

Access Point Selection for Improving Throughput Fairness in Wireless LANs

Vasilios A. Siris and Despina Evaggelatou
Institute of Computer Science, FORTH and
Department of Computer Science, Univ. of Crete
P. O. Box 1385, GR 711 10 Heraklion, Crete, Greece
Email: vsiris@ics.forth.gr

Abstract— We investigate the problem of access point selection in wireless LANs based on the IEEE 802.11 standard, when a station is within the vicinity of more than one access points. According to the proposed approach, the selection is based on the packet transmission delay for each access point that a station can associate to, which is related to the throughput that the station can achieve. Important features of the approach is that it considers the contention-based nature of IEEE 802.11's MAC layer, it can be applied to the 802.11e standard which can support different classes in terms of the minimum contention window, and it can be implemented solely at the wireless stations, which passively monitor the activity of each access point's channel, without requiring modifications to the access points. Experiments show that the proposed approach can achieve significantly higher throughput fairness compared to other approaches, without a significant decrease of the aggregate throughput.

I. INTRODUCTION

As wireless networks become more ubiquitous, efficient sharing of wireless resources will become increasingly important, especially since the capacity of wireless networks is at least an order of magnitude smaller than that of wired networks and is limited by spectrum availability. Additionally, wireless networks will be used by applications with different bandwidth and delay requirements, thus making the support for quality of service important. In current wireless hotspots and enterprise wireless networks it is common that a station is within the vicinity of more than one access points; this will increasingly be the case in the future, with the deployment of dense wireless LANs with enterprise or city-wide coverage. Up to now, the selection of an access point is based on the received signal strength of beacon frames transmitted by the available access points. However, such an approach does not take into account the actual load of each access point, and can result in an unbalanced distribution of stations to access points.

In this paper we propose and investigate an approach for access point selection based on the average packet delay, which determines the average throughput that a station would receive if it connected to a particular access point. The proposed approach can be used for self-configuration of stations in wireless LANs, in a manner that improves fairness of the throughput the

wireless stations obtain. Such self-configuration capabilities are significant, due to the distributed and ad hoc deployment of wireless access points. Important and novel features of the approach are that it considers the case where different stations have different values of the minimum contention window, which is supported by IEEE 802.11e, and the approach can be implemented solely at the wireless stations, which passively monitor the activity of the available wireless channels, without requiring any changes to the access points. The proposed method is compared with the current approach for access point selection based on the SNR value (using the Received Signal Strength Indicator - RSSI) of beacon frames and with an approach according to which a station is associated with the access point which would result in the highest aggregate throughput for all access points. Our results show that even though the SNR and aggregate throughput maximizing scheme can achieve high aggregate throughput, they both achieve low throughput fairness, which deteriorates significantly as the geographical distribution of wireless stations becomes more unbalanced. On the other hand, the proposed access point selection scheme based on the average packet delay achieves significantly higher fairness, which is not influenced by the geographical distribution of wireless stations, and without a significant decrease of the aggregate throughput. Indeed, one can argue that fairness is more important, since customer studies have shown that users value more consistent quality of telecommunication services, rather than bursty periods of high and low quality.

A number of works have addressed the issue of access point selection. We present a brief overview of such work in Section V. Our work differs from previous work in the following points, which constitute novel features of the approach proposed in this paper: 1) we consider a throughput model that accounts for both the multi-rate operation of wireless networks and the contention in the form of packet collisions that can arise due to the 802.11 MAC, 2) our model is applicable to IEEE 802.11e networks containing different throughput classes that have different minimum contention windows, and 3) the proposed approach can be implemented solely at the wireless stations, without requiring any information from or modifications to the access points; the stations passively monitor the wireless channel corresponding to each available access point in order to estimate a packet delay metric, which

This work has been supported by British Telecom, UK and the General Secretariat for Research and Technology, Greece, through the program 05-AKMON-80.

reflects the load of each access point.

The rest of the paper is organized as follows. In Section II we present a brief overview of the DCF and EDCA mechanisms in IEEE 802.11. In Section III we first discuss an analytical throughput model for multi-rate 802.11e networks, and then present the proposed access point selection approach based on the average packet delay. In Section IV we present experimental results that compare the proposed approach with the SNR-based and aggregate throughput maximizing schemes. Finally, in Section V we present a brief overview of related work and in Section VI we conclude the paper identifying future research directions.

II. CONTENTION-BASED DCF AND EDCA

IEEE 802.11's DCF (Distributed Coordination Function) is based on CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance). According to the collision avoidance mechanism of CSMA/CA, a station performs a backoff procedure before initiating the transmission of a frame. After detecting that the medium is idle for a DIFS (DCF interframe spacing) interval, the station selects a random backoff period from $[0, CW - 1]$, where CW is referred to as contention window. The station waits for the channel to be idle for a total time equal to this backoff period, after which it can transmit a data frame with the basic CSMA/CA procedure, or an RTS frame with the RTS/CTS procedure. The contention window CW has an initial value CW_{min} , and is doubled when a collision occurs, up to the maximum value CW_{max} . When a frame is successfully transmitted, the contention window is set to its initial value CW_{min} .

