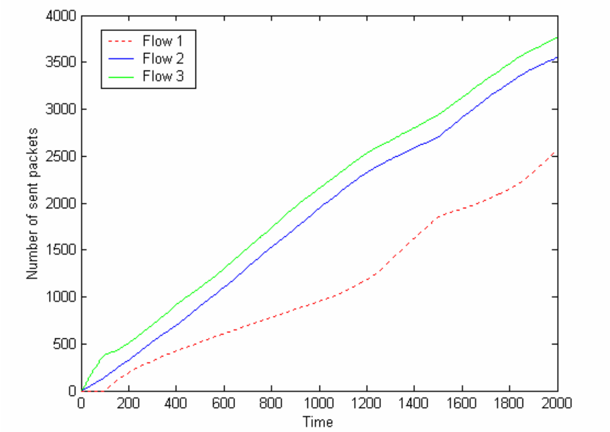

Figure 4: Throughput in OWFS (M=5, S=2)

C. *Support of RT and NRT flows*

All sources are Poisson. We consider 14 RT flows whose source rate is 64kbps. A packet whose delay exceeds 100 ms is dropped. Ratio between RT and NRT delay weight is one. The number of NRT sources increases from 10 to 20; their source rate is 384 kbps. Results are obtained over 10000 time units and averaged over 15 simulation runs. The NRT throughput (cf. Fig.5) is higher in OWFS than in EWFS since it exploits channel quality as soon as packets arrive. In fact, NRT packets are significantly delayed in EWFS compared to OWFS. In this case, NRT packets are not delayed because of their delay weight (indeed Φ_{RT}/Φ_{NRT}=1). It is due to the rate difference between RT and NRT. In EWFS, high rate sources suffer from crowded queues (delays are high so throughput is low) whereas in OWFS such sources send more packets thanks to presence of channel state into start tags. RT packets have higher delay in OWFS than in EWFS. At arrival, packets may be delayed because of channel quality (a bad channel increases the start tag). It can be seen in Fig.6 where OWFS RT drop rate becomes large when U_{NRT} =16 whereas EWFS RT drop rate stays low.

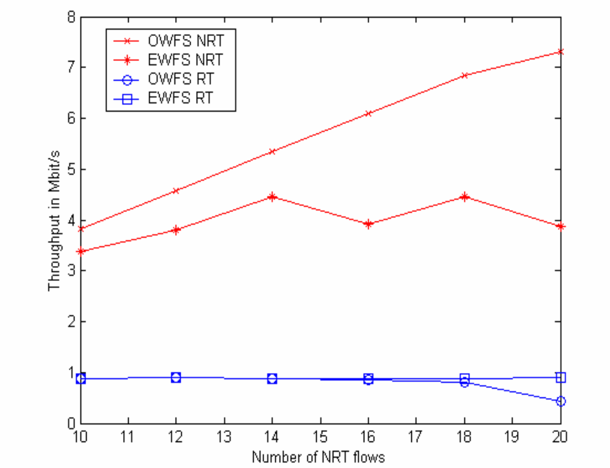

Figure 5: RT and NRT throughput

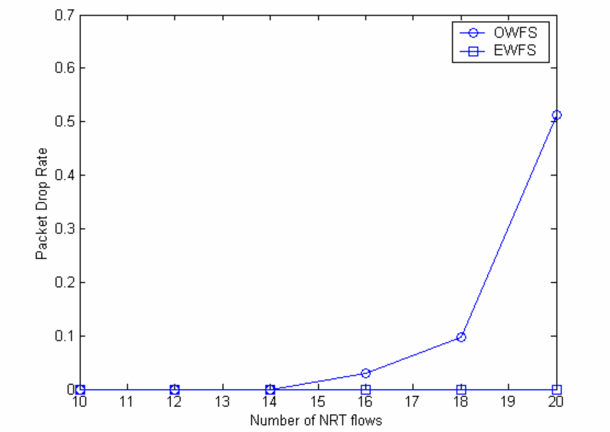

Figure 6: RT packet drop rate

D. *Influence of Delay Weight*

In this section, U_{RT} =14 and U_{NRT} =16. The ratio between RT and NRT delay weight increases from 1 to 20. It can be seen on Fig.7, that the OWFS drop rate decreases with the increase of the delay weight. EWFS drop rate does not vary much with the delay weight. OWFS drop rate becomes better than EWFS when the ratio between RT and NRT delay weight is higher than 6.

