On-Line AP Association Algorithms for 802.11n WLANs with Heterogeneous Clients

Dawei Gong, *Student Member*, *IEEE*, and Yuanyuan Yang, *Fellow*, *IEEE*

Abstract—As the latest amendment of IEEE 802.11 standard, 802.11n allows a maximum raw data rate as high as 600 Mbps, making it a desirable candidate for wireless local area network (WLAN) deployment. In typical deployment, the coverage areas of nearby access points (APs) usually overlap with each other to provide satisfactory coverage and seamless mobility support. Clients tend to associate (connect) to the AP with the strongest signal strength, which may lead to poor client throughput and overloaded APs. Although a number of AP association schemes have been proposed for IEEE 802.11 WLANs in the literature, the challenges brought by the new features in 802.11n have not been thoroughly studied nor the impact of legacy 802.11a/b/g clients in 802.11n WLANS on AP association. To fill in this gap, in this paper, we explore AP association for 802.11n with heterogeneous clients (802.11a/b/g/n). We first present a bi-dimensional Markov model to estimate the uplink and downlink throughput of clients and formulate AP association into an optimization problem, aiming at providing each client a bandwidth proportional to its usable data rate. Based on this Markov model, we propose an on-line AP association algorithm under the condition that each client can acquire timely information of all clients associated with nearby APs. Furthermore, for WLANs with densely deployed APs, we provide another on-line AP association algorithm with lower complexity, which takes full advantage of 802.11n transmissions by simply associating different types of clients with different APs. We have conducted extensive simulations and experiments to validate the proposed algorithms. The results show that our algorithms can significantly improve both 802.11n throughput and aggregated network throughput under various network scenarios, compared to previous AP association schemes. Our experiments also confirm the effectiveness of the algorithms in enhancing network throughput, maintaining proportional fairness among clients, and balancing load among APs.

Index Terms—Wireless local area networks (WLANs), IEEE 802.11n standard, AP association, frame aggregation, heterogeneous clients

1 Introduction

In the last decade, IEEE 802.11 based wireless local area networks (WLANs) have been widely deployed in universities, enterprises, public areas, and homes to establish high-speed inter-computer connections and provide access to the Internet. In the meanwhile, extensive research efforts [1]-[4], [7]-[10] have been devoted to WLANs to enhance network performance and explore novel applications. In a WLAN, clients need to associate with an access point (AP) to access the Internet. An AP and its associated clients are referred to as a basic service set (BSS). For convenience, we will use a *station* to represent an AP or a client in the rest of this paper. Typically, APs are deployed densely to provide any-where any-time WLAN connections. As a result, a client may be covered by multiple APs simultaneously and thus it has to determine which AP to associate with. In most current vendor implementations, a client selects the AP that has the highest signal to noise ratio (SNR), which is reflected by the received signal strength indicator (RSSI), to associate with. In such a scheme, the load among all APs is generally balanced if clients are evenly distributed. However, previous studies showed that client load is often distributed unevenly among APs [5], [6]. Clients tend to stay in particular areas of the network due to various reasons, such as proximity to power outlets, classrooms, gate area in airports, etc. As a result, APs in these areas get overloaded while APs in other areas are under-utilized. The performance of clients associated with an overloaded AP becomes unsatisfactory due to the intense contention and collisions among these clients.

One way to address this issue is to incorporate the AP load into the AP association algorithm. Various metrics, such as the number of associated clients, the measured channel busy ratio, and the delay of the Beacon messages have been used in the literature to describe the load of an AP. Most of these schemes aim at either maximizing the minimum throughput of all clients or balancing the load among all APs. In fact, it was proved in [7] that the objective of max-min fairness and load balance in AP association can be achieved simultaneously if one of them is achieved. In the meanwhile, WLAN stations can transmit packets at different data rates to ensure low packet error ratio under various channel conditions. It was pointed out in [8] that the throughput of multi-rate clients associated with an AP is in the same magnitude as that of the client with the lowest data rate, since each client has the same opportunity to access the wireless medium regardless of its data rate. Such phenomenon is called *performance anomaly*. Thus, the aggregated throughput of these max-min throughput schemes could be lower than that of the default RSSI-based scheme, even though the minimum throughput of all clients is maximized.

• *The authors are with the Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794. E-mail: {dawei.gong, yuanyuan.yang}@stonybrook.edu.*

Manuscript received 26 Aug. 2012; revised 30 Mar. 2013; accepted 04 July 2013. Date of publication 04 Aug. 2013; date of current version 14 Oct. 2014. Recommended for acceptance by Z. Tari.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TC.2013.156

To achieve a tradeoff between aggregated throughput and load balance, maximizing proportional fairness can be used as the objective of an AP association algorithm. Specifically, the utility function of proportional fairness is defined as the weighted sum of logarithmic client throughput. Some schemes in the literature have used this utility function for AP association [9], [10]. However, they assume that either the bandwidth or the time slot can be allocated to clients, which is difficult to implement in 802.11 based WLANs where all stations follow the distributed coordination function (DCF) to access the wireless medium.

Nevertheless, AP association in IEEE 802.11n [11] based WLANs has not been studied in depth. In 802.11n, several new features, such as multiple input multiple output (MIMO), channel bonding, frame aggregation, block acknowledgment, etc., have been introduced to boost physical data rates and improve the MAC efficiency. However, the benefits and challenges brought by these new features have not been fully explored in previous AP association schemes. In particular, clients in 802.11n WLANs will be a mixture of 802.11n clients and conventional 802.11a/b/g clients, which makes it difficult to fully utilize these new features in IEEE 802.11n to achieve high performance in practice. This is because that 802.11n standard is backward compatible and WLAN users tend to upgrade their devices gradually. The impact of conventional clients on AP association in IEEE 802.11n based WLANs has not been studied in the literature.

Therefore, in this paper we consider the problem of AP association in 802.11n WLANs with heterogeneous clients. First, we show via experiments that the performance of 802.11n transmissions could be severely affected by the association decisions of legacy 802.11a/b/g clients. We then introduce a network model for AP association, and develop a bi-dimensional Markov model to estimate client throughput in a BSS with heterogeneous clients. To describe the gain of the association, we define the MAC efficiency of a client as its achievable throughput divided by the optimal data rate it can use. We formulate the problem of AP association into an optimization problem, aiming at maximizing the MAC efficiency of all clients proportionally. To provide practical solutions, we propose an on-line AP association algorithm, named FAME, which maximizes the minimum MAC efficiency of the network when making association decisions. The algorithm is based on the bi-dimensional Markov model and requires timely knowledge of all clients in nearby BSSs. For WLANs where APs are densely deployed, we further propose another on-line algorithm with lower complexity, called Categorized AP association algorithm, which takes full advantage of 802.11n transmissions by associating different types of clients with different APs. We have conducted extensive simulations and experiments to evaluate the proposed algorithms under various network scenarios. The results demonstrate that both algorithms significantly outperform the compared schemes in terms of 802.11n throughput and network throughput.

The remainder of the paper is organized as follows. Section 2 reviews the related work. Section 3 gives an overview of new features and discusses the challenges of AP association in 802.11n WLANs. Section 4 introduces the Markov model to estimate client throughput and formulates the AP association problem into an optimization problem. Section 5 presents two on-line algorithms. Section 6 evaluates the performance of the proposed algorithms via simulations. Section 7 further implements and validates the proposed algorithms in a WLAN testbed. Finally, Section 8 concludes the paper.

2 Related Work

There have been a number of AP association schemes in the literature for load balance among APs. Various metrics were used in these studies to determine the AP load. In the AP association scheme proposed in [12], the sum of the reciprocal of data rates from associated clients is used to estimate the load of an AP. In [13] and [14], load balancing AP association was formulated into non-cooperative games, where the estimated packet access delay is used as the cost utility. It was proved in [13] that a Nash equilibrium can be achieved, and both centralized and localized algorithms were proposed in [14] to reach the equilibrium. In addition, the estimated file download time is adopted to indicate the AP load in the AP association scheme for web browsing in [15]. On the contrast, the effect of hidden terminals was considered in [18] as the main reason for AP performance degradation and used as the metric for AP association. Furthermore, the traffic intensity of clients was regarded as another impacting factor to AP load in the AP association scheme in [19]. However, it was pointed out [16] that greedy selection of the least-loaded AP does not guarantee optimal AP association, and the weighted sum of estimated throughput and usable data rate is used as the metric to make association decisions. In addition, it could be challenging to apply these schemes in realistic WLANs, since most of them require modifications to WLAN clients, which is not feasible for WLAN operators. Thus, in [20] a cell-breathing scheme was proposed for AP association, where APs balance their load by adjusting the transmitting power of Beacon frames, requiring no change at clients.

AP association has also been jointly considered with channel assignment and power assignment problems in WLANs. In [21], Gibbs sampler based algorithms were proposed for the joint AP association and channel assignment problem, with the objective of minimizing global interference and transmission delays. In [22], the joint problem of AP association and channel assignment was further studied from the perspective of a non-cooperative game, aiming at minimizing the aggregated packet transmission time for all clients. Moreover, an AP association algorithm for multi-channel WLANs was presented in [23], in which clients first select a best-signal AP on each channel to form a subset, then choose the AP that offers the highest throughput from this subset. Joint AP association and channel assignment in IEEE 802.11n WLANs was first studied in [24]. However, it mainly focuses on the impact of channel bonding while leaving the impact of conventional 802.11a/b/g clients untouched. AP association was also jointly explored with power assignment and bandwidth allocation in [25] and [26], respectively.

However, most of the above AP association schemes achieve throughput fairness among clients at the cost of reduced network throughput, due to aforementioned performance anomaly. Given that clients need to choose different physical data rates to adapt to various channel conditions, it is more desirable to achieve *airtime fairness* among clients, that is, all clients have generally the same medium access time regardless of their data rates. In the AP association schemes

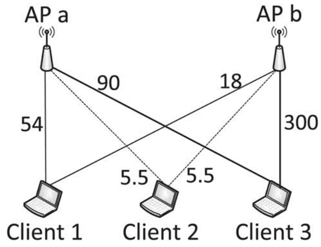


Fig. 1. An example WLAN, where lines denote the potential associations between APs and clients, and numbers by the links are corresponding data rates.

from [9] and [10], it was proved that airtime fairness can be achieved by implementing proportional fairness. Nevertheless, These schemes do not specify how to implement airtime or bandwidth allocation in WLANs to realize proportional fairness. To implement proportional fairness, AP association was jointly explored with rate adaptation and contention resolution in [30]. In particular, clients choose their associated APs, adjust their physical data rates and minimum contention window sizes to reach proportional fairness. To alleviate performance anomaly and achieve airtime fairness, several approaches other than AP association have also been proposed in prior work. In the WLAN virtualization scheme presented in [31], a controller was deployed to allocate uplink airtime to clients fairly, while clients regulate their uplink traffic based on commands from the controller. In [32], airtime usage control mechanisms based on enhanced distributed channel access (EDCA) were proposed, where stations control their airtime usage by choosing appropriate arbitration interframe space (AIFS) and the contention window size. Similarly, in the time fair CSMA (TFCSMA) scheme from [33], each station chooses a target throughput based on its optimal physical data rate, and adapts its minimum contention window size dynamically according to the ratio of measured throughput and the target throughput.

