Access Point Selection for Multi-Rate IEEE 802.11 Wireless LANs

Shicong Yang
Avideh Zakhor

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2014-104
http://www.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-104.html

May 16, 2014

Access Point Selection for Multi-Rate IEEE 802.11 Wireless LANs

by

Shicong Yang

A thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science
in

Engineering - Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Avideh Zakhor, Chair
Professor Jean Walrand
Professor Sylvia Ratnasamy

Spring 2014

The thesis of Shicong Yang, titled Access Point Selection for Multi-Rate IEEE 802.11 Wireless LANs, is approved.

Chair Date

Date

Date

University of California, Berkeley

Spring 2014

Access Point Selection for Multi-Rate IEEE 802.11 Wireless LANs

by

Shicong Yang

Abstract

Access Point Selection for Multi-Rate IEEE 802.11 Wireless LANs

by

Shicong Yang

Master of Science in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Avideh Zakhor, Chair

Access Point (AP) selection is an important problem in WLANs as it affects the throughput of the joining station (STA). Existing approaches to AP selection predominantly use received signal power and as such, do not take into account interference and collisions at each STA, and transmit opportunities (TXOPs) at APs. In this paper, we propose a class of AP selection algorithms based on the joining STA's expected throughput by taking into account the above factors. Specifically, we collect a binary-valued local channel occupancy signal, called busy-idle (BI) signal, at each node and require the APs to periodically broadcast their BI signal together with a quantity that represents TXOPs at the APs. This enables the joining STA to estimate expected throughput for each candidate AP before it selects one. To capture TXOPs at each AP, we propose a few different quantities, including the number of associated STAs, sum of inverse of MAC rates, and the average waiting and idle times. We use NS-2 simulations to demonstrate the effectiveness of our algorithms. For a random topology consisting of 24 APs and 60 STAs, our algorithms can increase average throughput of the joining STA by as much as 52% as compared to the traditional signal power based AP selection approach (*rxpwr*). In this case, the achieved average throughput is 97% of that obtained via the optimal selection. We also show that, in contrast to *rxpwr*, the throughput of our proposed algorithms remain close to optimal with the increase in AP density or STA density.

To Yu Shi and my family

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to express my very great appreciation to my research advisor, Professor Avideh Zakhor, for her continued guidance and support. She has taught me how to do research, from choosing topics, gaining intuitions, solving problems, and eventually to presenting results. Without her invaluable help, this thesis is not possible and I would not be close to where I am right now.

I would also like to thank Professor Jean Walrand and Professor Sylvia Ratnasamy for serving on my thesis committee. Professor Jean Walrand taught me fundamental and important tools in probabilities and random processes which provide building blocks for my research. Professor Sylvia Ratnasamy provided valuable comments which made this thesis better and complete.

My thesis would not be possible without the help from other Professors and friends. I am truly grateful for Professor Adam Wolisz and Professor Jan Rabaey for spending time to meet, discuss and feedback on this work. Their valuable suggestions helped steer this research to a more practical setting. Thank Plamen Levchev, Mo Chen, and Jiaqi Zhao, for the help during my thesis revision.

The funding for this thesis was provided by ARO MURI W911NF-08-1-0233. This thesis was also partially supported by the National Science and Engineering Research Council of Canada (NSERC). I really appreciate the generosity of ARO and NSERC for their financial support to enable me to accomplish this thesis.

Thanks to administrative staff at EECS Berkeley, especially to Shirley Salanio and Jennifer Gardner, for their patient help on administrative questions and requests.

My graduate student life would not be as exciting without other members of my group. In particularly thanks to Michael Krishnan, whose work formed the ground for my master thesis. I also want to thank Ricardo Garcia for his advice and suggestion in both professional and personal life. My thanks also extends to the other group members, in no particular order, Nicholas Corso, Eric Turner, Shangliang Jiang, John Kua, Richard Zhang, and Plamen Levchev. Also thanks to other friends I met in Berkeley for making my graduate life happy and memorable. In no particular order, this goes to Mo Chen, Ka Kit Lam, Chi Pang Lam, Stephan Adams, Vijay Kamble, Ramtin Pedarsani, Qie Hu, Alonso Silva, Longbo Huang, and Ana Ferreira.