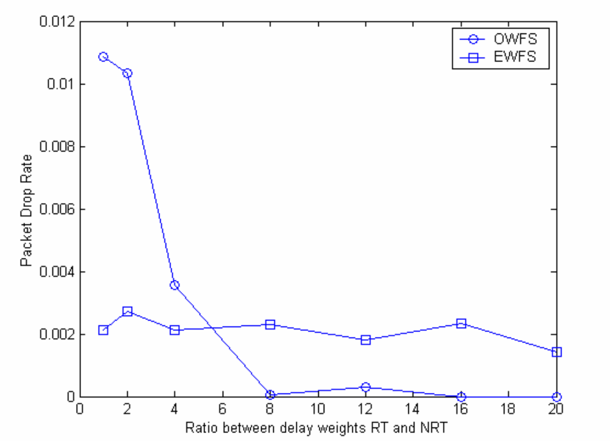

Figure 7: RT drop rate

V. CONCLUSION

In this paper, we have studied fair scheduling algorithms and illustrated WFS properties. We defined EWFS and OWFS, two extensions for OFDMA context. We saw that these algorithms present rate and delay decoupling, graceful service degradation and temporal fairness. However, in a multirate context, these algorithms present different degree of throughput fairness depending on the load. At high loads, EWFS achieve more throughput fairness than OWFS at the cost of global throughput performance. Indeed, OWFS presents good global throughput performances. If RT traffic is present, a delay weight adjustment in OWFS achieves a packet drop rate measure as good as EWFS.

VI. RFERENCES

[1] K. Parekh and R. G. Gallager, “A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The single-node case,” IEEE/ACM Transactions on Networking, Vol. 1, No. 3, June 1993, pp. 344—357.

[2] S. J. Golestani, “A Self-Clocked Fair Queueing Scheme for Broadband Applications,” Proceedings of IEEE INFOCOM, Toronto, Canada, Jun. 1994.

[3] P. Goyal, H.M. Vin, and H. Cheng, “Start-Time Fair Queueing: A Scheduling Algorithm for Integrated Services Packet Switching Networks,” IEEE/ACM Transactions on Networking, Vol. 5, No. 5, Oct. 1997, pp. 690-704.

[4] Lu S. , Nandagopal, T., and Bharghavan, V. “A wireless fair service algorithm for packet cellular networks”. In Proceedings of the 4th Annual ACM/IEEE international Conference on Mobile Computing and Networking (Dallas, Texas, United States, October 25 - 30, 1998). W. P. Osborne and D. Moghe, Eds. MobiCom '98. ACM Press, New York, NY, 10-20.

[5] P. Lin, B. Bensaou, Q.L. Ding, K.C. Chua. “A wireless fair scheduling algorithm for error-prone wireless channels”. International Workshop on Wireless Mobile Multimedia. Proceedings of the 3rd ACM international workshop in Wireless mobile multimedia. Year of Publication:2000. Pages: 11 – 20.

[6] Yuan, Y.; Gu, D.; Arbaugh, W.; Zhang, J., "Achieving Fair Scheduling Over Error-Prone Channels in Multirate WLANs", International Conference on Wireless Networks, Communicaitons and Mobile Computing *(WIRELESSCOM)*, Vol. 1, pp. 698-703, June 2005.

[7] Zhifeng Diao Dongxu Shen Li, V.O.K.“CPLD-PGPS scheduling algorithm in wireless OFDM systems». Global Telecommunications Conference, GLOBECOM '04. Date: 29 Nov.-3 Dec. 2004. Volume: 6, On page(s): 3732- 3736 Vol.6.

[8] C. Lengoumbi, P. Martins, P. Godlewski. “ An Opportunist Extension of Wireless Fair Service for Packet Scheduling in OFDMA”. IEEE 65th International Vehicular Technology Conference VTC2007-Spring. Pages 3001-3005. April 2007.

[9] Yang-Chung Tseng and Cheng-Sang Chang. «PGPS servers with Time-Varying Capacities». IEEE Communications Letters, Vol.2, No 9, september 1998.