3 Background and Challenges in IEEE 802.11n

In this section, we first give a brief overview of new features introduced in IEEE 802.11n. We then discuss the challenges of AP association in 802.11n WLANs with heterogeneous clients.

3.1 New Features of IEEE 802.11n

As the latest IEEE 802.11 standard, 802.11n has introduced several new technologies to boost WLAN performance. First, 802.11n stations can be equipped with multiple antennas and MIMO is used to boost physical rates and reliability by transmitting multiple spatial streams simultaneously or exploring spatial diversity. In addition, the maximum coding rate is increased from 3/4 to 5/6 and a short guard interval (400 ns) between orthogonal frequency-division multiplexing (OFDM) symbols is introduced to improve spectrum efficiency and thus maximum physical rates. Furthermore, a channel bonding technology, also known as 40 MHz channel, is applied to further enhance physical rates through combining two non-overlapping 20 MHz channels together for data transmissions. At the MAC layer, frame aggregation mechanism is employed such that multiple frames are aggregated into a single frame before transmission. Each aggregated frame is acknowledged by a block ACK frame, in which a bitmap is used to acknowledge all sub-frames. In such a way, both the MAC overhead and random backoff period due to carrier sense multiple access with collision avoidance (CSMA/CA) are greatly reduced.

TABLE 1
Client and Aggregated Throughput under Different Association Strategies

AP for Client2	Client3 Frame Size	Throughput			
		Client1	Client2	Client3	Aggregate
AP a	1.5KB	3.9Mbps	3.0Mbps	21.0Mbps	27.9Mpbs
AP b	1.5KB	15.8Mpbs	3.0Mbps	5.6Mbps	24.4Mbps
AP a	30KB	3.9Mbps	3.0Mbps	180Mbps	186.9Mbps
AP b	30KB	15.8Mbps	2.8Mbps	53Mbps	71.6Mbps

3.2 Challenges of AP Association in 802.11n WLANs

In 802.11n WLANs, clients consist of 802.11n clients and conventional 802.11a/b/g clients, as the 802.11n standard is backward compatible. For convenience, we use *legacy clients* to refer to the conventional 802.11a/b/g clients in the rest of the paper. A high throughput mixed (HT-mixed) preamble or an explicit protection mechanism, such as RTS/CTS, or CTS-to-self, needs to be used by 802.11n clients at the presence of 802.11a/g and 802.11b clients, respectively. The performance of 802.11n clients with these preambles and protections was studied in [27]. It was shown that with the RTS/CTS protection, the MAC efficiency of 802.11n is only 12%, implying that most of time the wireless medium is wasted. Furthermore, as discussed in the last section, all clients in multi-rate WLANs generally have the same throughput, which is dominated by the client with the lowest data rate. Such performance anomaly also holds in 802.11n WLANs with legacy clients [28]. A bi-dimensional Markov model for multi-rate WLANs in [29] further validates this performance anomaly theoretically.

The performance of 802.11n WLANs with heterogeneous clients highly depends on the strategy of AP association. Take the 802.11n WLAN in Fig. 1 as an example, where clients 1, 2 and 3 are of 802.11 g, 802.11b and 802.11n types, respectively. We assume that client 1 is already associated with AP a and client 3 is associated with AP b. Based on the RSSI-based AP selection strategy, client 2 can be associated with either AP as both APs have the same data rate. On the other hand, based on the least-load AP association strategy, client 2 will associate with AP b, since the load of AP b is lower by using either the reciprocal of the data rate or average packet delay as the metric. However, much higher aggregated throughput can be achieved if client 2 is associated with AP a. Such inefficiency becomes more evident when frame aggregation is enabled. Table 1 lists the throughput of each client and the entire network for different AP associations of client 2. From the table, we can see that the AP association strategy in 802.11n WLANs has a significant impact on the throughput of 802.11n clients as well as the aggregated throughput.

4 Throughput Estimation and Problem Formulation

As shown in the previous section, network throughput in 802.11n WLANs is highly related to AP association decisions.

In this section, we first present a network model for a 802.11n WLAN with heterogeneous clients, and discuss the constraints on AP association. We then analyze the time to transmit a frame by various clients. After that, we introduce a bi-dimensional Markov model to estimate the uplink and downlink throughput of each client. Finally, we formulate the AP association problem into an optimization problem, aiming at providing each client the throughput that is proportional to its usable data rate. In this way, the AP loads are balanced while the network throughput is boosted.

4.1 Network Model

We consider a WLAN consisting of multiple APs and a number of clients. Let set A denote the set of APs and set N denote the set of clients. Each AP has a limited coverage area and all clients are randomly distributed in the field. We assume that there are sufficient channel resources and each AP is assigned with a channel that is orthogonal with the channel of other APs in the neighborhood. For client i, we use a variable $t_i = 0, 1, 2, 3$ to denote that it is an 802.11a, 802.11b, 802.11 g or 802.11n client, correspondingly. As 802.11a stations operate on 5 GHz band and 802.11b/g stations operate on 2.4 GHz band, we will not consider 802.11a clients in this paper, though the network model can be applied directly to 802.11a clients. We define a variable $u_i \in (0, 1)$ to denote the probability that client i has packets to transmit to the AP, which will be referred to as *uplink traffic* in the rest of paper. Similarly, we define a variable $d_i \in (0, 1)$ to denote the probability that the AP has packets to transmit to client i, which will be referred to as *downlink traffic*.

We use set R to denote the set of data rates supported by IEEE 802.11n standard. As aforementioned, 802.11n is backward compatible with 802.11a/b/g standards, thus the data rates of 802.11a/b/g clients are a subset of R. We assume that a client can estimate the optimal data rate for downlink transmission from an AP by measuring the RSSI of Beacon packets from the AP. We further assume that the channel condition is symmetric thus downlink and uplink transmissions between a client and an AP have the same optimal data rate. For client i and AP a, we use $r_{i,a} \in R$ to denote the uplink and downlink data rates between them. For simplicity, we assume that channel changes slowly, and thus the optimal data rate $r_{i,a}$ remains unchanged during AP association. We also assume the uplink traffic and downlink traffic for client i have the same average frame length, and use variable L_i to denote it. The notations used in this paper are summarized in Table 2.

TABLE 2
Notations

Symbol	Semantics
A	Set of all APs
N	Set of all clients
R	Set of data rates supported by 802.11n
m	Number of backoff stages for DCF
W_0	Minimum contention window size for DCF
t_i	Type of client i
o_a	Operation mode of AP a
a_i	Associated AP for client i
N_a	Set of clients associated with AP a
N_{a_i}	Set of clients sharing the same AP with client i
$r_{i,a}$	Optimal data rate between client i and AP a
L_i	Average frame length for client i
d_i	Downlink traffic probability of client i
u_i	Uplink traffic probability of client i
q_i	Probability of station i having packets to transmit
$T_s(i)$	Frame transmission time for client i
$T_c(i)$	Collision detection time for client i
$b_{l,k}(i)$	Stationary probability of client i at state (l, k)
τ_i	Transmission probability of client i
p_i	Transmission failure probability of client i
$p_s(i)$	Successful transmission probability of client i
δ	Duration of a time slot
S_i^{up}	Estimated uplink throughput of client i
S_i^{down}	Estimated downlink throughput of client i
α_i	MAC efficiency of client i

4.2 Association Constraints

We now discuss the constraints that a client has to satisfy to associate with an AP. For each client i, we use a variable $a_i \in A$ to denote its associated AP. In addition, we define set N_a as the set of clients associated with AP a.

$$N_a = \{i | i \in N, a_i = a\}.$$

If client i is not within the coverage area of AP a, the potential data rate $r_{i,a}$ would be zero. Thus the following condition must be met.

$$r_{i,a_i} > 0.$$

Then we simply use variable $r_i = r_{i,a_i}$ to represent the data rate between client i and its associated AP a_i. As aforementioned, transmissions to and from 802.11n clients need to use HT-mixed preambles if there are 802.11a/g clients in the network. Moreover, protection mechanism is required by 802.11 g/n transmissions if 802.11b clients are associated with the same AP. Hence we define an operation mode variable o_a for each AP a to indicate the necessity of HT-mixed preambles and protections, which is defined as

$$o_a = \min\{t_i | \forall i \in N_a\}.$$

4.3 Frame Transmission Time for Various Clients

In this subsection, we analyze the time for a client to successfully transmit a frame or detect a collision, which will be used to estimate the client throughput in later subsections.

In 802.11 WLANs, all stations follow the distributed coordination function (DCF) to access the wireless medium. With DCF, a station first listens to the medium for a DCF inter frame space (DIFS) period if the station has pending traffic. If the medium is idle during DIFS, the station selects a random backoff time from the contention window to postpone its transmission. In case the channel becomes busy again during the backoff period, the station stops counting down its backoff timer and waits for the channel to be idle. Otherwise, the station begins its transmission after the backoff period. If the data frame needs to be protected from legacy clients, the station first transmits a RTS or CTS-to-self frame. After that, the station transmits the data frame. A preamble is transmitted ahead of the data payload such that the receiver can acquire the coding and modulation schemes for the payload. After successfully receiving the data frame, the receiver sends back an ACK frame, beginning with a preamble as well, after waiting for a short inter frame space (SIFS) period. The timing of an 802.11 frame transmission is shown in Fig. 2.

Note that for a specific client i, the frame transmission time of uplink and downlink traffic is identical, since the uplink and downlink traffic has the same data rate based on the assumption that the channel between a client and an AP is

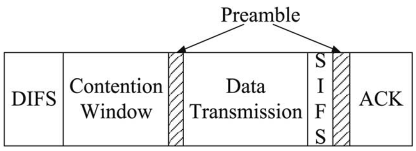


Fig. 2. Timing of an 802.11 frame transmission.

symmetrical. Thus, after obtaining the medium access opportunity, the frame transmission time to or from client i can be expressed as

$$T_s(i) = T_{DIFS} + T_{prot}(i) + T_{pre}(i) + T_{SIFS} + T_{data}(i) + T_{SIFS} + T_{ack}(i).$$

If a collision occurs, the required time for client i to detect the collision is

$$T_C(i) = T_{DIFS} + T_{prot}(i) + T_{pre}(i) + T_{SIFS} + T_{data}(i) + T_{ack\ timeout}(i).$$

In the above equations, T_{DIFS}, $T_{prot}(i)$, $T_{pre}(i)$, T_{SIFS}, $T_{data}(i)$, $T_{ack}(i)$ and $T_{ack\ timeout}(i)$ respectively stand for the DIFS duration, protection time, preamble time, SIFS period, data payload transmission time, ACK transmission time and ACK timeout time. DIFS duration and SIFS duration are constants. The payload transmission time for a client can be determined via dividing its average frame length by the data rate for the association. In addition, the protection time and preamble time for a client are specified in the 802.11n standard given the client type and the operation mode of its association AP.

4.4 Throughput Estimation

In this subsection, we first present a bi-dimensional Markov model to describe the DCF behavior of stations in 802.11n WLANs with heterogeneous clients. We then estimate the uplink and downlink throughput of each client, using the frame transmission time and the transmission and collision probabilities derived from the Markov model.