Special thanks to my famliy. Thanks to my parents for their everlasting support and encouragement to make me a better person. This thesis is also dedicated to my grandparents, who passed away during my graduate study. They were there for me during my early stage of life, but I was not able to accompany them during their last stage of lives. I wish I had spent more time with them. Wish they rest in peace.

Last but not least, I am indebted to Yu Shi, who has provided enormous patience, support and encouragement during my graduate study at Berkeley, though I had to spend more time on study and research. You are the most valuable asset I discovered during my graduate study.

Chapter 1

Introduction

Wireless local area networks (WLANs) have gained increasing popularity due to their convenience, flexibility, and mobility as compared to traditional wireline infrastructure. As a result, WLANs are becoming the preferred technology of high-speed broadband access in homes, offices, and other hotspots such as coffee shops, shopping malls, and airports. Each WLAN access point (AP) forms a Basic Service Set (BSS), and multiple BSSs can overlap to form an Extended Service Set (ESS) to provide seamless handoff for stations (STAs). Due to the dense deployment of WLANs and the use of ESS to provide roaming services, it is common for STAs to have multiple available APs to choose from. In addition, nearby BSSs often experience inter-BSS co-channel interference due to the limited number of orthogonal channels. The MAC rate and throughput for different APs can vary significantly depending on the physical channel conditions and the interference level. An inappropriate AP selection typically leads to compromised service, thus it is imperative for an STA to identify and select the AP that provides the highest data rate to improve user experience.

AP selection policy is not specified in IEEE 802.11 standards. Currently the most widely used scheme is to select the AP with the strongest received power. Stronger received signal implies that the wireless channel is in better condition and can potentially support higher MAC rates, resulting in higher throughput for STAs. While this strategy is straightforward and easy to implement with no modifications and overhead to existing standards, it is ineffective especially in hyper dense deployment scenarios where adjacent APs could use the same channel. For example, as shown in Figure 1.1, consider two nearby co-channel APs who cannot sense each other, and an "joining" STA within range of both APs. Assume without loss of generality, the joining STA is closer to AP 2 and hence experiences a higher received signal power from it as compared to AP 1. The traditional received signal power based method would result in selecting AP 2, even though it could experience more interference and a lower throughput. Furthermore, the strongest received signal power based algorithm cannot take into account the difference between the potential transmit opportunities (TXOPs) from the two APs, which is another important factor in determining

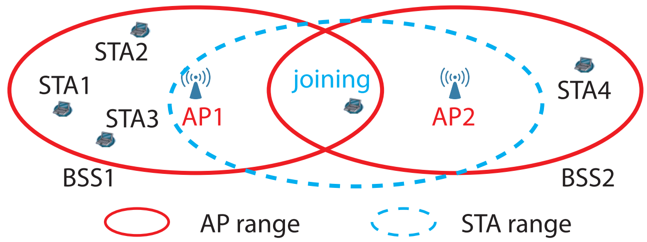


Figure 1.1. APs with Overlapping BSSs.

throughput. It is clear from this example that choosing an AP with the strongest received power is sub-optimal for hyper dense deployment of WLANs, and that interference, collisions and TXOPs should also be taken into account when selecting an AP among multiple available APs.

