Access Scheduling Based on Time Water-Filling for Next Generation Wireless LANs

Ertugrul Necdet Ciftcioglu†, Ozgur Gurbuz‡
†The Pennsylvania State University, University Park, PA 16802, ‡Sabanci University, Istanbul, 34956 Turkey
enc118@psu.edu, ogurbuz@sabanciuniv.edu

Abstract- **Opportunistic user access scheduling enhances the capacity of wireless networks by exploiting the multi user diversity. When frame aggregation is used, opportunistic schemes are no longer optimal, since users with high capacity links are frequently served, causing small queue sizes and low throughput. Recently, we have proposed schedulers that take queue and channel conditions into account jointly, to maximize the instantaneous throughput. In this paper, we extend this work to design a scheduler that performs block scheduling for maximizing network throughput over multiple transmission sequences. This scheduler makes use of the estimated evolution of the aggregation process by queueing theory and determines users' temporal access proportions using an approach based on the water-filling principle. Through detailed simulations, we show that our new algorithm with block scheduling offers further improvement in throughput over the previous schedulers, along with better fairness.**

Keywords- **Wireless LANs, opportunistic scheduling, queuing theory, scheduling and statistical multiplexing**

I. INTRODUCTION

Multiple Input Multiple Output (MIMO) systems significantly improve the quality, and hence the data rate of wireless links, by utilizing multiple antennas at the transmitter and receiver ends. The emerging new standards for wireless local area networks (WLANs), defined by the IEEE 802.11, specifically 802.11n task group provision physical layer data rates exceeding 200 Mbps with the realization of MIMO technology. However, the actual throughput to be experienced by the WLAN users is considerably lower than the promised physical (PHY) layer data rates. MAC efficiency is enhanced via the method of frame aggregation [1][2], where multiple MAC layer protocol data units are transmitted in one physical frame. This reduces the relative percentage of the time loss due to packet overhead and MAC coordination, as proposed by the new IEEE 802.11e MAC specification and the draft standard of 802.11n [3].

In multi user communication systems, such as WLANs, scheduling is an essential element determining which user should transmit or receive data in a given time interval. Opportunistic scheduling algorithms maximize system throughput by making use of the channel variations and multi user diversity [4-6]. In spatially greedy scheduling schemes, named as Maximum Rate Scheduling (MRS), the selection metric is the channel capacity of the user, which prefers the user with the best channel conditions to transmit at a given time instant [4]. In Proportional Fair Queuing (PFQ), the user with the best channel capacity relative to its own average capacity is selected [6]. The main aim of PFQ is to maximize the throughput while satisfying fair resource allocation. If the users of all channels deviate from their mean capacities in similar ways, all users are to access the medium for similar time durations.

In [7], we have proposed a queue-aware scheduling method, Aggregate Opportunistic Scheduling (AOS), where we extend the opportunistic approach by considering both the queue and channel states of the users. Instead of channel capacity, AOS selects the user that maximizes the instantaneous throughput, which is a function of aggregate, i.e., queue size and channel data rate, i.e., capacity. Through simulations, it has been shown that the performance of capacity based opportunistic algorithms, MRS and PFQ, is suboptimal with frame aggregation, while the AOS algorithm significantly improves the throughput and fairness. However, selecting the user that maximizes the instantaneous throughput only for a specific transmission opportunity (TXOP) can prevent transmitting with higher efficiencies in the subsequent TXOPs, reducing the overall throughput in the long term. In this paper, we propose a new scheduler which performs block scheduling by considering the statistical evolution of the user queues so as to maximize the throughput over a longer time scale. In order to estimate the throughput in the long term, 802.11n MAC frame aggregation and evolution of the queue states are modeled by extending the bulk service model from queuing theory [8]. and determine the optimal temporal access proportions of users such that total system throughput is maximized. In this paper, we apply a water-filling like approach to obtain the optimal access proportions, propose an iterative method for computation of these proportions and a method for realizing them We name this scheduler as, Predictive Scheduling with Time-domain Water-filling (P-WF), and through simulations, we compare P-WF with our previous scheduler, AOS [7], opportunistic MRS [4] and PFQ [6], and Longest Queue (LQ) [3] algorithm, which is a non-opportunistic scheme that selects users based on queue size. P-WF promises to offer highest throughput with lower delay and better fairness.