Fig. 3 shows the proposed bi-dimensional Markov model. In this model, we use $m+1$ backoff stages to describe the exponential backoff behavior of DCF for transmissions and retransmissions, where m is a constant value specified by the 802.11 standard. At each back off stage $l, 0 \le l \le m$, the station randomly chooses a back off value from the contention window size, W_l, which is given as follows

$$W_l = \begin{cases} W_0, & l = 0; \\ 2^l W_0, & 1 \le l \le m. \end{cases}$$

where W_0 is the minimum contention window size at stage 0. Thus, a station is at state (l,k) in the Markov model if it is in backoff stage l, $0 \le l \le m$, and its contention window size is $k, 0 \le k < W_l$. Moreover, a station is at state $(-1,0)$ if it has no packet to transmit when the channel is free.

We also use a variable q_i to denote the probability that station i has packets to transmit. Thus for any client $i \in N$, q_i equals its uplink traffic probability u_i. Meanwhile, for an AP $a \in A$, q_a is the summed downlink traffic probability of all of its associated clients. q_a is rounded to 1 if the summation is greater than 1.

Initially, station i is at the idle state $(-1,0)$. If the station has no packet to transmit in the next time slot, it transits back to

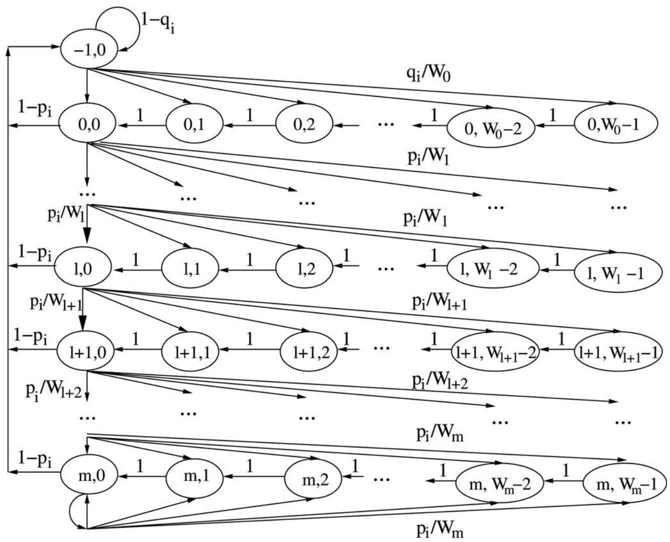


Fig. 3. Bi-dimensional Markov chain model for client i with saturated traffic, where W_l is contention window size at state l and p_i is the transmission failure probability of client i.

state $(-1,0)$. The transition probability is $(1-q_i)$, as q_i is the probability that station i has packets to transmit. Otherwise, station i transits to a random state $(0,k)$ at backoff stage 0. The transition probability is q_i/W_0 as the traffic probability of station i is q_i while the contention window size is randomly selected from $[0, W_0-1]$. If $k \ne 0$, station i transits from state $(0,k)$ to state $(0,k-1)$ with probability 1 in the next time slot. Otherwise, station i transmits a packet in the next time slot. If the transmission fails, the station transits to a random state $(1,k)$ in the next time slot with probability p_i/W_1, where p_i is the transmission failure probability of station i. If the transmission is successful, the station transits back to state $(-1,0)$. The transition probability is thus $(1-p_i)$. In general, a station transits from state (l,k) to state $(l,k-1)$ in the next time slot if $k > 0$. When $k = 0$, the station tries to transmit a frame. It transits back to state $(-1,0)$ if the transmission is successful; otherwise, it transits to a random state at backoff stage $l+1$.

Let $s(t)$ and $c(t)$ be the stochastic processes of backoff stage and backoff counter for station i, then its stationary probability at state (l,k) would be

$$b_{l,k}(i) = \lim_{t\to\infty} P\{s(t) = l, c(t) = k\}, 0 \le l \le m, 0 \le k < W_l.$$

Similarly, the stationary probability at the idle state can be expressed as

$$b_{-1,0}(i) = \lim_{t\to\infty} P\{s(t) = -1, c(t) = 0\}.$$

For any two states, (l,k) and (l',k'), we use condition probability $P\{l',k'|l,k\}$ to denote the one-step transition probability from state (l,k) to state (l',k'). Thus, the one-step transition probability for station i at state (l,k) is

$$\begin{cases} P\{-1,0|-1,0\} = 1-q_i & \\ P\{0,k|-1,0\} = q_i/W_0, & 0 \le k < W_0, \\ P\{-1,0|l,0\} = 1-p_i, & 0 \le l \le m, \\ P\{l+1,k|l,0\} = p_i/W_{l+1}, & 0 \le l < m, 0 \le k < W_{l+1}, \\ P\{l,k-1|l,k\} = 1, & 0 \le l \le m, 1 \le k < W_l, \\ P\{m,k|m,0\} = p_i/W_m, & 0 \le k < W_m. \end{cases} \quad (1)$$

By the chain regularities of the Markov model, the stationary probability of station i at state (l,k) is

$$b_{l,k}(i) = \frac{W_l - k}{W_l} \begin{cases} q_i \cdot b_{-1,0}(i), & l = 0, \\ p_i \cdot b_{l-1,0}(i), & 0 < l < m, \\ p_i \cdot [b_{m-1,0}(i) + b_{m,0}(i)], & l = m. \end{cases} \quad (2)$$

By (1) and (2), $b_{l,k}$ can be expressed as a function of $b_{-1,0}$, transmission failure probability p_i and traffic probability q_i. On the other hand, according to the normalization condition for a stationary Markov chain, the summation of stationary probabilities for all states should be equal to 1, which can be formally expressed as

$$b_{-1,0}(i) + \sum_{l=0}^{m} \sum_{k=0}^{W_l} b_{i,k}(i) = 1,$$

where $b_{-1,0}(i)$ can be derived from

$$b_{-1,0}(i) = \frac{2(1-p_i)(1-2p_i)}{2(1-2p_i)(1-p_i+q_i) + q_i W_0((1-p_i) + (2p_i)^m(1-2p_i))}.$$

Furthermore, we use variable τ_i to denote the transmission probability of station i. Then τ_i is equal to the summed probabilities that the counter equals zero at all backoff stages, which can be formally expressed by

$$\tau_i = \sum_{l=0}^{m} b_{l,0}(i) = \frac{2q_i(1-2p_i)}{2(1-2p_i)(1-p_i+q_i) + q_i W_0((1-p_i) + (2p_i)^m(1-2p_i))}.$$

For any client $i \in N$, its probability of transmitting a frame without colliding with its associated AP and other clients associated with the same AP is equal to the probability that client i transmits while all other clients and the AP are not transmitting. We define set N_{a_i} to be the set of clients associated with the same AP of client i. Then the successful transmission probability of client i is

$$p_s(i) = \tau_i \cdot (1 - \tau_{a_i}) \cdot \prod_{j \in N_{a_i}} (1 - \tau_j).$$

For any AP $a \in A$, its successful transmission probability is equal to the probability that AP a transmits while none of its associated clients are transmitting, which can be expressed as follows

$$p_s(a) = \tau_a \cdot \prod_{i \in N_a} (1 - \tau_i).$$

A transmission failure can be caused by either a collision or a channel error. Note that transmission failures caused by channel errors can be neglected in our model, as we have assumed that each station selects the optimal data rate based on the channel condition. In other words, we assume that all transmission failures are caused by collisions. Then for any client $i \in N$, its transmission failure probability p_i is equal to the probability that its associated AP or at least one of other clients associated with the same AP is transmitting, that is,

$$p_i = \tau_i \cdot [1 - (1 - \tau_{a_i}) \cdot \prod_{j \in N_{a_i}} (1 - \tau_j)].$$

For any AP $a \in A$, its transmission failure probability p_a is equal to the probability that at least one of its associated clients transmits while it is transmitting, which can be expressed as

$$p_a = \tau_a \cdot [1 - \prod_{i \in N_a} (1 - \tau_i)].$$

The theoretical throughput for client i can then be expressed as the length of successfully transmitted payload divided by the average duration of a time slot $T_{avg}(a_i)$ for all stations in the same BSS, that is,

$$S_i = \frac{p_s(i) \cdot L_i}{T_{avg}(a_i)}.$$

Similarly, the throughput for AP a can be expressed as

$$S_a = \frac{p_s(a) \cdot L_a}{T_{avg}(a)}.$$

Note that for AP a, it needs to transmit frames to all associated clients that have downlink traffic. We assume that the AP determines the receiving client of a transmission proportional to the downlink traffic probability of all clients. Then the expected frame length of AP a can be given by

$$L_a = \frac{\sum_{i \in N_a} d_i \cdot L_i}{\sum_{i \in N_a} d_i}.$$

The average time slot $T_{avg}(a)$ for AP a and all its associated clients can be further expressed as the summation of three expected slot durations

$$T_{avg}(a) = T_I(a) + T_S(a) + T_C(a),$$

where $T_I(a)$, $T_S(a)$ and $T_C(a)$ stand for the expected durations of an idle time slot, a successful frame transmission and a transmission failure due to collisions, respectively, for AP a and its associated clients.

The probability that AP a and its associated clients are not transmitting can be represented as

$$P_I(a) = (1 - \tau_a) \cdot \prod_{i \in N_a} (1 - \tau_i).$$

Thus the average duration of an idle time slot in the BSS where AP a resides is

$$T_I(a) = P_I(a) \cdot \delta,$$

where δ is the duration of a DCF backoff time slot.

The expected duration of a successful downlink transmission for AP a can be given by

$$T_s^{down}(a) = \frac{\sum_{i \in N_a} d_i \cdot T_s(i)}{\sum_{i \in N_a} d_i}.$$

Then the expected duration of a successful transmission for AP a and its associated clients is

$$T_S(a) = p_s(a) \cdot T_s^{down}(a) + \sum_{i \in N_a} p_s(i) \cdot T_s(i).$$

To determine the expected collision duration of an AP and its associated clients, we first sort them according to their collision duration T_C. For an AP a, its collision duration for downlink traffic depends on the data rate and frame length of the destination of a transmission. For simplicity, we will use the expected collision duration of all clients as the downlink collision duration, that is

$$T_C^{down}(a) = \frac{\sum_{i\in N_a} d_i \cdot T_C(i)}{\sum_{i\in N_a} d_i}.$$

Then assume that station i only collides with other stations in the same BSS that have a shorter collision duration. In other words, a collision between any two stations $i, j, (T_C(i) < T_C(j))$ will be counted by station j only, rather than both of them, when we calculate the expected collision duration of their BSSs. Then the collision probability of a station i can be regarded as the probability that station i transmits and at least another station with a shorter collision duration transmits simultaneously, while all stations with a longer collision duration are not transmitting. Then if the collision duration $T_C(i)$ of client i is less than the collision duration $T_C(a_i)$ of its associated AP, its collision probability can be rewritten as

$$p_i' = \tau_i \cdot \left[1 - (1-\tau_{a_i}) \cdot \prod_{j\in N_{a_i}}^{T_C(j)\le T_C(i)} (1-\tau_j)\right] \cdot \prod_{j\in N_{a_i}}^{T_C(j)>T_C(i)} (1-\tau_j).$$