To address these issues, a variety of schemes have been proposed in the literature [1]–[18], many of which try to optimize potential throughput and bandwidth. Nicholson *et al.* propose that STAs quickly associate with each AP and run a battery of tests to estimate the quality of each AP's connection [1]. Vasudevan *et al.* propose to use potential bandwidth as a metric to facilitate AP selection [2], where beacon delays are used to estimate potential uplink (UL) and downlink (DL) bandwidth. Sundaresan *et al.* propose to optimize AP selection based on expected throughput obtained from cross-layer information [3]. Abusubaih *et al.* consider the effect of newly arrived STA on total network UL and DL throughput [4]. Similarly, Miyata *et al.* propose an AP selection algorithm to optimize total network throughput as well as preserving newly arrived STAs throughput [5]. Luo *et al.* consider wireless mesh networks and propose that a STA should make its association based on end-to-end performance [6]. A number of approaches take interference and collision into account for selecting APs. For instance, Fukuda *et al.* propose to avoid interference when selecting an AP [7]. Du *et al.* propose a metric to capture the effect of hidden nodes and multiple MAC rates [8]. They use the channel utilization field in the beacon packets and suggest that the difference between the AP's and the STA's respective channel usage captures the hidden node effect. Abusubaih *et al.* consider interference between BSSs and develop a metric based on collision probability to facilitate AP selection [9]. Jang *et al.* exploit the retry field in the MAC header to estimate collision probability, and propose to use expected throughput as a metric to choose APs [10]. Some researchers approach AP selection problem from fairness point of view. For example, Bejerano *et al.* propose to select AP for max-min fair bandwidth allocation [11]. Gong *et al.* further propose a distributed max-min throughput AP selection [12]. Zhou *et al.* consider multi-AP wireless hotspots and propose a new fairness notion called Fulfillment-based Fairness to select AP [13]. Judd *et al.* notice AP load imbalance problem for received signal strength based AP selection algorithm [14]. To alleviate this problem, Chen *et al.* propose to use probe delay to capture the load and probability of collisions on each AP [15]. They argue that higher load results in higher collision probability and therefore longer backoff time, so the probe frame delay increases when traffic load is heavy. Moreover, Bahl *et al.* propose to utilize the well-known cell breathing concept in cellular telephony to balance load in WLANs [16]. Other work try to study the AP selection problem using game theory tools. Musacchio *et al.* approach wireless AP selection from the economic point of view and model the problem as a dynamic game [17]. Mittal *et al.* present a game-theoretic analysis of wireless AP selection by selfish STAs [18]. Jiang *et al.* shows joining the correct AP can be achieved by distributed solution under certain model [19].

In this thesis, we propose a class of AP selection algorithms to maximize a joining STA's DL expected throughput. Our proposed AP selection metrics not only consider TXOPs at APs, but also take into account the inter-BSS interference with a more accurate collision estimation technique. We use the framework in [20], which provides a method to estimate the collision probability for UL traffic at a given STA. The basic idea behind [20] is that all STAs and APs continually measure the spatial channel occupancy around them, with APs periodically broadcasting a compressed binary-valued busy-idle (BI) signal to indicate their local channel occupancy to all associated STAs. Each STA can then estimate UL collision probability by comparing its local BI signal with that of the AP's. We extend this framework to estimate DL collision probability, and then compute decision metrics at the joining STA to select an AP.

Throughout the thesis, we use STA to refer to a non-AP station, and use *node* to refer to either an AP or an STA. The remainder of the thesis is organized as follows: Chapter 2 discusses our packet loss model, and the method to estimate each component of packet loss; Chapter 3 describes our proposed algorithms; Chapter 4 presents the performance evaluations, and Chapter 5 concludes the thesis.

Chapter 2

Packet Loss Modeling and Estimation

We categorize packet loss in WLANs into two classes: collisions and channel errors. A collision is defined as a packet failure at the intended receiver due to interference from other transmitters which are in close proximity to the receiver. A channel error is defined as an unsuccessful decoding of a packet due to low received SNR, which is caused by large path loss or deep multipath fade, given that the packet does not suffer from collisions. The probability of total packet loss can be expressed as:

$$P_L = 1 - (1 - P_C)(1 - P_e) \tag{2.1}$$

where P_C is the packet loss probability due to collisions, and P_e is the packet loss probability due to channel error given that the packet does not experience collisions. Equivalently, the packet success rate P_S is given by:

$$P_S = 1 - P_L = (1 - P_C)(1 - P_e) \tag{2.2}$$

In this thesis, we assume none of the packets suffering from collisions are captured, and are therefore assumed to be lost.