The rest of the paper is organized as follows: In Section II, we present the system model including air interface, MAC framework and queuing formulation. In Section III, we present the proposed scheduling scheme including time proportion assignment using the queuing model and station ordering. Section IV presents our simulation model and results, and Section V involves our conclusions.

II. System and Queuing Model

2.1 Physical Layer

We consider the downlink of a MIMO wireless cellular system that consists of a single access point (AP) communicating with multiple WLAN users. The system is a closed-loop MIMO OFDM system such that the mobile users measure their channel states and send them as feedback to the AP. Based on the channel state, link capacities are calculated and 802.11n data rates are assigned at the AP according to available capacity [1]. The properties of the fading wireless channel are modeled in the channel matrix ***H***, considering *large-scale path loss*, *shadowing* and *small scale multi-path fading* effects. In this paper, the log distance path loss model and the Channel B fading channel model defined by the Task Group n (TGn) are considered. The fading characteristics between individual antenna pairs are spatially correlated and the correlation matrices depend on the angular spread. Further details of the channel model can be found in [10]. Due to low speeds of WLAN users, coherence time is large enough so that channel fading is slow, i.e. the channel is assumed stationary within one transmission opportunity.

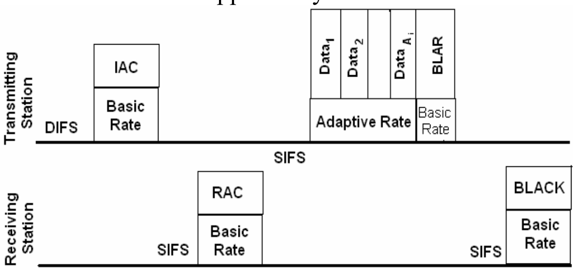

Figure 1: Example aggregate frame transmission

2.2 MAC Framework

We consider a time division system where only one user is served at a given time period, limited by a duration called transmission opportunity (TXOP). As defined by 802.11n draft standard, within a TXOP, a two-way handshake with frame aggregation can be performed as shown in Fig. 1 [3]. Initiator Aggregation Control (IAC) and Responder Aggregation Control (RAC) are RTS/CTS-like reservation messages, which also involve training sequences to help (MIMO) channel estimation and data rate selection. After IAC/RAC exchange, a number of data packets are aggregated in one frame and an acknowledgement is requested in the end via the Block ACK Request (BLAR) packet. The destination station replies with a Block ACK (BLACK) packet that contains the reception status of packets in the aggregation. The data packets are transmitted at the selected transmission rate, while the control packets (IAC, RAC, BLAR and BLACK) are transmitted at the basic rate, so that all stations can decode these packets. The inter frame spacing (DIFS, SIFS) values are as in the 802.11 specification. The throughput S^i for the i^{th} TXOP can be calculated as the data payload transmitted per transmission opportunity,

$$S_i = \frac{A_i \cdot L_P}{\frac{L_{IAC}}{r_0}+\frac{L_{RAC}}{r_0}+4.T_{PLCP}+DIFS+4.\tau+3.SIFS+\frac{L_{BLACK}}{r_0}+\frac{L_{BLAR}}{r_0}+\frac{A_i.(L_P+L_{MH})}{r_i}} \quad (1)$$

A_i is the aggregate size at transmission opportunity i; L_P, L_{IAC}, L_{RAC}, L_{BLACK}, L_{BLAR} are the length of the data, reservation, ACK and ACK request packets; L_{MH} is the MAC header in bits; TPLCP is training duration; τ is the one way propagation delay; r_0 is the basic rate and r_i is the selected data rate during data transmission.