If the collision duration $T_C(i)$ of client i is greater than the collision duration $T_C(a_i)$ of its associated AP, its collision probability is

$$p_i' = \tau_i \cdot (1-\tau_{a_i}) \cdot \left[1 - \prod_{j\in N_{a_i}}^{T_C(j)\le T_C(i)} (1-\tau_j)\right] \cdot \prod_{j\in N_{a_i}}^{T_C(j)>T_C(i)} (1-\tau_j).$$

Similarly, the collision probability of AP a can be rewritten as

$$p_a' = \tau_a \cdot \left[1 - \prod_{i\in N_a}^{T_C(i)\le T_C(a)} (1-\tau_i)\right] \cdot \prod_{i\in N_a}^{T_C(i)>T_C(a)} (1-\tau_i).$$

Then the expected collision duration for AP a and its associated clients is

$$T_C(a) = p_a' \cdot T_C(a) + \sum_{i\in N_a} p_i' \cdot T_C(i).$$

Finally, based on above equations, the estimated uplink throughput for client i can be represented as

$$S_i^{up} = \frac{p_s(i) \cdot L_i}{T_I(a_i) + T_S(a_i) + T_C(a_i)}. \tag{3}$$

For an AP a, the estimated downlink throughput for all of its associated clients can be given by

$$S_a^{down} = \frac{p_s(a) \cdot \sum_{i\in N_a}(d_i \cdot L_i)}{\sum_{i\in N_a} d_i \cdot [T_I(a) + T_S(a) + T_C(a)]}.$$

Accordingly, the estimated downlink throughput for client i can be expressed as

$$S_i^{down} = \frac{d_i}{S_{a_i}^{down} \cdot \sum_{j\in N_{a_i}} d_j}. \tag{4}$$

4.5 Formulation of AP Association Problem

As analyzed in the previous subsection, the downlink and uplink throughput of a client is related to not only its own traffic load and data rate, but also the client type, traffic load and data rates of all other clients in the same BSS. The composition of a BSS is eventually determined by the AP association strategy. Our goal of AP association is to provide each client the throughput proportional to its data rate. In this way, the throughput of a 802.11n client is determined by its signal quality to its associated AP, rather than the throughput of legacy clients in the same BSS. Then both 802.11n throughput and overall throughput can be boosted. In this subsection, we first define a MAC efficiency for each client as the metric to evaluate an AP association decision. After that, we formulate the AP association problem into an optimization problem.

The MAC efficiency of a client should reflect both the uplink throughput and downlink throughput of the client. In addition, the MAC efficiency should reflect the traffic load of a client as well, since it is inaccurate to say a client has poor MAC efficiency simply because it has no data to transmit and thus has throughput close to zero. Moreover, the data rate a client may use should be considered in the MAC efficiency as well, because it determines the highest achievable throughput of a client. Therefore, we define the MAC efficiency for client i as

$$\alpha_i = \frac{S_i^{up} + S_i^{down}}{\min\{1, u_i + d_i\} \cdot r_{i,a_i}}. \tag{5}$$

Given an AP association assignment, the clients in each BSS of the network are determined. Then the uplink and downlink throughput of each client can be estimated using the equations in the previous subsection. Accordingly, the MAC efficiency of each client can be determined. Then the AP association problem becomes the problem of finding an association assignment among all association possibilities, so that each client receives satisfactory MAC efficiency. As the utility function of summed logarithmic of the MAC efficiency of all clients can guarantee proportional fairness among all clients, we will use it as the objective function in the optimization.

We can now formulate the AP association problem into an optimization problem. Given a WLAN consisting of multiple 802.11n APs and a number of heterogeneous clients, find an AP association assignment, such that the MAC efficiency of all clients is proportionally maximized, while all association constraints are satisfied. The MAC efficiency of each client is determined using the estimated throughput from the previous subsection. The optimization problem can be described as follows

Maximize

$$\sum_{\forall i\in N} \log \alpha_i,$$

Subject to

$$a_i \in A, \forall i \in N, \tag{6}$$

$$\tau_i = \frac{2q_i}{2(1-p_i+q_i) + q_i W_0 (2p_i)^m + \frac{q_i W_0 (1-p_i)}{1-2p_i}}, \tag{7}$$

$$T_I(a) = \delta \cdot (1-\tau_a) \cdot \prod_{i\in N_a} (1-\tau_i), \tag{8}$$

$$T_S(a) = p_s(a) \cdot T_s^{down}(a) + \sum_{i\in N_a} p_s(i) \cdot T_s(i), \tag{9}$$

$$T_C(a) = p_a' \cdot T_C(a) + \sum_{i\in N_a} p_i' \cdot T_C(i), \tag{10}$$

$$S_i^{up} = \frac{p_s(i) \cdot L_i}{T_I(a_i) + T_S(a_i) + T_C(a_i)}, \tag{11}$$

$$S_a^{down} = \frac{p_s(a) \cdot \sum_{i \in N_a} (d_i \cdot L_i)}{\sum_{i \in N_a} d_i \cdot (T_I(a) + T_S(a) + T_C(a))}, \quad (12)$$

$$S_i^{down} = \frac{d_i}{S_{a_i}^{down} \cdot \sum_{j \in N_{a_i}} d_j}, \quad (13)$$

$$\alpha_i = \frac{S_i^{up} + S_i^{down}}{\min\{1, u_i + d_i\} \cdot r_{i,a_i}}. \quad (14)$$

In the above formulation, constraint (7) specifies the transmission probability for each station. Constraints (8), (9) and (10) determine the expected duration of an idle slot, successful transmission, failed transmission due to collision, respectively, for all clients associated with AP a. The estimated uplink throughput, downlink throughput and MAC efficiency for each client are given in constraints (11), (13) and (14).

Note that in the above formulation, there is an integrity constraint on x_i variables. Also, the problem is non-linear and non-concave. Thus, the complexity of this optimization problem grows exponentially as the network size increases. Since it is difficult to solve the optimization problem directly, in the next section, we will propose an on-line AP association algorithm based on the client throughput and MAC efficiency estimated in this section. In addition, for WLANs where APs are densely deployed, we will propose another algorithm with lower complexity than the first algorithm.

5 AP ASSOCIATION ALGORITHMS

In this section, we propose two on-line AP association algorithms for 802.11n WLANs with heterogeneous clients. The first algorithm is called *FAir Mac Efficiency (FAME)* algorithm, in which each client achieves the throughput proportional to its usable data rate, by associating with the AP that maximizes the minimum MAC efficiency of all associated clients. The second algorithm is referred to as *Categorized* algorithm, in which the potential of 802.11n transmissions is fully exploited by associating different types of clients with different APs. FAME algorithm has no preference on the network structure while Categorized algorithm performs the best in WLANs where APs are densely deployed. The algorithms are described in detail in following subsections.

5.1 FAME AP Association Algorithm

We first present FAME algorithm, with the primary objective to maximizing the minimum MAC efficiency of all clients. The MAC efficiency of a client can also be regarded as the fraction of medium time used by the client, since it has been defined as the ratio of achievable throughput to the usable data rate of the client. Thus, by maximizing the minimum MAC efficiency during AP association, all clients would have generally the same airtime. On the other hand, some low-rate legacy clients may do not have sufficient bandwidth to fulfill application demands, if their airtime is the same as other clients with much higher data rates. The AP association strategy to maximize the minimum throughput of all clients is more desirable in such scenarios. To take advantage of both strategies, a weight is assigned to the MAC efficiency in FAME and each client associates with the AP that maximizes the weighted MAC efficiency of all clients. For client i, its weighted MAC efficiency α_i^w can be given by

$$\begin{aligned}\alpha_i^w &= (1 + w(r_{i,a_i} - 1)) \cdot \alpha_i \\ &= (1 + w(r_{i,a_i} - 1)) \cdot \frac{S_i^{up} + S_i^{down}}{\min\{1, u_i + d_i\} \cdot r_{i,a_i}}, \end{aligned} \quad (15)$$

where w is a balancing factor ranging from 0 to 1. When w is 0, the weighted MAC efficiency is the MAC efficiency itself and then FAME algorithm provides airtime fairness. When w is 1 and a client has saturated traffic, its weighted MAC efficiency equals the client throughput and then FAME algorithm provides throughput fairness. w can also take an intermediate value between 0 and 1 to reach a balance between airtime fairness and throughput fairness. w is set to 0 be default as the primary goal of FAME is that each client achieves the throughput proportional to its physical data rate.

In FAME algorithm, each AP broadcasts a MAC efficiency entry for every associated client in the *Beacon* frame, including the client type, current data rate, average frame length, traffic probability, etc. In addition, a client estimates the potential data rate for each associable AP by measuring the RSSI of Beacon frames from all nearby APs. The client then determines the minimum weighted MAC efficiency of each nearby AP, by taking the MAC efficiency entries into the Markov model presented in the previous section. After that, the client associates with the AP that maximizes the minimum weighted MAC efficiency. The corresponding AP then updates its entries of MAC efficiency accordingly to include the new client. In the worst scenario, a client is in the coverage area of all APs in set A and an AP has at most $|N|$ associated clients. Then the time complexity of FAME algorithm is $O(|A| \cdot |N|)$, assuming that the time to estimate the minimum MAC efficiency of an AP is proportional to the number of associated clients.

FAME algorithm can be implemented in WLANs by extending the Beacon and *Associate Request* frames defined in the 802.11 standard. As discussed earlier, each AP maintains the MAC efficiency entries for all associated clients and broadcasts them in the Beacon frames. In particular, an AP acquires the uplink data rate, downlink data rate, and the average length of aggregated frames of a client from recently received and transmitted frames. An AP estimates the downlink traffic probability of a client from the number of pending frames to that client, while the client estimates the uplink traffic probability from the number of pending frames to the AP. After receiving the Beacon frames of all nearby APs, A client associates with the AP that maximizes the minimum weighted MAC efficiency by sending an extended Associate Request frame, including its own uplink traffic probability. To cope with client mobility, every client scans other channels to receive Beacon frames from nearby APs every a few seconds, so as to acquire updated MAC efficiency entries. A client then runs the FAME algorithm to determine whether the minimum MAC efficiency can be improved by associating with a different AP. If so, the client measures the received signal strength of a few more Beacon frames from the new AP. The client associates with the new AP only if the signal strength becomes stronger or at least remains the same, so as to avoid unnecessary associations when moving away from the new AP. The pseudo code of FAME algorithm is given in Table 3.