Krishnan *et al.* proposed a framework to estimate UL collision probabilities at STAs, using the local channel occupancy at the STA as well as the periodically broadcasted BI signal associated with the STA's AP, which reflects the AP's local channel occupancy [20]. We now generalize the estimator in [20] to estimate the collision probability on link (Tx, Rx) as follows:

$$P_C(Tx, Rx) = f(BI_{Tx}, BI_{Rx}) \tag{2.3}$$

where BI_{Tx} and BI_{Rx} are BI signals collected at the transmitter and the receiver, respectively. For DL, suppose AP i is the Tx and STA j is the Rx, hence:

$$P_C(i, j) = f(BI_{AP_i}, BI_{STA_j}) \tag{2.4}$$

We classify collisions into three types: direction collisions (DCs), staggered collisions of type 1 (SC1), and staggered collisions of type 2 (SC2) [20]. A DC for a given node is a collision in which the node under consideration finishes its backoff period and starts transmitting at the same time as other nodes. An SC1 for a given node is a collision in which the node under consideration transmits first and is then interrupted by a hidden node. An SC2 for a given node is a collision in which the node under consideration interrupts the transmission of a hidden node. Intuitively, for the node under consideration, an SC2 occurs when another node is already transmitting to the intended receiver before the node starts to transmit, a DC occurs when another node starts transmitting at the same time the node starts to transmit, and an SC1 occurs when another node starts transmitting later than, but interrupts, the node's transmission. Based on the above description, $(1 - P_C)$ can be expanded into [20]:

$$(1 - P_C) = (1 - P_{SC2})(1 - P_{DC})(1 - P_{SC1}) \tag{2.5}$$

where P_{SC2} denotes the probability of SC2, P_{DC} denotes the probability of DCs given that it does not experience SC2, and P_{SC1} denotes the probability of SC1 given that it experiences neither SC2 nor DC [20]. Due to the way collisions are counted, SC2 is the dominant type of collision for high traffic scenarios [20], and can therefore be used to approximate the total DL collision probability in a traffic-saturated WLAN network as:

$$P_C(i,j) \approx P_{SC2}(i,j) = \frac{\sum_t \mathbb{1}\{BI_{APi}(t) = 0, BI_{STAj}(t) = 1\}}{\sum_t \mathbb{1}\{BI_{APi}(t) = 0\}} \tag{2.6}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. The intuition is that this is the probability that the channel is busy at the STA given that it is idle at the AP, and hence if at time t a packet was transmitted by the AP when AP senses the channel to be idle, i.e., $BI_{APi}(t) = 0$, it would have experienced collision at the STA with probability $P_C(i,j)$.

An 802.11 packet uses PHY modulation rate R_{PHY} for preamble and PLCP header, and potentially higher modulation rates R_{MAC} for MAC frame. The probability of channel error for packets from AP i to STA j can be expressed as [21]

$$\begin{aligned} P_e(i,j) = 1 - &(1 - BER_{R_{\mathrm{PHY}}}(SNR_{ij}))^{L_{\mathrm{PHY}}} \\ &(1 - BER_{R_{\mathrm{MAC}}}(SNR_{ij}))^{L_{\mathrm{MAC}}} \end{aligned} \tag{2.7}$$

where L_{PHY} and L_{MAC} are the lengths of the preamble and PLCP header, and MAC frame, respectively. $BER_R(SNR)$ denotes the bit error rate which is assumed to be a known function of modulation rate R and SNR. SNR_{ij} can be estimated as:

$$SNR_{ij} = \frac{Pr_{ij}}{Noise} \tag{2.8}$$

where Pr_{ij} is the received power of beacon packets from AP i to STA j, and $Noise$ is the thermal noise that can be estimated from:

$$Noise(\mathrm{dBm}) = -174 + 10\log_{10}(W) + N_f \tag{2.9}$$

where W is the bandwidth of wireless transmission, and N_f is the noise figure of the wireless system, which is a property of hardware. Substituting Equation (2.9) into Equation (2.8), SNR_{ij} can be estimated and consequently the channel error probability $P_e(i,j)$ can be computed as Equation (2.7).