At each TXOP, the AP transmits to a selected station using frame aggregation. Station selection is to be done according to one of the scheduling algorithms.

2.3 Queuing Model

In this section, we devise a queuing model for aggregate frame transmissions by extending the bulk service model in [8]. In the bulk service model, the packets are served collectively in groups and incoming packets are enqueued as shown in Fig. 2. Packets arrive in a Poisson fashion with an average rate of λ. All of the packets in the queue are served together if the number of packets is less than the bulk size, L. If the queue length exceeds L, only the first L packets are served. The bulk service rate, μ, is defined as the rate of serving bulks, which is assumed constant for all states [8]. This assumption implies that the service rate in terms of bits per second is increased in a proportional manner with the bulk size. This is actually not valid for transmissions over a physical link, since the channel data rate is unchanged irrespective of the bulk size. Moreover, realistic aggregate frame transmissions MAC and PHY overhead are also taken into account.

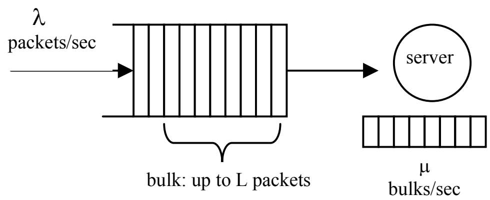

Figure 2: Bulk service system

The service rate, μ_j, for queing model of aggregate transmission, in packets/sec is obtained as:

[1] In MIMO-OFDM based systems, the channel capacity is calculated by partitioning the system into multiple sub-channels that correspond to different sub-carriers as follows [9]:

$$C=\frac{B}{N_c}\sum_{k=0}^{N_c-1}\log_2(\det(\mathbf{I}_{MN_c}+\rho\mathbf{H}(e^{j2\pi\frac{k}{N_c}})\mathbf{H}^H(e^{j2\pi\frac{k}{N_c}}))), \quad \text{with}$$

$$\mathbf{H}(e^{j2\pi\theta})=\sum_{l=0}^{L-1}\mathbf{H}_l e^{-j2\pi l\theta} \quad (N_c\text{: Number of subcarriers})$$

$$\mu_j = \begin{cases} \frac{\mu}{j} \cdot \left(\frac{j.L_P}{j.(L_P + L_{MH}) + L_{overhead} + r.T_{IFS}} \right) & 1 \le j < L, \\ \frac{\mu}{L} \cdot \left(\frac{L.L_P}{L.(L_P + L_{MH}) + L_{overhead} + r.T_{IFS}} \right) & j \ge L, \end{cases} \quad (2)$$

where j is the number of packets involved in the aggregation; μ is the rate of serving bulks; $L_{overhead}$ accounts for the total overhead including PHY ad MAC headers; T_{IFS} is the sum of interframe durations; r is the channel data rate for the current TXOP. This rate is determined according to the channel conditions which vary over time due to small scale fading.

Fig.3 depicts the Markov chain representation of this queuing model of aggregate frame transmissions, defining the state as the number of packets in the queue. Packets arrive one-by-one at rate λ and they are served at rate μ_j (Eq.2). Next, we derive the state probabilities to calculate, i.e., predict, the average queue size and throughput values.

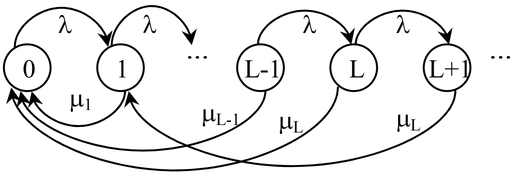

Figure 3: Markov-chain representation of aggregate frame transmission

The balance equations of the above system can be obtained to solve for the steady state probabilities of each state, i.e., p_1 p_2, ..., p_L, as follows:

$$\lambda p_0 = \mu_1 p_1 + \mu_2 p_2 + ... + \mu_L p_L \Rightarrow p_0 = (1/\lambda) \sum_{j=1}^{L} \mu_j p_j \quad (3)$$

$$(\lambda + \mu_j) p_j = \mu_L p_{j+L} + \lambda p_{j-1} \qquad 1 \le j \le L \quad (4)$$

$$(\lambda + \mu_L) p_j = \mu_L p_{j+L} + \lambda p_{j-1} \qquad j \ge L \quad (5)$$