TABLE 3
FAME AP Association Algorithm

```
Input:
  Set of APs A
  Set of clients N
  Client Type Vector T = {t_i | ∀i ∈ N}
  Weight factor w
Output:
  AP Association Matrix X = {x_{i,a} | ∀i ∈ N, a ∈ A}
Algorithm:
  for each client i ∈ N
    for each AP a ∈ A
      if client i receives Beacon frame from a
        Add AP a into subset A_i;
        Determines r_{i,a} between client i and AP a;
      end if
    end for
    Determine probability of uplink traffic u_i;
    for each AP a' ∈ A_i
      Send a Probe Request to AP a';
      Receive a Probe Response from AP a';
      Let N_{a'} be the set of clients associated with a';
      Determine MAC efficiency of all clients in N_{a'} ⋃ i;
      α'_a = min{α_j | j ∈ N_{a'} ⋃ i};
      Weighted MAC efficiency α^{w'}_a = (1 + w(r_{i,a'} − 1)) · α'_a;
    end for
    c = arg max_{a'} {α^{w'}_a | a' ∈ A_i};
    Associate client i with AP c;
    x_{i,c} = 1;
    Update MAC efficiency entries for AP c;
  end for
```

TABLE 4
Categorized AP Association Algorithm

```
Input:
  Set of APs A
  Set of clients N
  Client Type Vector T = {t_i | ∀i ∈ N}
Output:
  AP Association Matrix X = {x_{i,a} | ∀i ∈ N, a ∈ A}
Algorithm:
  for each AP a ∈ A
    Set its category c_a to zero;
  end for
  for each client i ∈ N
    for each AP a ∈ A
      if client i receives Beacon frames from AP a
        Get the category c_a of AP a;
        Get the lowest data rate r^{min}_a among clients of AP a;
        Add AP a into subset A_i;
      end if
    end for
    if ∃A'_i ⊂ A_i such that A'_i = {a ∈ A_i | c_a = t_i}
      b = arg max_a {r_{i,a} | a ∈ A'_i};
      Associate client i with AP b;
    else if ∃A''_i ⊂ A_i such that A''_i = {a ∈ A_i | c_a = 0}
      b = arg max_a {r_{i,a} | a ∈ A''_i};
      Associate client i with AP b;
      Set the category c_b of AP b to t_i;
    else
      b = arg min_a {r_{i,a} − r^{min}_a | a ∈ A_i};
      Associate client i with AP b;
    end if
    x_{i,b} = 1;
  end for
```

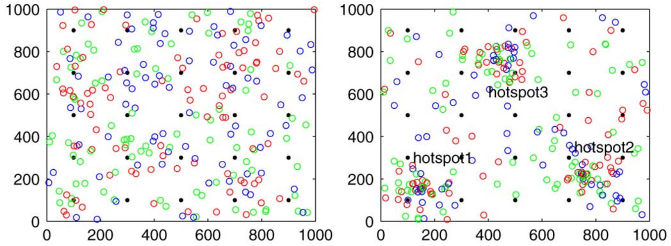
5.2 Categorized AP Association Algorithm

As discussed in the previous subsection, FAME algorithm requires the data rate and traffic load information of all clients in nearby BSSs to make association decisions. Thus each client should re-execute FAME algorithm whenever the channel condition or traffic load of other clients changes, to ensure it is associated with the best AP. This may affect ongoing applications and place extra computing overhead on clients. In this subsection, we present Categorized algorithm, which is much less sensitive to network variation.

Categorized algorithm minimizes the impact of legacy clients on 802.11n transmissions by associating different types of clients with different APs. APs are categorized by the type of their preferred clients. The rationale behind this is that in dense WLAN deployments, every client is within the coverage of multiple APs most of time. Thus a client is able to use satisfactory data rates by associating with a categorized AP, which does not necessarily have the best signal quality. In this way, performance degradation of 802.11n transmissions resulted from coexistence with legacy clients in the same BSS can be greatly alleviated.

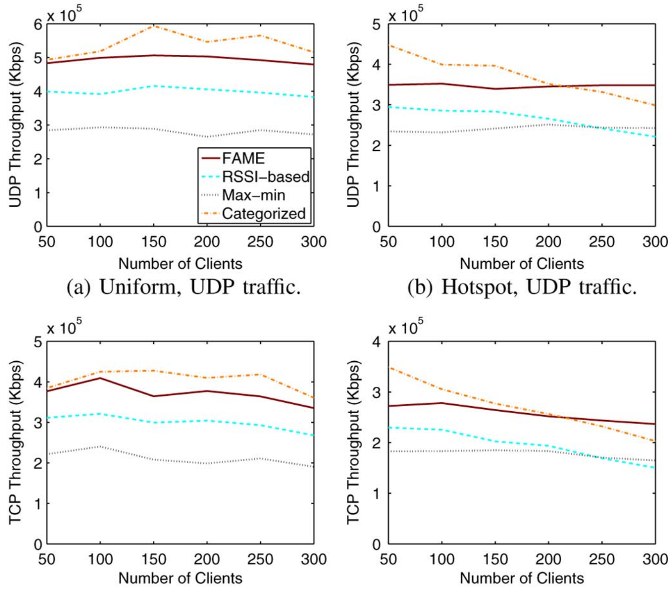
Initially, all APs are not assigned to any category. When a new client i joins the network, it first checks whether it is within the coverage area of one or more APs that prefer the type of client i. If more than one APs qualify, client i associates with the AP that has the best signal quality. If there is no such an AP, client i examines whether it is covered by APs that have not been categorized yet. If so, client i associates with the AP resulting in the maximum data rate for the association and the AP updates its category to the type of client i. Clearly, it cannot be guaranteed that each client is able to associate with an AP with matched preference, especially when APs are deployed sparsely. In such a case, client i associates with the AP whose minimum data rate is closest to the data rate of client i if associated. The reason is that if the preamble and protection overhead caused by legacy clients is inevitable, we want to minimize the data rate difference among all clients associated with the same AP. In the worst case, a client is covered by all APs in set A. For each AP, the client takes constant time to determine whether the AP is the best choice, given the category and the minimum data rate among all associated clients of the AP. Therefore, the complexity of Categorized algorithm is $O(|A|)$.

Similar to FAME algorithm, Categorized algorithm can be implemented on deployed WLANs by extending the *Beacon* frame of 802.11. In the Beacon frame, an additional category field and a new data rate field are appended. The category field indicates the preferred client type of an AP, while the new data rate field specifies the lowest data rate used by all associated clients of the AP, for both uplink and downlink traffics. After receiving Beacon frames from all associable APs, a client can make association decisions based on simple comparisons. Each client also scans other channels every a few seconds to receive Beacon messages from nearby APs, so as to cope with client mobility. The client then executes the Categorized algorithm to determine whether there exists a better AP to associate with. Similar to FAME algorithm, the client measures the signal strength of a few more Beacon frames from the new AP, and associates with the new AP only if the signal strength becomes stronger or remains the same. An AP resets its preference to uncategorized if all of its associated clients disassociate from it. The pseudo code of the algorithm is given in Table 4.



(a) Uniform distribution. (b) Hotspot distribution.

Fig. 4. An example of client distribution in the simulation. The black dots denote APs, and the colored dots denote clients. In particular, red, green and blue dots represent 802.11b, 802.11 g and 802.11n clients, respectively: (a) uniform distribution and (b) hotspot distribution.



(a) Uniform, UDP traffic. (b) Hotspot, UDP traffic. (c) Uniform, TCP traffic. (d) Hotspot, TCP traffic.

Fig. 5. Network throughput under uniform and hotspot client distributions with respect to various client densities: (a) uniform, UDP traffic; (b) hotspot, UDP traffic; (c) uniform, TCP traffic; and (d) hotspot, TCP traffic.

6 Simulation Results

In this section, we evaluate the performance of the proposed AP association algorithms via simulations, and compare them with the RSSI-based and the max-min throughput association schemes. We first evaluate the performance of the proposed algorithms in terms of network throughput, MAC access delay and fairness among clients in WLANs where all clients are stationary. After that, we examine their performance in WLANs where there exist mobile clients.

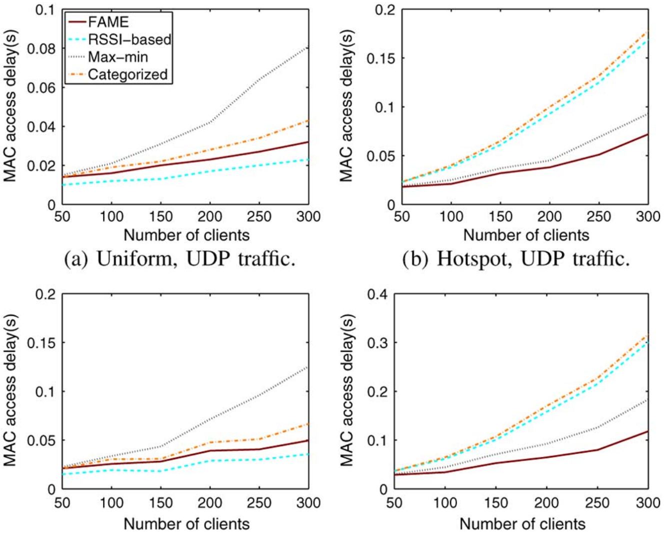
In the simulation, 25 APs are deployed in a $1000 \times 1000m^2$ field, and the APs are evenly placed on a 5×5 grid to provision fully coverage to the field. As co-channel interference among neighboring BSSs is not considered in this paper, we assign an orthogonal 40 MHz channel to each AP to eliminate potential interference. An AP uses one part of its assigned 40 MHz channel for transmissions to/from legacy clients. A number of clients are deployed in the field and the type of each client is randomly selected among 802.11b, 802.11g and 802.11n. The collision-aware rate adaptation (CARA) algorithm [34] is employed by for rate adaptation, as it differentiates transmission failures resulted from channel variations from transmission failures caused by collisions, and adapts data rate only for channel variations. For 802.11b and 802.11g transmissions, all data rates specified in the IEEE 802.11b and 802.11g standards are used in CARA. For 802.11n transmissions, only single spatial-stream, short guard interval (SGI) data rates at 40 MHz are used, since the CARA algorithm does not support adaptation between single spatial-stream and double spatial-stream data rates. However, it should be pointed out that other rate adaptation algorithms for WLANs could also be used in the simulation, as the proposed AP association algorithms do not rely on a specific rate adaptation algorithm. TCP and UDP traffics have been adopted separately in the simulation to emulate various applications. Every client always has saturated traffic in both uplink and downlink directions. For each network configuration, the simulation is run 300 seconds.

6.1 Performance in WLANs with Stationary Clients

In this subsection, we evaluate the network performance of the proposed algorithms in WLANs where all clients are static. As shown in Fig. 4, we consider two types of client distributions. (1) Uniform: all clients are randomly distributed over the entire field; (2) Hotspot: 75% clients are placed in three circle-shaped hotspot areas, while other clients are randomly distributed in the rest of the field.

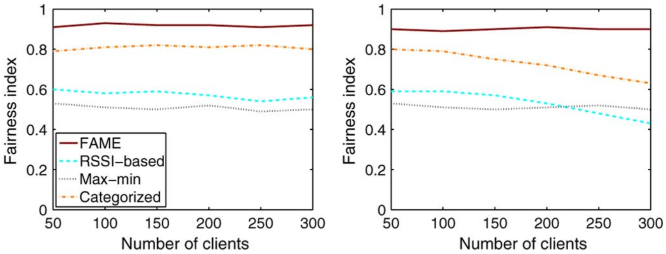
We first study the network throughput of FAME and Categorized algorithms under various client densities. The UDP throughput under uniform and hotspot client distributions is plotted in Figs. 5a and 5b, respectively, where the number of clients increases from 50 to 300 in a step of 50. We can see that both FAME and Categorized algorithms can achieve much higher UDP throughput than the compared schemes under both types of client distribution, regardless of the client density. When 200 clients are uniformly distributed in the field, the UDP throughput of FAME and Categorized algorithm is 25% and 36% higher than that of RSSI-based algorithm, and 92% and 107% higher than that of max-min algorithm, respectively. This can be attribute to the fact that the impact of legacy clients, especially the protection overhead from RTS/CTS transmissions, is alleviated in the proposed algorithms. We can also observe that the Categorized algorithm leads to the highest UDP throughput under uniform client distribution, as most 802.11n transmissions are isolated from 802.11b/g transmissions while no AP is overloaded. However, under hotspot distribution, the UDP throughput of Categorized algorithm drops below that of FAME algorithm when the number of clients grows beyond 200. The UDP throughput of RSSI-based algorithm also decreases as the client density increases. This is because APs in the hotspots may become overloaded when the client density is high, as the AP load is not considered in Categorized and RSSI-based algorithms when making association decisions. The TCP throughput under uniform and hotspot client distributions is plotted in Figs. 5c and 5d. The TCP throughput in both cases is lower than the UDP throughput, as the transmission of TCP ACK consumes medium access time and intensifies collisions. Nevertheless, both proposed algorithms achieve much higher TCP throughput than the compared schemes for similar reasons.