The IEEE 802.11e standard supplement addresses the issue of QoS support in wireless LANs. The MAC protocol of 802.11e is the Hybrid Coordination Function (HCF), which supports both contention-based and controlled channel access. The contention-based access of HCF is supported by the Enhanced Distributed Channel Access (EDCA) mechanism, which is an extension of the DCF mechanism that enables distributed differentiated access to the wireless channel with the support of multiple access categories (ACs). A higher priority access category has a smaller minimum contention window CW_{min} , thus has a higher probability to access the channel. Additionally, different access categories can have a different maximum contention window CW_{max} and interframe spacing interval (IFS). Although the IEEE 802.11e standard defines a number of parameters that can be used to achieve service differentiation, it does not define how these parameters should depend on the network load and traffic characteristics in order to efficiently utilize the shared wireless channel.

III. ACCESS POINT SELECTION MODEL

In this section we first discuss a simple throughput model for multi-rate IEEE 802.11e networks using the contention-based EDCA mechanism. In our prior work, this model has been used to develop a framework for selecting the values of the minimum contention window to efficiently utilize the

wireless channel [1], and has been evaluated in an actual test-bed [2]. Based on this model, we next propose an access point selection scheme whose selection metric is the average delay a packet will experience for each access point a station can associate to.

A. Throughput model for EDCA

Several analytical studies have approximated IEEE 802.11's congestion avoidance procedure with a p-persistent model [3], [4]. In a p-persistent model, the probability p that a station tries to transmit in a time slot is independent of the success or failure of previous transmission attempts. The p-persistent model closely approximates the throughput of the actual congestion avoidance procedure when the average backoff is the same [3]; moreover, the saturation throughput has a small dependence on the exact backoff distribution [5].

If $E[CW]$ is the average contention window, then the approximate p-persistent model has transmission probability [3]

$$p = \frac{2}{E[CW] + 1}.$$

If the probability of a frame being involved in more than one collision is very small, then $E[CW] \approx CW_{min}$ [4]. In IEEE 802.11e, different wireless stations can have a different minimum contention window, hence using the same arguments as above [4], the corresponding transmission probability of station i in the p-persistent model is

$$p_i = \frac{2}{CW_{min,i} + 1}.$$

The MAC operation of IEEE 802.11 can be viewed in time as involving three different types of time intervals: a successful transmission interval, a collision interval, and an idle time interval. We denote the length of each interval type as T^{suc} , T^{col} , and T^{idl} , respectively. The duration of each time interval depends on the physical layer encoding and the MAC layer operations. For basic CSMA/CA in 802.11b, T^{suc} is given by¹:

$$T^{suc} = 2T_{PHY} + T_{SIFS} + \frac{8(O + L)}{R} + T_{ACK} + T_{DIFS}, \quad (1)$$

where L is the frame length, $O = 34$ or 28 bytes is the MAC overhead, R is the transmission rate, and T_{PHY} , T_{SIFS} , T_{DIFS} , T_{ACK} are the physical layer overhead, SIFS interval, DIFS interval, and ACK transmission time, respectively. The collision interval T^{col} is

$$T^{col} = T_{PHY} + \frac{8(O + L)}{R} + T_{DIFS}. \quad (2)$$

When the RTS/CTS procedure is used, the successful transmission and collision intervals can be computed in a similar manner, taking into account that in 802.11b the ACK, RTS, and CTS frames are always transmitted at the basic rate (1 or 2 Mbps), hence their transmission times are independent of the rate R .

¹We assume that the propagation delay is very small, hence do not consider it.

The average throughput for station i , considering a renewal assumption, can be expressed as the ratio of the average amount of data transmitted by that station in one time interval over the average time interval $x_i = \frac{E[X_i]}{E[T]}$ [6], [3], [4]. The average data transmitted by station i in one time interval, considering a p-persistent model and assuming that the station always has a frame ready to transmit, is

$$E[X_i] = p_i \prod_{j \neq i, j \in I_k} (1 - p_j) L,$$

where I_k is the set of stations associated to access point k , L is the frame size, which for simplicity we assume is the same for all stations. The average time interval is a weighted sum of the three types of intervals. If we assume that the intervals T^{suc} and T^{col} are normalized to the size of the idle slot time, and if all stations have the same transmission rate, then the average time interval is

$$E[T] = \sum_k p_k \prod_{j \neq k} (1 - p_j) T^{suc} + \left[1 - \prod_j (1 - p_j) - \sum_k p_k \prod_{j \neq k} (1 - p_j) \right] T^{col} + \prod_j (1 - p_j), \quad (3)$$

where $j \in I_k$. From the above, the average throughput x_i for station i is approximately

$$x_i = \frac{p_i \prod_{j \neq i, j \in I_k} (1 - p_j) L}{E[T]}, \quad (4)$$

where $E[T]$ is given by (3). Note that the above expression is valid under saturation conditions, when stations always have a packet to transmit, and can be applied to all versions of 802.11, provided all stations have the same transmission rate. The specific version of 802.11, and whether the CSMA/CA or RTS/CTS procedure is used, will determine the values of T^{suc} and T^{col} , which we have taken to be normalized to the duration of the idle interval.