With the estimates of collision probability $P_C(i,j)$ and channel error probability $P_e(i,j)$, the total loss probability $P_L(i,j)$ can be computed as Equation (2.1). We use $P_L(i,j)$ to estimate average backoff time if STA j associates with AP i, and to compute our proposed decision metric in Chapter 3.

Chapter 3

Proposed AP Selection Algorithm

In this chapter we describe a class of AP selection algorithms which take into account the TXOPs, MAC rates, interference and collisions at the STA. We begin by describing our system model. We assume WLAN operates in infrastructure mode with DCF, and hence no RTS/CTS is used. All traffic flows have the same priority, and packets have Poisson arrival whose rate depends on the application layer data rate. When serving MAC Service Data Units (MSDUs), an AP does not switch to a new MSDU until the previous MSDU is successful or dropped due to its retransmission limit being exceeded. The network is assumed to be saturated, i.e., the application layers always have backlogs in their queues. The MAC rate is determined by the path loss from an STA to its serving AP, and no rate adaptation is assumed to be used. In this analysis we assume APs to be on the same channel, and focus on one STA j joining the network while all other STAs are already associated to and exchanging traffic with their desired APs. For ease of notation, we use $P_C(i)$, $P_e(i)$ and $P_S(i)$ in place of $P_C(i,j)$, $P_e(i,j)$ and $P_S(i,j)$, respectively, since only one joining STA j is considered.

In our proposed algorithm, both APs and the joining STA record their BI signals at a resolution of $10\mu sec$ as suggested in [20]; this sampling period provides a good balance between estimation error and transmission overhead. APs broadcast the BI signals every $3sec$ with the overhead to send BI signal being about 3% in the 802.11b network [20]. Before associating to any AP, the joining STA stays idle and records its local BI signal for the first $3sec$.

Our approach to AP selection is to maximize DL expected throughput (eTP):

$$AP_{sel} = arg \max_{i \in \mathcal{A}} (eTP(i)) \tag{3.1}$$

where $\mathcal{A}$ is the set of candidate APs that the joining STA can choose from. We define $eTP(i)$ as:

$$\begin{aligned} eTP(i) :&= \frac{\text{total successful MSDU in bits from AP } i \text{ to joining STA}}{\text{total time}} \\ &= \frac{\text{total time to send MSDU by AP } i \text{ to joining STA}}{\text{total time}} \end{aligned}$$

$$\times \frac{\text{total successful MSDU in bits from AP } i \text{ to joining STA}}{\text{total time to send MSDU by AP } i \text{ to joining STA}} \tag{3.2}$$

$$= t_{\text{alloc}}(i) \times eTMR(i) \tag{3.3}$$

where we denote the first term in Equation (3.2) as $t_{\text{alloc}}(i)$, representing the percentage of channel time that AP i can allocate to the joining STA, and define the second term as ($eTMR(i)$), representing the expected true MAC rate from AP i to the joining STA. Note the MSDU in Equation (3.2) is also known as MAC payload, and we will use them interchangeably in this thesis. We refer the second term as "true" MAC rate because it measures the real MAC layer data rate that the STA experiences by counting only the successfully delivered MAC payloads. In contrast, the nominal MAC rate is the MAC layer data rate that is used by an AP to modulate packets to an STA, which does not take packet loss and overhead into account. The potential TXOP from AP i is captured in $t_{\text{alloc}}(i)$ term, while the effect of MAC rate, interference and collisions is taken into account in both $t_{\text{alloc}}(i)$ and $eTMR(i)$. We elaborate on how to estimate these two terms in the following sections.