Converting the balance equations into the alternative form by taking the z-transform, we obtain $P(z)$ in rational form, $P(z)=N(z)/D(z)$:

$$P(z) = \frac{\sum_{j=1}^{L} [z^{L+j}(\mu_j - \mu_L) - z^L(\mu_j + \frac{\mu_L \mu_j}{\lambda}) + \mu_L z^j + \frac{\mu_L \mu_j}{\lambda}] p_j}{\lambda z^{L+1} - (\lambda + \mu_L) z^L + \mu_L} \quad (6)$$

The global sum of probabilities should be equal to 1, requiring $P(1)$=1 to be satisfied. Since both $N(1)$=0 and $D(1)$=0, we need to utilize the L'Hospital rule. As such, we require that

$$\frac{N'(z)}{D'(z)} = 1 \,. \quad (7)$$

The next step is to obtain state probabilities by taking the inverse transform of $P(z)$. The fact that the bulk service rates are state-dependent has caused the order of $N(z)$ to be greater than the order of $D(z)$, so $P(z)$ cannot be simplified. We take an alternative approach as follows: Similar to the bulk service model solution in [8], out of the $(L+1)$ roots of $D(z)$, $(L-1)$ roots are located within the unit circle. Due to the fact that the z-transform of a probability distribution is analytical inside the unit circle, $P(z)$ should be bounded, which implies that $(L-1)$ zeros of $P(z)$ must also be the roots of the numerator $N(z)$. $N(z)$ must also vanish at each of the (L-1) roots of $D(z)$ inside the unit circle. This constraint results in a set of $(L-1)$ equations. Including the equation provided by Eq. (7), we obtain L equations for probabilities p_1, p_2, ..., p_L. Eq. (3) provides the solution for p_0. The set of equations is solved via numerical computations in MATLAB, obtaining the steady-state probabilities of the system for all the states up to the aggregation limit L.

Next, we find the expected aggregate size and expected throughput by weighted averaging using calculated state probabilities, as follows:

$$\overline{A} = \sum_{j=1}^{L} j.p_j + L.(1 - \sum_{j=0}^{L} p_j) \,. \quad (8)$$

Throughput is a function of aggregate size. The average throughput can be calculated by considering state probabilities of the aggregate size, as follows:

$$\overline{S} = \sum_{j=0}^{L} S(A_j).p_j + S(L).(1 - \sum_{j=0}^{L} p_j) \,, \quad (9)$$

where $S(A_j)$ is the throghput achieved with aggregate size A_j as given in (1). With this analysis, it has been observed that the throughput values are equal to the applied load until the load reaches the maximum service rate, which is the service rate for the maximum aggregate size. When the load exceeds the maximum service rate, throughput is saturated at that value.

III. Predictive Scheduling with Time Water-Filling

In this section, we utilize the queuing model of aggregate transmissions for designing a new scheduler, Predictive Scheduling with Time-domain Water-filling (P-WF). P-WF maximizes the total network throughput over a long time scale, as opposed to the previous schedulers [4-7] that consider the upcoming transmission opportunity only. Over the scheduling duration, the temporal access proportion of each user is varied with an effort to maximize the total throughput.