Next we consider the case where different stations have different transmission rates. In 802.11b with the RTS/CTS procedure, the transmission rate does not affect the collision interval, since the latter involves RTS frames which are always transmitted at the basic rate (1 or 2 Mbps). Hence, for 802.11b with RTS/CTS, the average time interval is

$$E[T] = \sum_k p_k \prod_{j \neq k} (1 - p_j) T_k^{suc} + \left[1 - \prod_j (1 - p_j) - \sum_k p_k \prod_{j \neq k} (1 - p_j) \right] T^{col} + \prod_j (1 - p_j), \quad (5)$$

where the duration of the successful transmission interval T_k^{suc} depends the station's transmission rate through (1).

In 802.11 with the basic CSMA/CA procedure, the collision interval also depends on the transmission rate. In this case the second term in (5) needs to be modified appropriately, to take into account the transmission rate of the stations involved in

the collision. In particular, when two packets from stations with different transmission rates collide, then the channel cannot be used for an interval equal to the time required to transmit a packet from the slowest station. One can consider an approximation which assumes that collisions involving 3 or more stations are negligible. Based on this assumption, in equations (3) and (5) we replace T^{col} with a weighted average $\overline{T^{col}}$ containing two terms: the first corresponds to the case where one of the two stations involved in a collision is a low rate station, and the second corresponds to the case where both stations involved in a collision are high rate stations; the corresponding weights for each case are the probability that a collision involves at least one low transmission rate station, and the probability that it involves only high transmission rate stations.

B. Delay-based access point selection

Based on the above analysis, the average time, which includes the channel access delay and the transmission delay, for a new station i that associates with access point k to transmit a packet is approximately

$$\frac{E[T]}{p_i \prod_{j \neq i, j \in I_k} (1 - p_j)},$$

from which we consider the following factor

$$PD = \frac{E[T]}{\prod_{j \neq i, j \in I_k} (1 - p_j)}, \quad (6)$$

where $E[T]$ takes into account the transmission probabilities of the stations I_k already associated to access point k . Note that PD considers only parameters reflecting the access point's load, and not p_i , which corresponds to the congestion avoidance procedure of station i and appears in the average transmission delay of station i if it associates to any access point.

A smaller value of PD corresponds to a less loaded access point, i.e. PD reflects an access point's load. The proposed approach involves selecting the access point with the smallest value of PD . Note that this is a selfish policy, since the station tries to minimize his average packet delay, hence maximize his average throughput. Nevertheless, this selection policy also results in a balanced load among the different access points. Moreover, the metric (6) considers the increased delay due to channel contention, which is not the case with [7], [8], which define load as the inverse of the maximum achievable bit rate.

Note that even though the above packet delay metric PD is based on (4), which is valid under saturation conditions, PD still corresponds to a lower bound for the packet delay (hence, an upper bound for the achievable throughput), and as such can justify its use as an access point selection metric.

C. Online estimation of AP selection metric

In this section we discuss the online estimation of the average packet delay metric (6), which only requires that a wireless station passively monitors the activity of the wireless channel of each access point within its vicinity.

Equation (3) can be written as follows:

$$E[T] = P^{suc}T^{suc} + P^{col}T^{col} + P^{idl} \quad (7)$$

where $P^{suc}, P^{col}, P^{idl}$ are the probability of a successful interval, a collision interval, and an idle interval, respectively. If wireless stations have different transmission rates, then the last equation can be modified as follows:

$$E[T] = \sum_r P_r^{suc}T_r^{suc} + P^{col}T^{col} + P^{idl} \quad (8)$$

where P_r^{suc}, T_r^{suc} are the probability and the duration of a successful interval respectively, for a station with transmission rate r .

The above probabilities can be estimated by measuring the number of intervals of each type within some measurement period. Hence, if $N^{suc}, N^{col}, N^{idl}$ are the number of successful, collision, and idle intervals, respectively, then the above probabilities can be estimated from

$$P^{suc} \stackrel{est}{=} \frac{N^{suc}}{N^{suc} + N^{col} + N^{idl}}, \quad P^{col} \stackrel{est}{=} \frac{N^{col}}{N^{suc} + N^{col} + N^{idl}}, \quad P^{idl} \stackrel{est}{=} \frac{N^{idl}}{N^{suc} + N^{col} + N^{idl}}. \quad (9)$$

Additionally, we can consider the following approximation $P^{idl} \approx \prod_{j \neq i, j \in I_k} (1 - p_j)$, and hence the packet delay metric PD can be estimated from

$$PD \stackrel{est}{=} \frac{E[T]}{P^{idl}}, \quad (10)$$

where $E[T]$ is estimated from (7) or (8).

IV. PERFORMANCE EVALUATION

In this section we present experimental results that evaluate the access point selection approach based on the average packet delay (DBS - Delay-Based access point Selection) presented in the previous section. In particular, we compare the DBS scheme with the SNR-based selection scheme and a scheme where a new station selects the access point that yields the highest aggregate throughput (rather than the highest throughput for the new station, which corresponds to the selection criteria of the DBS scheme). The aggregate throughput maximizing scheme works as follows, in the case of two access points, AP 1 and AP 2: let T_1, T_2 be the current aggregate throughput for AP 1 and AP 2 respectively, and T'_1, T'_2 be the corresponding throughput if the new station connects to AP 1 or 2, respectively. The station chooses to associate with AP 1 if $T'_1 + T_2 > T_1 + T'_2$, or else it selects AP 2; this approach (Agg.Thr) is similar to the aggregate throughput maximizing selection discussed in [9], but considers the more accurate expression for each station's throughput given by (4).