We then examine the average *MAC access delay* of the proposed algorithms, so as to assess the intensity of medium access contention, the transmission overhead, and the applied



(a) Uniform, UDP traffic. (b) Hotspot, UDP traffic. (c) Uniform, TCP traffic. (d) Hotspot, TCP traffic.

Fig. 6. Average MAC access delay under uniform and hotspot client distributions with respect to various client densities: (a) uniform, UDP traffic; (b) hotspot, UDP traffic; (c) uniform, TCP traffic; and (d) hotspot, TCP traffic.



(a) Uniform distribution. (b) Hotspot distribution.

Fig. 7. Fairness index of MAC efficiency under various client densities: (a) uniform distribution and (b) hotspot distribution.

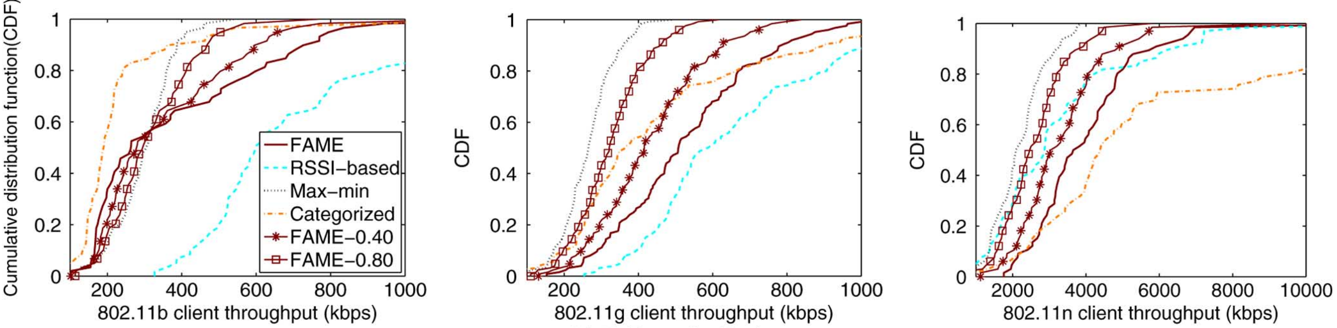
data rates of different AP association algorithms. Here MAC access delay is defined as the duration from the time that a packet enters the queue at the MAC layer to the time that the packet is transmitted. The average MAC access delay for UDP traffic in uniform client distribution and hotspot client distribution is given in Figs. 6a and 6b. It can be noted that the MAC access delay of all AP association algorithms increases along with the number of clients, because the contention intensity within each BSS is high when there are many clients in it. In the uniform distribution case, RSSI-based algorithm has the shortest MAC access delay, since every client is associated with the AP that has the best signal quality and thus uses a high data rate, while the traffic load is in general evenly distributed among all APs. The MAC access delay of FAME and Categorized algorithms is very close to that of RSSI-based algorithm and lower that of max-min algorithm, indicating the benefits of avoiding protection overhead when applicable. On the contrast, under hotspot distribution, FAME has the shortest MAC access delay while Categorized has the longest MAC access delay. Moreover, max-min has similar MAC access delay to FAME, while RSSI-based algorithm has similar MAC access delay to Categorized. The reason behind this is that only FAME and max-min algorithms are capable of associating clients in hotspots to APs out of the hotspots. Similar results can be observed for TCP traffic in Figs. 6c and 6d, except that the MAC access delay of all algorithms is longer because each station need to transmit both TCP data and TCP ACK packets.

Next, we evaluate the fairness of the proposed AP association algorithms. As aforementioned, the primary objective of our proposed algorithms is that each client achieves the throughput that is proportional to its physical data rate. In other words, we want to achieve fairness among the MAC efficiency of all clients. We adopt the fairness index from [35] to quantify the fairness of MAC efficiency. The fairness index $F(\alpha)$ of MAC efficiency is defined as

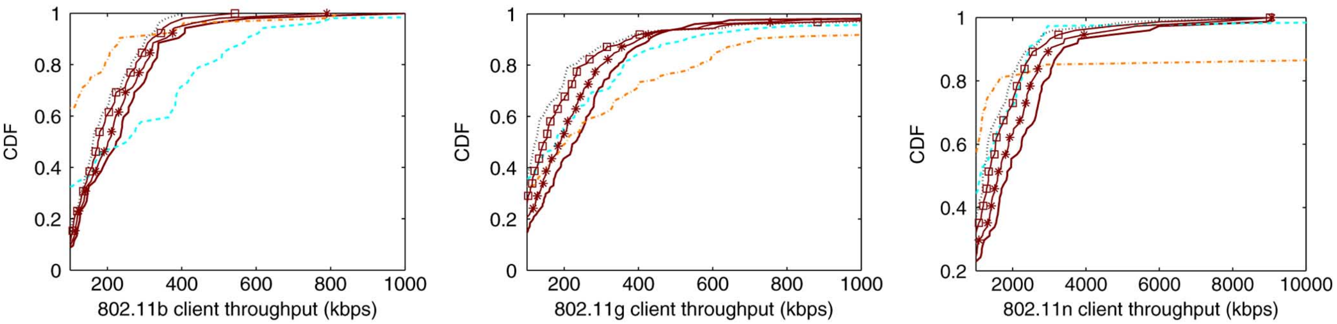
$$F(\alpha) = \frac{\left(\sum_{i\in N}\alpha_i\right)^2}{|N|\sum_{i\in N}\alpha_i^2}, \qquad (16)$$

where N is the set of clients and α_i is the MAC efficiency of client i (Equation (5)). $F(\alpha)$ takes a value from 0 to 1, and reaches 1 is all clients have the same MAC efficiency. The fairness index of MAC efficiency for UDP traffic is plotted in Fig. 7, where the number of clients varies from 50 to 300. We can observe that the fairness index of FAME algorithm is always above 0.9 under both uniform and hotspot distributions regardless of client density, indicating that 90% of clients achieve a fair MAC efficiency. The fairness of Categorized algorithm is also above 0.8 under uniform distribution, validating the effectiveness of isolating 802.11b, 802.11 g and 802.11n transmissions from each other. The fairness index of RSSI-based and max-min algorithm is below 0.6. This is mainly because that 802.11g transmissions are severely suppressed by low-rate 802.11b transmissions and aggregated-frame 802.11n transmissions. We can also see that under hotspot distribution, the fairness index of Categorized algorithm degrades as the number of client increases. The reason is that the MAC efficiency of clients in the hotspots becomes very low when the client density is high, as the associated APs are overloaded. While the MAC efficiency of clients in the rest area of the field is much higher, since the associated APs are underutilized. The fairness index of RSSI-based algorithm is worse when the client density is high for the same reason. We have also evaluated the fairness index of the proposed algorithms for TCP traffic. In Figs. 7-9, the results for TCP traffic is very similar to the results of UDP traffic and thus are not presented due to space limitation.

We also assess the fairness of the proposed algorithms and verify the effectiveness of the balancing factor in FAME algorithm by examining detailed client throughput in an example WLAN with 200 clients. The cumulative distribution function (CDF) of client throughput for UDP traffic is given in Fig. 8, where FAME-0.4 and FAME-0.8 denote variants of FAME algorithm where the balancing factor w is set to 0.4 and 0.8, respectively. Different types of clients are plotted separately for clear presentation. Under uniform distribution, it can be observed that although about 60% 802.11b clients have higher throughput if the max-min algorithm is used, over 90% of 802.11 g and 802.11n clients have much higher throughput if FAME or Categorized algorithm is used. In addition, most 802.11b clients in FAME and Categorized algorithms can achieve at least one half of the throughput achieved by 802.11b clients in max-min algorithm, which is acceptable for most applications. On the other hand, even if the max-min algorithm is used, the average throughput of 802.11n clients (around 2.5 Mbps) is still much higher than the average throughput of legacy clients (around 250 Kbps). In other words, the max-min algorithm cannot provide throughput fairness



(a) Uniform distribution.

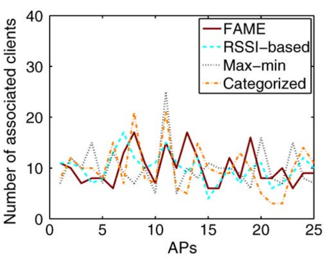


(b) Hotspot distribution.

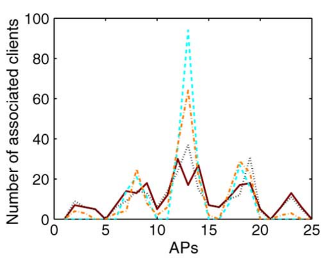
Fig. 8. Cumulative distribution function (CDF) for throughput of 802.11b, 802.11 g and 802.11n clients in an example WLAN with 200 clients: (a) uniform distribution and (b) hotspot distribution.

among 802.11n clients and legacy clients, because of the frame aggregation mechanism of 802.11n. Under hotspot distribution, it can be noted that regardless of the client type, a large number of clients have very low throughput in Categorized and RSSI-based algorithms due to AP overload. FAME algorithm can greatly reduce the amount of such clients while ensuring that the throughput of each client is comparable to other AP association schemes. More importantly, it is obvious that as the balancing factor increases, the throughput of 802.11b clients increases while the throughput of 802.11n clients decreases. In particular, the CDF of FAME algorithm is very close to the CDF of max-min algorithm when the balancing factor equals 0.8. This validates that FAME algorithm can reach a balance between airtime fairness and throughput fairness by choosing an appropriate balancing factor.

Finally, we evaluate the AP load distribution of various algorithms, by examining the number of associated clients on each AP in an example WLAN with 200 clients. The simulation results are shown in Fig. 9. Note that under uniform distribution (Fig. 9a), the number of associated clients on each AP is very close to the average value of all algorithms, although



(a) Uniform distribution.



(b) Hotspot distribution.

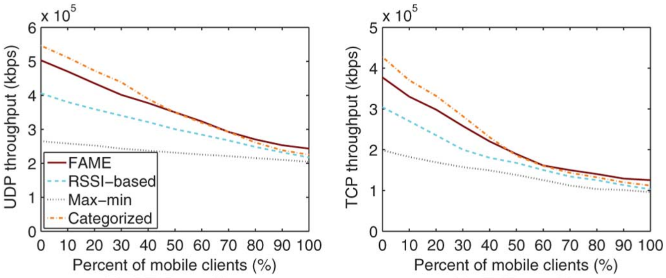
Fig. 9. Number of associated clients on each AP in an example WLANs with 200 clients: (a) uniform distribution and (b) hotspot distribution.