The devised queuing model provides us the average expected aggregate size and throughput, given the service rate and applied load for a single queue (user). Considering the multiuser scenario with time-division multiplexed traffic, the input parameters for the queuing model are assigned as: The downlink load per user is modified by dividing it by the user's access proportion in time (π_n Є [0, 1]) , to obtain the effective load, and the service rate is determined by the data rate of the served user's link. The aggregate size and throughput values are calculated through (8) and (9) for each user, and the total network throughput is obtained as the weighted average of the individual throughput values,

$$S_{total} \triangleq \sum_{n=1}^{N} \pi_n S_n \ . \tag{10}$$

π_n is the temporal proportion of access for user n, and N is the total number of users. The throughput maximization problem is described as:

$$\arg\max \sum_{n=1}^{N} \pi_n S_n \qquad \text{s.t.} \ \sum_{n=1}^{N} \pi_n = 1 \tag{11}$$

In this work, the aim is to maximize the overall throughput over a sequence of transmissions via scheduling.In order to maximize the total throughput, we perform a search by varying proportion values for all users, i.e., the proportion vector $\overline{\pi} = (\pi_1, \pi_2, ..., \pi_N)$ and computing S_{total}. P-WF aims to maximize the total throughput by applying the principle of water-filling, commonly used in information theory, to the time proportions π_n to find proportions that maximize the problem defined in (11). We call this method as "*time-domain waterfilling*". The principle of waterfilling is commonly used in the field of information theory. In waterfilling problems, the aim is to maximize a weighted average with a constraint. An example formulation is to find the optimal (x_1 ,x_2, ... x_N) in order to

$$\max \sum_{i=1}^{N} (\beta + \gamma_i x_i) \ \text{with the constraint} \ \sum_{i=1}^{N} x_i = 1 \tag{12}$$

The waterfilling solution to the problem is given as

$$x_i^{opt} = (\mu - \frac{\beta}{\gamma_i})_+ \ , \ i = 1, ..., N \ , \tag{13}$$

where $(\theta)_+$ denotes $max(\theta,0)$. Some modes may be unused.

By comparing (11) with (12), we can exploit the mathermatical analogy between these equations. Even though both the x_i and time proportions are weighting factors, (12) also includes an additive term which is crucial for the remaining of the waterfilling analysis. In order to achieve a full analogy between the equation pairs, we add a constant to each term in the summation of (10) :

$$S' \triangleq \sum_{n=1}^{N} (\beta + \pi_n S_n^i) \ , \tag{14}$$

Maximizing S' is equivalent to maximizing S_{total} , so the waterfilling solution is given as:

$$\pi_n = \left(\varsigma - \frac{\beta}{S_n} \right)_+ \tag{15}$$

We can not compute π_n values directly, since S_n depends on π_n. In order to overcome this coupling in the waterfilling terms, we apply an iterative procedure to π_n values.

According to our queueing model results, we can express S_n in terms of the per-user load λ_n, time access proportions π_n, supported data rate r_n and MAC overhead as follows:

$$S_n = f(\pi_n) = \begin{cases} \dfrac{\lambda_n}{\pi_n} \ , & \dfrac{\lambda_n}{\pi_n} < S(L) \\ S(L) \, , & \dfrac{\lambda_n}{\pi_n} > S(L) \end{cases} , \tag{16}$$

where $S(L)$ is the maximum throughput offered by using data rate r_n , which is equal to the throughput that can be achived with the highest aggregate size, L allowed.

Our iterative time-domain waterfilling algorithm can be described as follows:

i. First,initial proportions π_n^0 is initialized as $1/N$ for n=1...N.

ii. For iteration i=1,2,..I, access proportions are calculated using the formula

$$\pi_n^{i+1} = \left(\varsigma - \frac{\beta}{f(\pi_n^i)} \right)_+ \tag{17}$$

The threshold ς is also evaluated for each iteration, using the constraint

$$\sum_{i=0}^{N} \left(\varsigma - \frac{\beta}{f(\pi_n^i)} \right)_+ = 1 \tag{18}$$

After a finite number of iterations, the access proportions πn converge and are determined. These proportions indicate transmission durations of the users relative to the total transmission sequence in which scheduling is applied.

$$\overline{\pi}^* = [\pi_1^*, \pi_2^*, ..., \pi_N^*] \qquad \overline{A}^* = [A_1^*, A_2^*, ..., A_N^*] \tag{19}$$

Where the average aggeragate size, A_i^*, are calculated from the analytical model, taking into account the data rates and effective load values, $\lambda_n^e = \lambda_n^0 / \pi_n^*$ for each served user.