3.1 Estimating $eTMR(i)$

The $eTMR(i)$ from AP i to the joining STA is defined as successful number of MAC payload bits transmitted over the time that AP i spent for delivering those data, including packet transmission time and all associated overhead time. $eTMR(i)$ can be expressed as:

$$\begin{aligned} eTMR(i) := & \frac{\text{total successful MSDU in bits from AP } i \text{ to joining STA}}{\text{total time to send MSDU by AP } i \text{ to joining STA}} \\ = & \frac{\sum_m L_i(m) \times \mathbb{1}\{A_i(m)\}}{\sum_m t_i(m)} \end{aligned} \tag{3.4}$$

where $L_i(m)$ is the MSDU size in bits from AP i on the mth Physical layer Protocol Data Unit (PPDU) transmission, $t_i(m)$ is the time that AP i spent on the mth PPDU transmission, backoff and protocol overhead, $A_i(m)$ is the event that the mth PPDU sent by AP i to the joining STA is successful, and $\mathbb{1}\{\cdot\}$ is the indicator function defined by:

$$\mathbb{1}\{A_i(m)\} = \begin{cases} 1 & \text{if } m\text{th PPDU sent by AP } i \text{ succeeds} \\ 0 & \text{if } m\text{th PPDU sent by AP } i \text{ fails} \end{cases} \tag{3.5}$$

If we assume the maximum MAC payload size L is used for each packet, Equation (3.4) can be rewritten as:

$$\begin{aligned} eTMR(i) = & \frac{L \times \sum_m \mathbb{1}\{A_i(m)\}}{\sum_m t_i(m)} \\ = & \frac{L \times P_S(i)}{t(i)} \end{aligned} \tag{3.6}$$

where $P_S(i)$ is the packet success probability from AP i to the joining STA given by Equation (2.2), $t(i)$ is the average time that AP i allocates to the joining STA for one packet transmission including backoff time, PPDU transmission time and protocol overhead, given by:

$$t(i) = t_p(i) + t_{OH}(i) \tag{3.7}$$

where $t_p(i)$ is the time for AP i to transmit MAC payload to the joining STA, and $t_{OH}(i)$ is the average overhead of one MSDU transmission from AP i. Substituting Equation (3.7) into (3.6) and rearranging the terms, we obtain:

$$\begin{aligned} eTMR(i) &= \frac{L \times P_S(i)}{t(i)} = \frac{L \times P_S(i)}{t_p(i) + t_{OH}(i)} \\ &= \frac{L}{t_p(i)} \times P_S(i) \times \frac{t_p(i)}{t_p(i) + t_{OH}(i)} \\ &= R_{\text{MAC}}(i) \times P_S(i) \times \frac{t_p(i)}{t_p(i) + t_{OH}(i)} \end{aligned} \tag{3.8}$$

where $R_{\text{MAC}}(i) = L/t_p(i)$ is the MAC rate used by AP i for the joining STA to modulate MAC payload. Substituting Equation (2.2) into (3.8), we obtain the following expression for $eTMR(i)$ at AP i:

$$eTMR(i) = R(i) \times (1 - P_C(i)) \times (1 - P_e(i)) \times \frac{t_p(i)}{t_p(i) + t_{OH}(i)} \tag{3.9}$$

where $P_C(i)$ is the DL collision probability given by Equation (2.6), and $P_e(i)$ is the DL channel error probability given by Equation (2.7). Next we explain how to estimate each component in Equation (3.9) in order for the joining STA to optimally select the AP.

$R(i)$ depends on SNR_i from AP i to the joining STA. Assuming the function to map SNR to MAC rate is known, the MAC rate $R(i)$ used by AP i can be predicted as long as SNR is estimated as in Equation (2.8).

$t_p(i)$ is the time to transmit MAC payload, i.e., MSDU. It depends on the payload size L in bits and MAC rate $R_{\text{MAC}}(i)$:

$$t_p(i) = \frac{L}{R_{\text{MAC}}(i)} \tag{3.10}$$

Maximum payload is typically used in WLANs to improve transmission efficiency.