RSSI-based algorithm has a smaller variation. It is reasonable since the client distribution is the only factor affecting the load of each AP in that algorithm. On the other hand, as shown in Fig. 9b, the load of RSSI-based and Categorized algorithms is not so balanced compared to other algorithms under hotspot distribution, which is because that the load on APs is not considered in RSSI-based and Categorized algorithms when making association decisions. Nevertheless, the maximum number of associated clients on an AP in Categorized is still much lower than that of RSSI-based algorithm (64 vs. 94), as some clients in the hotspots need to associate with far APs whose category matches their client types, if Categorized algorithm is used.

6.2 Performance in WLANs with Mobile Clients

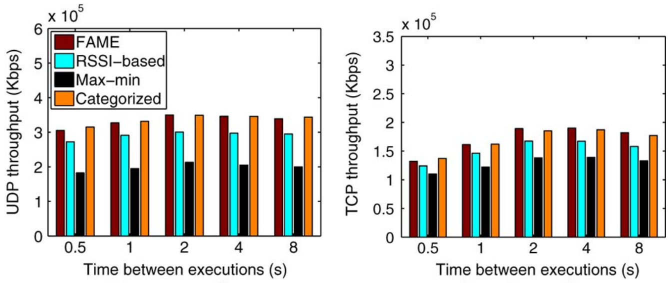
In this subsection, we evaluate the network throughput of FAME and Categorized algorithms in WLANs with mobile clients. The mobility model for mobile clients is as follows: every mobile client randomly chooses a direction and a moving speed ranging from 0 to 5 m/s, moves for a fixed duration t_1, and then pauses for a fixed duration t_2. The mobile client repeats this procedure until the end of the simulation. If the client moves into the boundary of the field during moving, it chooses a new direction and keeps moving at the same speed until duration t_1 is reached. If not otherwise specified, 50% clients are mobile and other clients are stationary. In addition, the moving duration t_1 and pausing duration t_2 are set to 6 and 4 seconds, respectively. All AP association algorithms are executed every 2 seconds.

We first study the network throughput of FAME and Categorized algorithms in WLANs with various percentages of mobile clients. The UDP and TCP throughput of all algorithms is plotted in Fig. 10, where the number of clients is fixed at 200 while the percent of mobile clients increases from zero



(a) UDP traffic. (b) TCP traffic.

Fig. 10. Network throughput in a 200-client WLANs under various percentages of mobile clients: (a) UDP traffic and (b) TCP traffic.



(a) UDP traffic. (b) TCP traffic.

Fig. 11. Impact of the duration between two consecutive executions of AP association algorithms on network throughput: (a) UDP traffic and (b) TCP traffic.

to 100%. We can see that the network throughput of both TCP and UDP traffic decreases when the percent of mobile clients increases. This is because that mobile clients need to choose low data rates for transmissions to adapt to degraded channel quality, when they move to the boundary of neighboring BSSs. In addition, TCP throughput drops more drastically than UDP throughput as the percent of mobile clients grows, which can be attributed to the additive increasing multiplicative decreasing congest control mechanism of TCP traffic. When a client moves away from its associated AP, its TCP throughput decreases quickly if a few packets are dropped due to deteriorating channel conditions; when a client moves towards its associated AP, its TCP throughput increases slowly even if the rate adaptation algorithm has already chosen higher data rates for transmission. We can also observe that when the percent of mobile clients grows beyond 60%, the network throughput of FAME algorithm becomes higher than that of Categorized algorithm. The reason is that when most clients move around the field, some clients in Categorized algorithm may fail to find APs with matched category. Then they need to share the same AP with different types of clients, which undermines the throughput gain of isolating different types of clients. Nevertheless, the proposed algorithms always lead to higher network throughput than the compared schemes regardless of the percent of mobile clients, validating the benefits of minimizing protection overhead and encouraging airtime fairness.

We then evaluate the impact of the frequency at which the AP association algorithms are executed on network throughput. The network throughput of a WLAN with 200 clients is shown in Fig. 11, in which all AP association algorithms are executed every 0.5, 1, 2, 4, and 8 seconds, respectively. It can be noted that the network throughput increases when the duration between two consecutive executions of the AP association algorithms grows from 0.5 to 2 seconds. This is because that clients cannot transmit or receive data frames when they are scanning nearby APs in all channels for re-association. Thus the network throughput is reduced if the AP association algorithms are executed too frequently. In addition, the throughput improvement is more obvious for TCP traffic when the duration increases from 0.5 to 2 seconds, as frequent scanning of nearby APs may result in time out for some TCP packets and trigger retransmissions. On the other hand, the network throughput of the proposed algorithms decreases when the duration between two consecutive executions grows beyond 4 seconds. The reason behind this is that after moving to new locations, mobile clients remain associated with far APs and use low data rates for transmission, if the association algorithms are not executed timely.

From above simulation results, we can see that in WLANs where clients are uniformly distributed or there are few mobile clients, Categorized algorithm is a better choice for deployment, as it has much lower complexity, can lead to higher network throughput while the throughput of most legacy clients is acceptable. In WLANs where clients follow the hotspot distribution or there are a number of mobile clients, FAME algorithm performs the best, as it can distribute load among APs, and provides all clients generally the same medium access time by maximizing the minimum MAC efficiency. More importantly, FAME algorithm can allocate more medium access time to legacy clients than 802.11n clients if necessary, by plugging in a positive balancing factor into FAME algorithm. In the same WLAN, both algorithms can be deployed simultaneously and the best one can be activated according to the dynamic characteristics of the network.

7 Experimental Results

In this section, we further verify the performance of the proposed algorithms in a WLAN testbed. As shown in Fig. 12, three 802.11n APs are deployed in two adjacent rooms as the testbed, with each AP operating on channel 1, channel 6 and channel 11 on 2.4Ghz frequency band, respectively. In addition, two 802.11b clients (indexed by 1 to 2), three 802.11g clients (indexed by 3 to 5) and five 802.11n clients (indexed by 6 to 10) are randomly placed in the two rooms. The wireless signal can penetrate the wall between the two rooms without obvious degradation. We run the test at midnight when no traffic is observed over the deployed WLAN of the building, thus external interference can be neglected. During the test, all

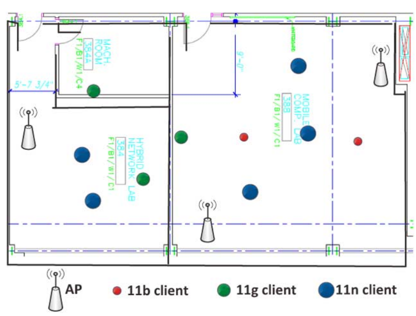
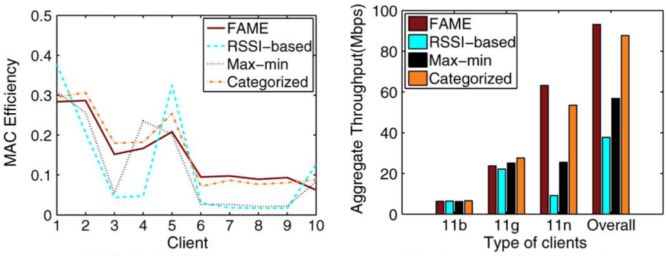
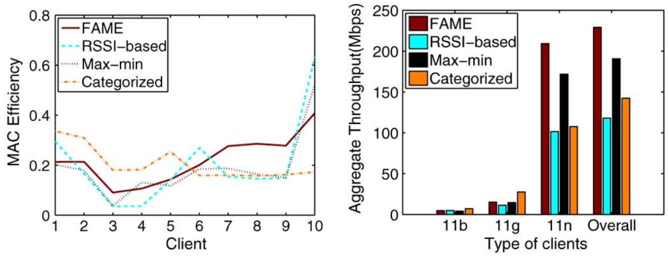


Fig. 12. Locations of APs and various clients for the 802.11n testbed.



(a) MAC efficiency. (b) Aggregated throughput.

Fig. 13. MAC efficiency and aggregated throughput for the testbed when frame aggregation of 802.11n clients is disabled: (a) MAC efficiency and (b) aggregated throughput.



(a) MAC efficiency. (b) Aggregated throughput.

Fig. 14. MAC efficiency and aggregated throughput for the testbed when frame aggregation of 802.11n clients is enabled: (a) MAC efficiency and (b) aggregated throughput.

clients transmit saturated UDP traffic to a PC, which is connected to all three APs via Gigabit Ethernet. The PC also transmits UDP traffic to every client. We have two sets of test scenarios. The frame aggregation feature of 802.11n clients is disabled in one scenario and enabled in the other.

We first study MAC efficiency and aggregated throughput for all clients when the frame aggregation feature on 802.11n clients is disabled. The test results are shown in Fig. 13. We can see that by using the proposed AP association algorithms, the MAC efficiency of all clients is more balanced, especially for clients of the same type. Note that the MAC efficiency of 802.11n clients is much lower than 802.11b clients regardless of AP association algorithm, which is inevitable since without frame aggregation, 802.11n clients take much less time to transmit data payload, while their transmitting overhead at the MAC and physical layers is identical to or even higher than that of legacy clients. As for aggregated throughput, both the overall throughput and the 802.11n throughput are greatly improved by the proposed algorithms, since the protection overhead for 802.11n clients is completely avoided. Another interesting observation is that FAME provides higher aggregated throughput for 802.11n clients compared with Categorized algorithm. This is because that with FAME, 802.11n clients share APs with 802.11g clients, resulting in more balanced load among all APs, at the cost of slightly increased preamble durations for 802.11n clients.

We now evaluate the proposed algorithms when frame aggregation on 802.11n clients is enabled, which is the default option in reality. The MAC efficiency and aggregated throughput for all clients are plotted in Figs. 14a and 14b. Similar to the previous scenario, the proposed algorithms can have more balanced MAC efficiency among all clients. However, different from the previous scenario, we notice that the MAC efficiency of 802.11n clients is comparable to or even higher than that of legacy clients for all AP association algorithms this time. The reason is that 802.11n clients can transmit many data packets in each transmission by aggregating them into one 802.11 frame, although they have the same opportunity to access the medium as other clients and even need extra overhead to protect their transmissions. In contrast, legacy clients transmit one data packet in each transmission. Note that the MAC efficiency of a 802.11g client would be very poor if it shares the AP with 802.11b and 802.11g clients, because both 802.11b and 802.11n clients occupy the medium for much longer time for each obtained transmission opportunity. This is confirmed by the aggregated 802.11g throughput of the algorithms shown in Fig. 14b. Only for Categorized algorithm, the throughput of 802.11g clients is proportional to their data rates, because the medium is not shared with either 802.11b or 802.11n clients. Furthermore, 802.11n clients can achieve high throughput even if they share an AP with 802.11b clients, with the help of frame aggregation. For the overall aggregated throughput, FAME outperforms all other algorithms by minimizing the protection overhead for 802.11n clients, reducing the impact of 802.11n and 802.11b clients on 802.11 g clients, and balancing network load at the same time.