The next step in scheduling is to realize a sequence of transmissions over all users so as to ensure the allocation of the optimal proportions. In order to define the transmission sequence, we assign each user a *turn number,* which indicates the number of times the user will be given access throughout the total scheduling duration. The turn number is determined in two steps. First, the ratio of the access proportion of each user to the transmission duration of serving that user once is found by,

$$t_n^0 = \frac{\pi_n^*}{T_n} = \frac{\pi_n^*}{((A_n^* . L_P)/r_n + T_{overhead})} \ , \tag{20}$$

where T_n is the transmission duration of serving user n once and $T_{overhead}$ refers to the overhead terms in (1). Next, the lowest turn number of the served users is determined, and the turn numbers of other stations are scaled with respect to the minimum. In other words,

$$t' = \min_{\pi_i^* > 0} \left(\frac{\pi_1^*}{T_1}, \frac{\pi_2^*}{T_2}, ..., \frac{\pi_N^*}{T_N} \right), \ t_1 = \frac{t_1^0}{t'}, t_2 = \frac{t_2^0}{t'}, ..., t_N = \frac{t_N^0}{t'} \tag{21}$$

The transmission sequence is determined as the ascending order of calculated turn numbers, and scheduling is performed by assigning transmissions starting with the smallest turn number.

IV. PERFORMANCE ANALYSIS

We have evaluated the performance of P-WF in comparison with the aforementioned scheduling disciplines via simulations. The simulations have been carried out in the OPNET simulation environment, modeling the wireless channel, physical layer parameters, 802.11 MAC layer with 802.11n enhancements and compared scheduling algorithms. In the wireless channel, the log-distance path loss is modeled with path loss exponent of 2 within a distance of 5 meters from the transmitter and 3.5 for distances larger than 5 meters. Log-normal shadowing term is taken as 3 dB up to 5 meters and 5 dB afterwards. For the fading model, the Channel B model developed by TGnSync group for small office environments and non line-of-sight conditions is implemented with an rms delay spread of 15 ns and Doppler frequency of 5 Hz. In the physical layer, a 2x2 MIMO configuration is assumed. OFDM parameters such as guard interval, number of subcarriers etc. are chosen according to the 802.11n specifications in [3]. IEEE 802.11n data rates are adaptively selected from the set {24,36,48,72,96,108,144,192,216} Mbps according to the instantaneous channel conditions. The basic rate, i.e. the common rate for control packet transmission is selected as 24 Mbps. Finally, some of the MAC related parameters of the simulation model are given in Table I. The maximum number of packets allowed in frame aggregation, L, is assumed as 63. The downlink traffic is modeled by fixed size (1024 bytes) packets that arrive due to the Poisson distribution. We assumed similar load levels for all stations. In our simulations, we used topologies with an AP and 12 stations distributed in an area with a radius of 25 m from the AP.

We first evaluate the scheduling algorithms under varying load, changing the network load between 50 Mbps and 200 Mbps. Our results in terms of network throughput can be seen in Fig. 4. For all scheduling disciplines, the throughput and load is about the same as long as the total load is below the "network service rate", which depends on the physical medium- i.e. maximum data rates that can be supported, and the MAC efficiency.

The LQ algorithm outperforms MRS using frame aggregation since with MRS, stations with good channel states are served more frequently without filling their queues, leading to low aggregate sizes and reduced throughput by (1). When the arrival rate exceeds the network service rate, the total throughputs of the scheduling algorithms start to deviate from each other.