The experiments target at identifying how the three access point selection methods compare in terms of both the aggregate throughput and fairness. Fairness is evaluated using R. Jain's well-known fairness index

$$Fairness\ Index = \frac{(\sum y_i)^2}{N \cdot \sum y_i^2},$$

where $y_i = x_i/w_i$, with x_i, w_i being station i 's throughput and weight (in the case of different throughput classes) respectively, and N is the number of stations. The fairness index varies from $1/N$ (where one station obtains all the capacity) to 1 (where all stations achieve equal throughput).

A. Experiment scenarios

We consider a simple topology containing two access points, Figure 1. Stations at a distance smaller than d_1 from each access point can transmit at 11 Mbps, whereas stations at a distance between d_1 and d_2 can transmit at rate 2 Mbps. The 802.11b protocol of course supports the intermediate rate 5.5 Mbps, but for simplicity we consider only the first two rates, 11 and 2 Mbps. Moreover, the simple topology in Figure 1 will allow us to gain insight on how the proposed access point selection method performs for different scenario parameters, such as the distance between the two access points, the number of stations and the unbalance of demand, both alone and in comparison to the other two methods (SNR and aggregate throughput maximizing selection - Agg.Thr).

In the first scenario we consider, Figure 1(a), the distance between the two access points is $l = 10$ m; this corresponds to the case where a second access point is introduced in a particular area in order to increase the wireless network's aggregate capacity, rather than its coverage. Note that in this scenario there is a large percentage of stations that are within the range of both access points, to which they can connect at rate 11 Mbps. In the second scenario we consider, Figure 1(b), the distance between the two access points is $l = 40$ m; this scenario corresponds to the case where a second access point is introduced in order to increase the coverage. In this scenario there are no stations within the range of both access points, to which they can connect at rate 11 Mbps.

In our experiments we consider an unbalance factor u that reflects the balance of the distribution of wireless stations around the two access points. In particular, the location of each station is selected as follows: First, with probability p_1 and p_2 the station will be located in the circle of radius d_2 around AP 1 and AP 2, respectively; these probabilities are related to the unbalance factor through $u = p_2/p_1$, and satisfy $p_1 + p_2 = 1$. Hence, a higher value of the unbalance factor u corresponds to a more skewed distribution of stations closer to AP 2. After the above selection, the station is randomly placed within the circle of radius d_2 around AP 1 or AP 2.

The values for the physical and MAC layer parameters appear in Table I, and correspond to the IEEE 802.11b standard with the basic CSMA/CA procedure. Note that in this case, the collision interval in (5) depends on the transmission rate, but for our comparative study of the aggregate throughput and fairness, setting T^{col} to the collision interval of the low rate station is sufficient. We consider the minimum collision interval to be $CW_{min} = 64, 128$, since these are the optimal values, i.e. the values that maximize the aggregate throughput, when there are around 5-10 stations at an access point [1]. We also present experiments with the 802.11e standard, and in particular when there are two classes with different minimum

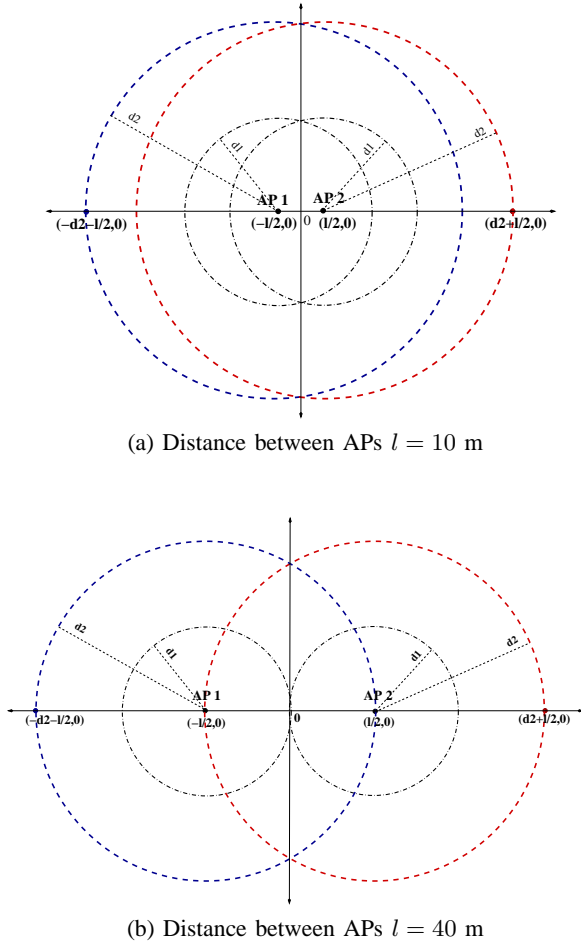


Fig. 1. Experiment scenarios. Distance between two APs is $l = 10$ or 40 m, range for achieving 11 Mbps transmission rate is $d_1 = 20$ m and for 2 Mbps is $d_2 = 40$ m.

contention window CW_{min} values. Finally, we consider UDP traffic with packet size 1044 bytes, which includes the UDP/IP headers.