$t_{OH}(i)$ is the average overhead time for AP i to deliver one MSDU, which includes preamble, PLCP header, MAC header and CRC, inter-frame spacing time, possible ACK time, and backoff time. In 802.11 each MSDU has a predefined retransmission limit N. Let $n \in [0, \cdots, N-1]$ be the retransmission state of an MSDU, $\pi(n)$ be the probability that a node is in retransmission state n, and $t_{OH}(i, n)$ be the overhead for a single MSDU transmission on the nth retry. Then $t_{OH}(i)$ is given by:

$$t_{OH}(i) = \sum_{n=0}^{N-1} \pi(n) t_{OH}(i, n) \tag{3.11}$$

and $t_{OH}(i,n)$ is given by:

$$t_{OH}(i,n) = \bar{t_{\mathrm{b}}}(i,n) + t_h(i) + t_{\mathrm{protocol}} \tag{3.12}$$

where $\bar{t_{\mathrm{b}}}(i,n)$ is the average backoff duration on the nth MSDU retry when AP i is transmitting to the joining STA, $t_h(i)$ is the time to transmit the PHY layer preamble, the PLCP header, the MAC header, and MAC CRC, and t_{protocol} is the overhead introduced by the 802.11 protocol for a single PPDU transmission, including inter-frame spacing, backoff time, and possible ACK time given by:

$$t_{\mathrm{protocol}} = \begin{cases} T_{\mathrm{SIFS}} + T_{\mathrm{ACK}} + T_{\mathrm{DIFS}} & \text{if PPDU succeeds} \\ T_{\mathrm{EIFS}} & \text{if PPDU fails} \end{cases} \tag{3.13}$$

where T_{ACK} is the time to transmit an ACK packet if the ACK packet is modulated at lowest rate, T_{SIFS} is the short inter-frame spacing, T_{DIFS} is the DCF inter-frame spacing, and T_{EIFS} is the extended inter-frame spacing defined by the 802.11 standards as:

$$T_{\mathrm{EIFS}} = T_{\mathrm{SIFS}} + T_{\mathrm{ACK}} + T_{\mathrm{DIFS}} \tag{3.14}$$

This implies that t_{protocol} in Equation (3.13) is the same regardless of whether or not the packet transmission is successful, if the ACK packet is modulated at lowest rate.

$t_h(i)$ in Equation (3.12) does not depend on the retransmission state n, and can be calculated as:

$$t_h(i) = \frac{L_{pre}}{R_{\mathrm{PHY}}} + \frac{L_{plcp}}{R_{\mathrm{PHY}}} + \frac{L_{mo}}{R_{\mathrm{MAC}}(i)} \tag{3.15}$$

where L_{pre} is the length of preamble in bits, L_{plcp} is the length of PLCP header in bits, L_{mo} is the length of MAC header and CRC in bits, R_{PHY} is the PHY modulation rate defined for a particular 802.11 protocol, and $R_{\mathrm{MAC}}(i)$ is the modulation rate used for MAC frame by AP i. Combining Equations (3.11) and (3.12), we obtain:

$$\begin{aligned} t_{OH}(i) &= \sum_{n=0}^{N-1} \pi(n) t_{OH}(i,n) \\ &= \sum_{n=0}^{N-1} \pi(n) \times (\bar{t_{\mathrm{b}}}(i,n) + t_h(i) + t_{\mathrm{protocol}}) \\ &= t_h(i) + t_{\mathrm{protocol}} + \sum_{n=0}^{N-1} \pi(n) \bar{t_{\mathrm{b}}}(i,n) \end{aligned} \tag{3.16}$$

Denoting $\bar{t_{\mathrm{b}}}(i)$ to be the average backoff duration for packets from AP i to the joining STA:

$$\bar{t_{\mathrm{b}}}(i) = \sum_{n=0}^{N-1} \pi(n) \bar{t_{\mathrm{b}}}(i,n) \tag{3.17}$$

Equation (3.16) can be simplified to:

$$t_{OH}(i) = t_h(i) + t_{\mathrm{protocol}} + \bar{t_{\mathrm{b}}}(i) \tag{3.18}$$

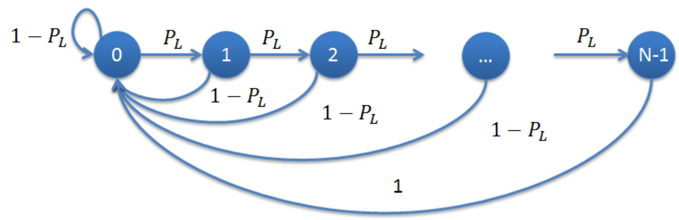


Figure 3.1. Markov Chain Model for 802.11 DCF BEB.