TABLE I
SOME MAC RELATED PARAMETERS

Parameter	Value
SIFS	16 μ sec= 16 X 10^{-6} sec.
DIFS	34 μ sec= 34 X 10^{-6} sec.
PLCP overhead	44.8 μ sec= 448 X 10^{-7} sec.
T_{IAC}	11.2 μ sec = 112 X 10^{-7} sec.
T_{RAC}	8.7 μ sec = 87 X 10^{-7} sec.
T_{BLACK}	48.7 μ sec = 487 X 10^{-7} sec.
T_{BLAR}	9 μ sec = 90 X 10-7 sec.

P-WF algorithm significantly outperforms all opportunistic algorithms, by 50 % over PFQ [4], 39 % over MRS [3] and 4% over AOS [1], and by 25 % over non-opportunistic LQ [2].

In Fig. 5 delay performance of the schedulers is presented, where the sample mean of average delays experienced by each user is plotted as a function of increasing load. The simulations have been carried out for 5 seconds. MRS performs poorly since some users may never be selected due to poor channel conditions and low link capacity. Delay performance of AOS is similar to PFQ and LQ, since despite offering higher throughput, AOS may also cause some users to starve.

In order to evaluate fairness, we employ our measure of unfairness, which is the ratio of the standard deviation of station throughputs to the mean value of station throughput, $UF=\sigma/S_{av}$. The fairness performance of the algorithms under varying load can be seen in Fig. 6. The queue based LQ algorithm performs best in terms of fairness, since it does not consider the channel conditions. Likewise, the MRS algorithm is poorest, since it takes only the channel conditions into account. AOS significantly improves MRS since both channel and queue states are considered; P-WF improves fairness of our previous algorithm, AOS, approaching the performance of PFQ.

Finally, we investigate the performances of the scheduling approaches with different topologies. In Fig. 7, the total throughputs for all algorithms are depicted for five different uniformly distributed topologies. The aggregate load is set as 200 Mbps, the maximum aggregation size is again 63. Our results indicate that the relative performances of all the algorithms are similar to previous results for the different topologies. P-WF consistently outperforms other algorithms since it can avoid low capacity users if necessary.

V. CONCLUSIONS

In this paper, we propose a new scheduling algorithm, which exploits multi user channel diversity and queue diversity to maximize the throughput of WLANs. A queuing model is developed for frame aggregation mode of next generation WLANs, and later utilized in throughput maximization. An optimization approach is also proposed for the scheduler. Through detailed simulations, we have shown that P-WF significantly outperforms the existing schedulers MRS, PFQ and LQ, and offers better fairness and delay performance as compared to AOS, justifying the concept that selecting the user which maximizes the instantaneous scheduling metric may not provide maximum performance throughout the entire time duration.

VI. REFERENCES

[1] I. Tinnirello and S. Choi, "Efficiency Analysis of Burst Transmissions with Block ACK in Contention-Based 802.11e WLANs," *in Proc. IEEE ICC'2005*, Seoul, Korea, May 16-20, 2005.

[2] C. Liu and A. P. Stephens.; "An analytic model for infrastructure WLAN capacity with bidirectional frame aggregation," *Wireless Communications*

and Networking Conference, 2005 IEEE Volume 1, 13-17 March 2005 Page(s):113 - 119 Vol. 1.

[3] S. A. Mujtaba, "TGn Sync Proposal Technical Specification 3," *TGn Sync Technical Proposal R00*,August 13, 2004.

[4] R. Knopp and P. Humblet ,"Information capacity and Power Control in Single Cell Multi-user Communications," *Proc. IEEE ICC'95*, Seattle, USA, vol. 1., pp. 331-335, June 1995.

[5] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. On Info. Theory,* Vol.48, No. 6, Page(s): 1277-1294, June,2002.