B. Performance in terms of aggregate throughput and fairness

Next we compare the three access point selection methods in terms of the aggregate throughput and fairness. The aggregate throughput is the sum of the throughput of both access points, and is calculated by taking the sum of the throughput for each station, which is given by (4).

Each experiment is performed as follows: We start from a network with no stations. Then we assume that new stations gradually appear, up to the total number of stations corresponding to each experiment. With probability p_1 the station will be located in a circle around AP 1 with radius d_2 , and with probability p_2 it will be located in a circle around AP 2 with radius d_2 ; these probabilities satisfy $u = p_2/p_1$ and $p_1 + p_2 = 1$, where u is the unbalance factor for the particular experiment. The new station selects the access point to associate with, according to the three schemes we compare. The graphs we present show the average from 200 runs of the same experiment (with the same parameters). The

TABLE I
PARAMETERS FOR THE EXPERIMENTS

Parameter	Values
Slot Time	20 μ s
$T_{DIFS}, T_{SIFS}, T_{ACK}$	50, 10, 112 μ s
T_{PHY}	192 μ s
CW_{min}	64, 128
distance between APs, l	$l = 10, 40$ m
Range for 11 Mbps transmission	$d_1 = 20$ m
Range for 2 Mbps transmission	$d_2 = 40$ m
Unbalance factor u	$u = 1, 2, 3, 5$
Total # of stations N	$N \in [8, 60]$

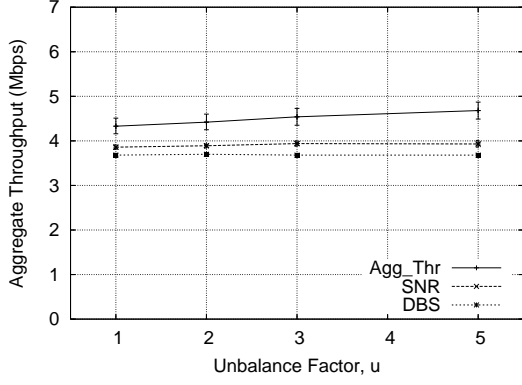
95% confidence interval is also shown in the graphs for the aggregate throughput.

Figure 2(a) shows the aggregate throughput achieved by the three access point selection schemes, for different values of the unbalance factor u , and for the topology shown in Figure 1(a). Observe that the aggregate throughput achieved by the SNR and DBS schemes is close, with the throughput achieved by the DBS scheme being smaller by less than 6%. Moreover, both achieve smaller throughput than the aggregate throughput maximizing scheme by approximately 11 – 16% (SNR) and 15 – 21% (DBS). Also observe that for all three schemes, the unbalance factor has a small influence on the aggregate throughput. The fact that the aggregate throughput maximizing scheme achieves the highest throughput is expected, since the selection policy seeks to maximize the aggregate throughput. On the other hand, the performance of the SNR selection scheme shows that the selection based on the SNR does not have a significant impact on the aggregate throughput, but has a significant impact on the fairness, as we will see next.

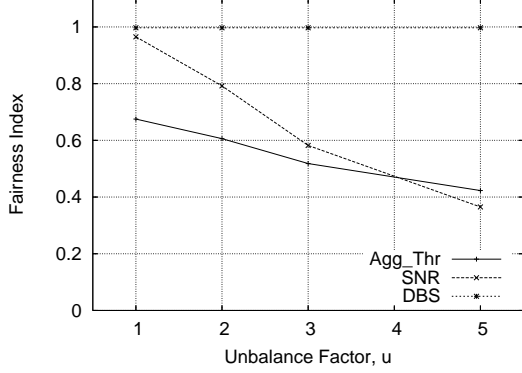
Comparison of the three schemes in terms of fairness appears in Figure 2(b), which shows that the proposed DBS scheme achieves significantly higher fairness than the other two schemes, without a significantly smaller aggregate throughput compared to the aggregate throughput maximizing scheme and an aggregate throughput which is very close to that achieved by the SNR scheme, as discussed above. Additionally, the figure shows that the fairness of the SNR and aggregate throughput maximizing scheme decreases significantly as the unbalance factor increases; indeed, the fairness of the SNR scheme decreases faster, compared to the aggregate throughput maximizing scheme.

To understand the degree of variance that corresponds to a particular value of the fairness index, we note that a fairness index of 0.61 corresponds to the case where the lowest rate can be 127 Kbps and the highest 2149 Kbps, which is an almost 17-fold difference! On the other hand, a fairness index of 0.97 corresponds to the case where the lowest rate can be 195 Kbps and the highest 407 Kbps.