8 Conclusions

In this paper, we have studied AP association for IEEE 802.11n based WLANs with heterogeneous clients, in particular, addressed the new challenges introduced by the high data rates and frame aggregation mechanism of 802.11n. We first showed via experiments that 802.11n throughput and overall network throughput can be severely affected by legacy clients. After that, we presented a bi-dimensional Markov model to estimate client throughput, and formulated the AP association problem into an optimization problem. Based on the Markov model introduced in the problem formulation, we provided an AP association algorithm with which each client achieves the throughput proportional to its data rate. For WLANs where APs are densely deployed, we further provided a simple but effective AP association algorithm that guarantees the performance of 802.11n transmissions by associating different type of clients with different APs. Finally, we conducted extensive simulations and experiments to assess their performance. Our simulation and experimental results demonstrate that not only high network throughput, but also fairness and balanced load can be achieved by the proposed algorithms.

References

[1] S. Lo, G. Lee, and W. Chen, "An Efficient Multipolling Mechanism for IEEE 802.11 Wireless LANs," *IEEE Trans. Computers*, vol. 52, no. 6, pp. 764-778, June 2003.
[2] Z. Zhang, Y. Yang, and M. Zhao, "Enhancing Downlink Performance in Wireless Networks by Simultaneous Multiple Packet Transmission," *IEEE Trans. Computers*, vol. 58, no. 5, pp. 706-718, May 2009.
[3] R. Liu, M. Hedley, and X. Yang, "WLAN Location Service with TXOP," *IEEE Trans. Computers*, vol. 62, no. 3, pp. 589-598, Mar. 2013.

[4] M. Ma and Y. Yang, "A Novel Contention-Based MAC Protocol with Channel Reservation for Wireless LANs," *IEEE Trans. Wireless Comm.*, vol. 7, no. 10, pp. 3748-3758, Oct. 2008.

[5] M. Balazinska and P. Castro, "Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network," *Proc. 1st Int'l Conf. Mobile Systems, Applications and Services*, pp. 303-316, May 2003.

[6] D. Schwab and R. Bunt, "Characterising the Use of a Campus Wireless Network," *Proc. 23rd Ann. Joint Conf. IEEE Computer and Comm. Soc. (INFOCOM)*, pp. 862-870, Mar. 2004.

[7] Y. Bejerano, S. Han, and L. Li, "Fairness and Load Balancing in Wireless LANs Using Association Control," *IEEE/ACM Trans. Networking*, vol. 15, no. 3, pp. 560-573, June 2007.

[8] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance Anomaly of 802.11b," *Proc. 22nd Ann. Joint Conf. IEEE Computer and Comm. Soc. (INFOCOM)*, pp. 836-843, Apr. 2003.

[9] L. Li, M. Pal, and Y.R. Yang, "Proportional Fairness in Multi-Rate Wireless LANs," *Proc. 27th IEEE Conf. Computer Comm. (INFOCOM)*, pp. 1004-1012, Apr. 2008.

[10] S.C. Liew and Y.J. Zhang, "Proportional Fairness in Multi-Channel Multi-Rate Wireless Networks-Part I: The Case of Deterministic Channels with Application to AP Association Problem in Large-Scale WLAN," *IEEE Trans. Wireless Comm.*, vol. 7, no. 9, pp. 3446-3456, Sept. 2008.

[11] IEEE 802.11n Standard, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Enhancements for Higher Throughput*. IEEE, Oct. 2009.

[12] H. Gong, K. Nahm, and J.W. Kim, "Distributed Fair Access Point Selection for Multi-Rate IEEE 802.11 WLANs," *Proc. 5th IEEE Consumer Comm. Networking Conf.*, pp. 528-532, Jan. 2008.

[13] O. Ercetin, "Association Games in IEEE 802.11 Wireless Local Area Networks," *IEEE Trans. Wireless Comm.*, vol. 7, no. 12, pp. 5136-5143, Dec. 2008.

[14] W. Xu, C. Hua, and A. Huang, "A Game Theoretical Approach for Load Balancing User Association in 802.11 Wireless Networks," *Proc. 2010 IEEE Global Telecomm. Conf. (GLOBECOM'10)*, pp. 1-5, Dec. 2010.

[15] B.K. Pradeepa and J. Kuri, "An Estimated Delay Based Association Policy for Web Browsing in a Multirate WLAN," *IEEE Trans. Network and Service Management*, vol. 9, no. 3, pp. 1-13, June 2012.

[16] G.S. Kasbekar, P. Nuggehalli, and J. Kuri, "Online Client-AP Association in WLANs," *Proc. 4th Int'l Symp. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pp. 1-8, Apr. 2006.

[17] N. Papaoulakis and C.Z. Patrikakis, "A Proactive, Terminal Based Best Access Point Selection Mechanism for Wireless LANs," *Proc. IEEE Int'l Conf. Global Comm. (GLOBECOM) Workshops*, pp. 1-4, Dec. 2008.

[18] L. Du, Y. Bai, and L. Chen, "Access Point Selection Strategy for Large-Scale Wireless Local Area Networks," *Proc. IEEE Wireless Comm. Networking Conf. (WCNC)*, pp. 2161-2166, Mar. 2007.

[19] S. Keranidis, T. Korakis, I. Koutsopoulos, and L. Tassiulas, "Contention and Traffic Load-Aware Association in IEEE 802.11 WLANs: Algorithms and Implementation," *Proc. Int'l Symp. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pp. 334-341, May 2011.

[20] P. Bahl, M.T. Hajiaghayi, K. Jain, S.V. Mirrokni, L. Qiu, and A. Saberi, "Cell Breathing in Wireless LANs: Algorithms and Evaluation," *IEEE Trans. Mobile Computing*, vol. 6, no. 2, pp. 164-178, Feb. 2007.

[21] B. Kauffmann, F. Baccelli, A. Chaintreau, V. Mhatre, K. Papagiannaki, and C. Diot, "Measurement-Based Self Organization of Interfering 802.11 Wireless Access Networks," *Proc. 26th IEEE Int'l Conf. Computer Comm. (IEEE INFOCOM'07)*, pp. 1451-1459, May 2007.

[22] W. Xu, C. Hua, and A. Huang, "Channel Assignment and User Association Game in Dense 802.11 Wireless Networks," *Proc. IEEE Int'l Conf. Comm. (ICC)*, pp. 1-5, June 2011.

[23] T. Bonald, A. Ibrahim, and J. Roberts, "The Impact of Association on the Capacity of WLANs," *Proc. 7th Int'l Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, pp. 1-10, June 2009.

[24] M. Arslan, K. Pelechrinis, I. Broustis, S. Krishnamurthy, S. Addepalli, and K. Papagiannaki, "ACORN: Auto-Configuration framework for 802.11n WLANs," *IEEE/ACM Trans. Networking*, pp. 896-909, vol. 21, no. 3, June 2013.

[25] C. Huang and P. Ramanathan, "Joint Association and Power Adaption in WiFi Community Networks," *Proc. 45th Ann. Conf. Information Sciences and Systems*, pp. 1-6, Mar. 2011.

[26] H. Ko, J. Shin, D. Kwak, and C. Kim, "A Joint Approach to Bandwidth Allocation and AP-Client Association for WLANs," *Proc. 35th IEEE Local Computer Networks*, pp. 576-581, Oct. 2010.

[27] T. Selvam and S. Srikanth, "Performance Study of IEEE 802.11n WLANs," *Proc. 1st Int'l Comm. Systems and Networks and Workshops (COMSNETS)*, pp. 1-6, Jan. 2009.

[28] V. Shrivastava, S. Rayanchu, J. Yoonj, and S. Banerjee, "802.11n under the Microscope," *Proc. 8th ACM SIGCOMM Conf. Internet Measurement*, pp. 105-110, Oct. 2008.

[29] M. Laddomada, F. Mesiti, M. Mondin, and F. Daneshgaran, "On the Throughput Performance of Multirate IEEE 802.11 Networks with Variable-Loaded Stations: Analysis, Modeling, and a Novel Proportional Fairness Criterion," *IEEE Trans. Wireless Comm.*, vol. 9, no. 5, pp. 1594-1607, May 2010.

[30] K. Hung, B. Bensaou, and R. Li, "Cross Layer Association Control for Throughput Optimization in Wireless LANs with Inter-AP Interference," *Proc. 13th ACM Int'l Conf. Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWIM)*, pp. 210-217, Oct. 2010.

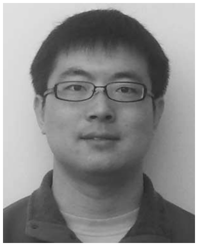
[31] G. Bhanage, D. Vete, I. Seskar, and D. Raychaudhuri, "SplitAP: Leveraging Wireless Network Virtualization for Flexible Sharing of WLANs," *Proc. IEEE Global Telecomm. Conf. (GLOBECOM'10)*, pp.1-6, Dec. 2010.

[32] C. Chou, K.G. Shin, and N. Sai Shankar, "Contention-Based Airtime Usage Control in Multirate IEEE 802.11 Wireless LANs," *IEEE/ACM Trans. Networking*, vol. 14, no. 6, pp. 1179-1192, Dec. 2006.

[33] T. Joshi, A. Mukherjee, Y. Yoo, and D.P. Agrawal, "Airtime Fairness for IEEE 802.11 Multirate Networks," *IEEE Trans. Mobile Computing*, vol. 7, no. 4, pp. 513-527, Apr. 2008.

[34] J. Kim, S. Kim, S. Choi, and D. Qiao, "CARA: Collision-Aware Rate Adaptation for IEEE 802.11 WLANs," *Proc. 25th IEEE Int'l Conf. Computer Comm. (IEEE INFOCOM'06)*, pp. 1-11, Apr. 2006.

[35] R. Jain, D. Chiu, and W. Hawe, "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Systems," DEC Research Report TR-301, Sept. 1984.



Dawei Gong received the BEng degree from Zhejiang University, China, and the MS degree from Tsinghua University, Beijing, China. He is currently a PhD student in the Department of Electrical and Computer Engineering, Stony Brook University, New York. His research interests include wireless sensor networks and wireless LANs.

Yuanyuan Yang received the BEng and MS degrees in computer science and engineering from Tsinghua University, Beijing, China, and the MSE and PhD degrees in computer science from Johns Hopkins University, Baltimore, Maryland. She is a professor of computer engineering and computer science at Stony Brook University, New York, and the director of Communications and Devices Division at New York State Center of Excellence in Wireless and Information Technology (CEWIT). Her research interests include wireless networks, optical networks, high-speed networks, and parallel and distributed computing systems. Her research has been supported by the US National Science Foundation (NSF) and US Army Research Office (ARO). She has published more than 280 papers in major journals and refereed conference proceedings and holds seven US patents in these areas. She is currently an associate editor-in-chief for the *IEEE Transactions on Computers* and a subject area Editor for the *Journal of Parallel and Distributed Computing*. She has served as an associate editor for the *IEEE Transactions on Computers* and *IEEE Transactions on Parallel and Distributed Systems*. She has served as a general chair, program chair, or vice chair for several major conferences and a program committee member for numerous conferences.