[6] A. Jalali, R. Padovani, R. Pankaj, "Data Throughput of CDMA-HDR: A High Efficiency-High Data Rate Personal Communication Wireless System," *Proc. IEEE VTC Spring '00*, Tokyo, Japan, pp. 1854-1858, May 2000.

[7] E.N. Ciftcioglu, O. Gurbuz ,"Opportunistic Scheduling with Frame Aggregation for Next Genaration Wireless LANs," IEEE *International Conference on Communications* (ICC) 2006, Istanbul, Turkey, June 2006

[8] Kleinrock, L; "Queuing Systems , Volume I:Theory", Wiley-Interscience, 1975

[9] H. Bolcskei, D. Gesbert and A. J. Paulraj ,"On the Capacity of OFDM-Based Spatial Multiplexing Systems," *IEEE Transactions on Communications*, no. 2, Feb 2002 pp. 225-234.

[10] V. Erceg, and et al., "TGn Channel Models," *IEEE 802.11 - 03/940r4*, May 2004.

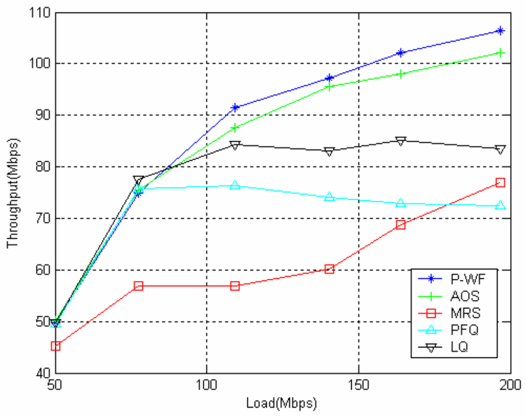

Figure 4: Performance of schedulers with varying load

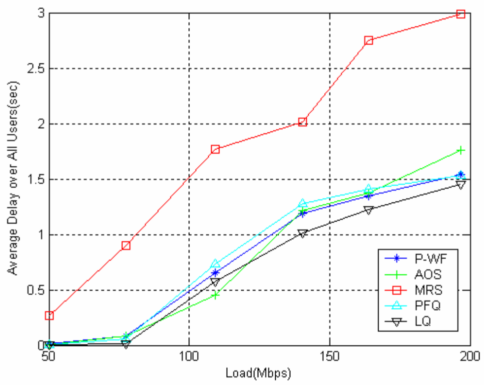

Figure 5: Mean user delay vs load.

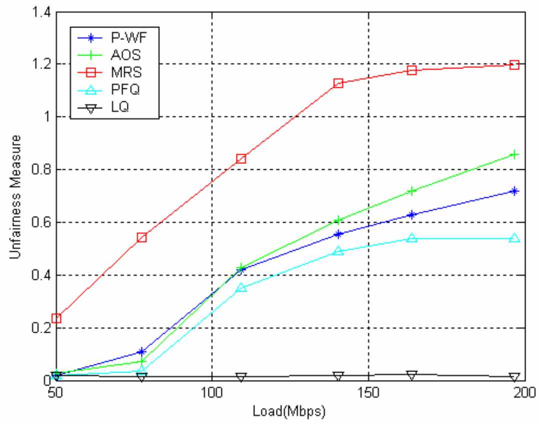

Figure 6: Fairness performance of proposed and existing schedulers

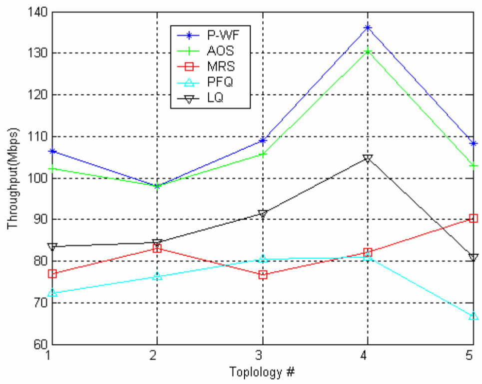

Figure 7: Performance of the schedulers with different topologies