Figures 3(a) and 3(b) show the aggregate throughput and fairness for a different total number of stations. As above, the aggregate throughput of the SNR and DBS schemes are close, and smaller than that of the aggregate throughput maximizing scheme; also observe that the aggregate throughput decreases

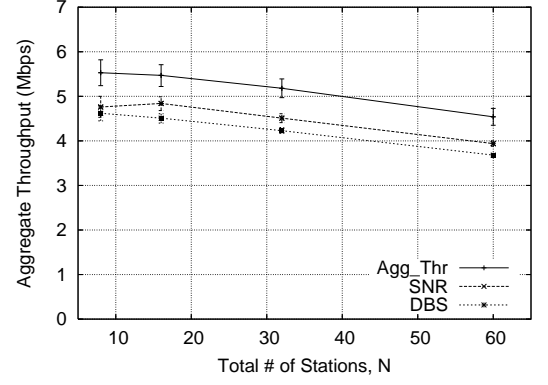


(a) Throughput

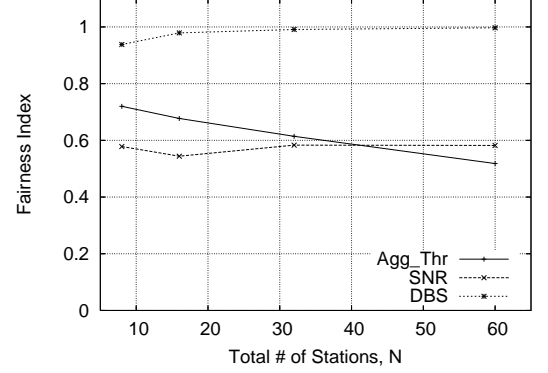


(b) Fairness Index

Fig. 2. Throughput and fairness for different unbalance factors u , distance between APs $l = 10$ m, number of stations $N = 60$, $CW_{\min} = 128$.



(a) Throughput



(b) Fairness Index

Fig. 3. Throughput and fairness for different total number of stations N , distance between APs $l = 10$ m, unbalance factor $u = 3$, $CW_{\min} = 128$.

with the number of stations, indicating that the network is becoming overloaded. Recall that the values of throughput that appear in Figure 3(a) are calculated analytically, by summing the throughput achieved by each station (4). However, comparison with simulation results, using the ns-2 simulator, show that the analytical values of throughput differ from the values obtained with simulation by less than 5% for $N \leq 32$ and less than 15% for $N = 60$.

Figure 3(b) shows that the DBS scheme achieves significantly higher fairness than the other two schemes. Moreover, observe that there is a slight increase of fairness for the DBS scheme, as the number of stations increases. On the other hand, the fairness for the aggregate throughput maximizing scheme decreases, while there is no clear trend for the fairness of the SNR scheme.

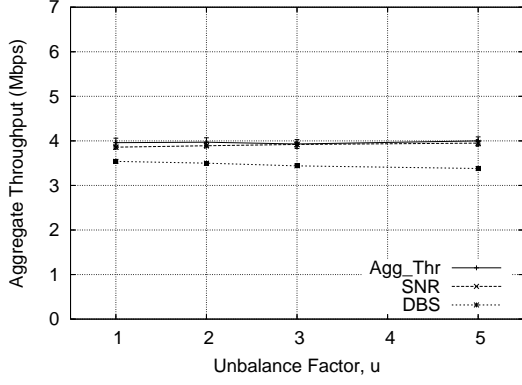
Figures 4(a) and 4(b) show the aggregate throughput and fairness for a different total number of stations, for the topology shown in Figure 1(b) where the distance between the two access points is $l = 40$ m. Observe that the aggregate throughput achieved by the SNR scheme is very close to that achieved by the aggregate throughput maximizing scheme, while the aggregate throughput achieved by the DBS scheme is approximately 12% lower. In terms of fairness, as in the previous experiments, Figure 4(b) shows that the DBS scheme achieves significantly higher fairness, which is independent of

the unbalance factor. On the other hand, the fairness of the other two schemes decreases significantly as the unbalance factor increases.

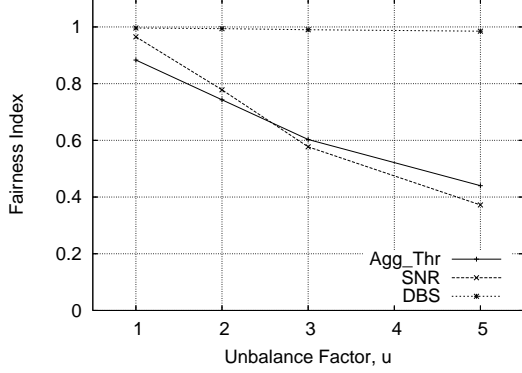
Figures 5(a) and 5(b) show the aggregate throughput and fairness when there are flows belonging to two different classes, which correspond to different minimum contention windows CW_{\min} . Note that in this case the fairness index is computed for the values $x_i \cdot CW_{\min,i}$, since a station's throughput, hence its weight, is approximately inversely proportional to the minimum contention window CW_{\min} . Observe that DBS maintains significantly high fairness, whereas the fairness of the aggregate throughput maximizing and SNR schemes are lower than the single class case, Figure 2(b); this is because the proposed DBS scheme explicitly takes into account different minimum contention window values. Additional results show that the DBS scheme still maintains a significantly high fairness in cases where the CW_{\min} values differ more, and when there is an uneven distribution of stations to different classes. On the other hand, in these cases the fairness of the other two schemes is smaller than that shown in Figure 5(b).

C. Evaluation of online implementation

Figure 6 compares the analytically computed values of the packet delay metric PD using (6), with those estimated online using (10). Observe that the online estimated simulation values



(a) Throughput



(b) Fairness Index

Fig. 4. Throughput and fairness for different unbalance factors u , distance between APs $l = 40$ m, number of stations $N = 60$, $CW_{\min} = 128$.

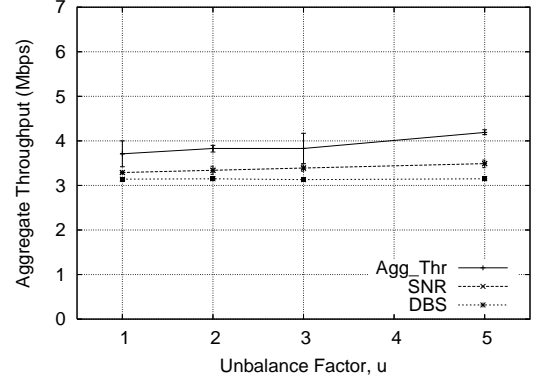
follow the analytical values quite well for a smaller number of stations. In any case, what is important for the access point selection problem is that the online estimated value of the packet delay metric increases with the number of stations.

One issue with the proposed online estimation of the packet delay metric using (6), which is also the reason that for a large number of stations the value estimated online does not closely follow the analytical value, is that for a small value of the minimum contention window the throughput and the average packet delay has a very small dependence on the number of stations. This is due to the exponential backoff procedure of 802.11, and is not an issue if (6) were applied directly, which however requires knowledge of the number of stations and the minimum contention window. One approach to address this issue is to consider an additional metric, such as the number of collisions, in cases where the packet delay metric for two access points is very close. Further work is investigating this direction.

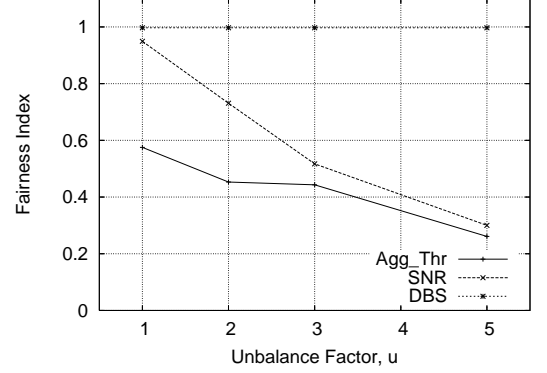
Finally, Figure 7 shows that the proposed online estimation approach can reach a stable estimate of the packet delay metric in a very short time, approximately 1 to 2 seconds.

V. RELATED WORK

In this section we present a brief overview of related work, identifying where it differs from the approach proposed in the



(a) Throughput



(b) Fairness Index

Fig. 5. Throughput and fairness for different unbalance factors u , distance between APs $l = 10$ m, number of stations $N = 60$, two classes with $CW_{\min} = 64$ (high) and $CW_{\min} = 128$ (low), 50% of the stations in each class.

current paper. The work in [7] investigates the problem of assigning stations to access points in order to achieve max-min fairness. The approach considers the fact that in wireless networks different stations can connect at different physical layer transmission rates, but does not account for the sharing model enforced by the IEEE 802.11 MAC protocol. The works in [9], [8] considers a simple model for the throughput that a station achieves, which takes into account the multi-rate operation of IEEE 802.11, but does not consider the decrease of throughput due to contention from packet collisions, nor the possibility offered in 802.11e to have stations belonging to different throughput classes. Moreover, [9] heuristically considers a linear combination of the transmission rate and throughput for access point selection. The work in [10] considers the packet error rate and the number of stations already associated to an access point as a metric for access point selection, which is not accurate when their transmission rates or minimum contention windows differ. The work of [11] proposes an access point selection scheme based on the quality of a link in both directions, in order to maximize the link rate; however, the scheme does not account for different physical layer transmission rates nor the 802.11 MAC sharing model. Another direction that exploits the cooperation and adaptation

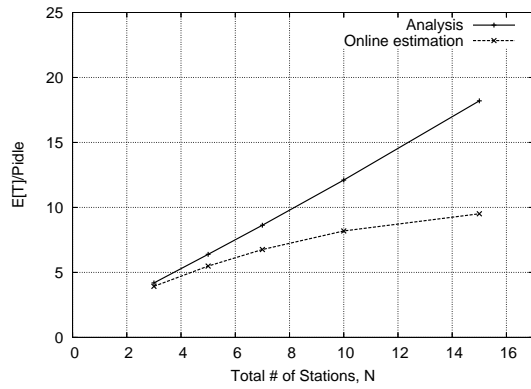


Fig. 6. Comparison of packet delay metric PD estimated using analysis (6) and online (10) through simulation. $CW_{\min} = 128$.

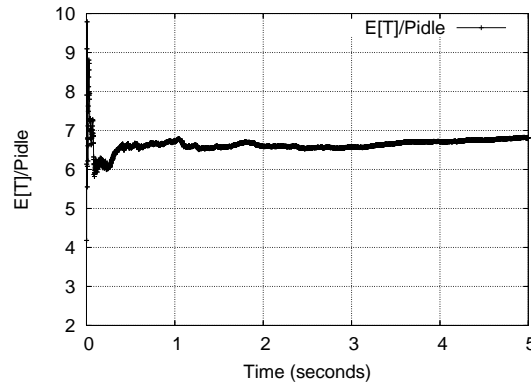


Fig. 7. Online estimation of $PD = E[T]/P^{idl}$ as function of time. $N = 7$, $CW_{\min} = 128$.

from both the user and the wireless network side to improve the performance of wireless hotspots is investigated in [12]. A common characteristic to most of the above work is that they require sharing of information between access points and wireless stations. The work in [13] considers an approach for selecting the access point based on the potential bandwidth that a station can achieve; the potential bandwidth is calculated from the channel access delay, which is estimated from the access delay that beacon frames encounter. A shortcoming of this approach is that it cannot be applied to IEEE 802.11e networks, where beacon frames have a separate queue, and hence the delay they encounter does not reflect the channel access delay that would be experienced by data packets.

VI. CONCLUSIONS

We have proposed an approach for access point selection that considers a packet delay metric for selecting the access point to associate to. Novel features of the approach are that it considers the contention-based nature of IEEE 802.11e's EDCA mechanism, it can be applied to the case of different classes with different minimum contention windows, and it can be implemented solely at the wireless stations, which passively monitor the wireless channel, without requiring any modifications to the access points. Experiments verify that

the approach can achieve significantly higher fairness than other schemes, without a significant decrease of the aggregate throughput. Another important point is that our scheme is a selfish scheme: each station selects the access point with which the packet delay is the smallest, hence his throughput largest. Hence, users have clear incentives to implement such an approach. This is not the case with other approaches where the access point selection is based on some global metric [8].

We are currently investigating the issue that arises when the minimum contention window is small and the number of stations large, in which case the average packet delay (and throughput) of two access points would not differ significantly. As noted in Section IV-C, one way to address this issue is to consider another metric, such as the number of collisions, in addition to the packet delay. Finally, in addition to simulation experiments with more complex topologies, we are investigating the implementation of the proposed online estimation approach in an actual test-bed. Recall that this requires measurements of the success and idle periods, which appears to be possible in current wireless interfaces; the more subtle variable is the number of collision intervals, which however can be estimated by the intervals that have not been identified as either success or idle intervals.

Finally, in other related work we are investigating the problem of channel selection and power control at the access point. Together with access point selection, which is the focus of this paper, all three control mechanisms affect the overall performance of wireless LANs, hence their joint investigation is important.

REFERENCES

- [1] V. A. Siris and C. Courcoubetis, "Resource Control for the EDCA Mechanism in Multi-Rate IEEE 802.11e Networks," in *Proc. of IEEE Int'l Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2006.
- [2] V. A. Siris and G. Stamatakis, "Optimal CW_{\min} Selection for Achieving Proportional Fairness in Multi-Rate 802.11e WLANs: Test-bed Implementation and Evaluation," in *Proc. of 1st ACM Int'l Workshop on Wireless Network Testbeds, Experimental evaluation and CHaracterization (WiNTECH)*, 2006.
- [3] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Trans. on Networking*, vol. 8, no. 6, pp. 785–799, 2000.
- [4] D. Qiao and K. G. Shin, "Achieving efficient channel utilization and weighted fairness for data communications in IEEE 802.11 WLAN under DCF," in *Proc. of IEEE/IFIP IWQoS*, May 2002.
- [5] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, "New Insights from a Fixed Point Analysis of Single Cell IEEE 802.11 WLANs," in *Proc. of IEEE INFOCOM*, 2005.
- [6] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Select. Areas Commun.*, vol. 18, no. 3, pp. 535–547, March 2000.
- [7] Y. Bejerano, S.-J. Han, and L. Li, "Fairness and load balancing in wireless lans using association control," in *Proc. of ACM International Conference on Mobile Computing and Networking (MOBICOM)*, 2004.
- [8] B. Kauffmann, F. Baccelli, A. Chaintreau, K. Papagiannaki, and C. Diot, "Self Organization of Interfering 802.11 Wireless Access Networks," INRIA, Tech. Rep. RR-5649, August 2005.
- [9] G. Kasbekar, J. Kuri, and P. Nuggehalli, "Online Association Policies in IEEE 802.11 WLANs," in *Proc. of 4th Int'l Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 2006.

- [10] Y. Fukuda, T. Abe, and Y. Oie, "Decentralized access point selection architecture for wireless LANs," in *Proc. of Wireless Telecommunications Symposium*, May 2004.
- [11] T. Korakis, O. Ercetin, S. V. Krishnamurthy, L. Tassiulas, and S. K. Tripathi, "Link Quality based Association Mechanism in IEEE 802.11h compliant Wireless LANs," in *Proc. of 1st Workshop on Resource Allocation in Wireless Networks (RAWNET)*, 2005.
- [12] A. B. P. Bahl and G. Voelker, "Hot-Spot Congestion Relief in Public-area Wireless Networks," in *Proc. of 4th IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*, 2002.
- [13] S. Vasudevan, K. Papagiannaki, C. Diot, J. Kurose, and D. Towsley, "Facilitating Access Point Selection in IEEE 802.11 Wireless Networks," in *Proc. of ACM Internet Measurement Conference*, October 2005.