We now explain how $\pi = (\pi(0), \cdots, \pi(n), \cdots, \pi(N-1))$ and $\bar{t_{\text{b}}}(i,n)$ can be estimated. IEEE 802.11 standards use Binary Exponential Backoff (BEB) to achieve random channel access. Before each node transmits, it is required to wait for a random amount of time slots uniformly chosen from $[0, CW(n)]$, where $CW(n)$ is the contention window size on the nth retry:

$$CW(n) = \min\left\{(CW_{\min} + 1) \times 2^n - 1, CW_{\max}\right\} \tag{3.19}$$

where $CW_{\min}$ and $CW_{\max}$ are predefined maximum and minimum contention window size, respectively. The average backoff interval on the nth retry is:

$$\bar{t_{\text{b}}}(i,n) = \frac{CW(n)}{2} \times T_{\text{slot}} \tag{3.20}$$

where T_{slot} is the slot time defined for a particular 802.11 PHY layer.

To estimate $\pi(n)$, the probability that a node is in retransmission state n, we model the BEB behavior for MSDUs from AP i to the joining STA by a Markov Chain (MC) shown in Figure 3.1. The MC state $n \in [0, \cdots, N-1]$ is the retransmission count of an MSDU, and π is the stationary distribution of the MC. The $N \times N$ transition matrix $P(i)$ of the MC for AP i can be constructed as:

$$P(i) = \begin{pmatrix} 1-P_L(i) & P_L(i) & 0 & 0 & \cdots & 0 \\ 1-P_L(i) & 0 & P_L(i) & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 1-P_L(i) & 0 & 0 & 0 & \cdots & P_L(i) \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{3.21}$$

where $P_L(i)$ is given by Equation (2.1). The stationary distribution π of MC is the eigenvector of $P(i)$ corresponding to eigenvalue of 1, and can therfore be estimated from $P_L(i)$. Combining Equations (3.17), (3.18) and (3.20), the average overhead time is

$$t_{OH}(i) = t_h(i) + t_{\text{protocol}} + \sum_{n=0}^{N-1} \pi(n) CW(n) T_{\text{slot}}/2 \tag{3.22}$$

Once $t_{OH}(i)$ and the corresponding values in Equations (2.6), (2.7), and (3.10) are estimated by the joining STA, Equation (3.9) can be used to evaluate $eTMR(i)$.

3.2 Estimating $t_{\textbf{alloc}}(i)$

The $t_{\text{alloc}}(i)$ is a unit-less term representing the expected percentage of DL channel time that AP i can allocate to the joining STA. We propose four different ways to estimate this quantity, resulting in four different decision metrics for selecting AP:

1. $eTMR$: $t_{\text{alloc}}(i)$ is a constant.
2. eTP_n: $t_{\text{alloc}}(i)$ is estimated by the number of STAs that are already associated with AP i.
3. eTP_r: $t_{\text{alloc}}(i)$ is estimated by the MAC rates of STAs that are already associated with AP i.
4. eTP_t: $t_{\text{alloc}}(i)$ is estimated by calculating the average waiting time between consecutive unique MSDUs from AP i to the joining STA.

We now describe each method, and provide detailed evaluations in Chapter 4.

3.2.1 $eTMR$

The simplest way to estimate t_{alloc} is to treat it as a constant. Without loss of generality, we can set $t_{\text{alloc}} = 1$. Substituting this into Equation (3.3), the resulting metric becomes:

$$eTP(i) = t_{\text{alloc}}(i) \times eTMR(i) = eTMR(i) \tag{3.23}$$

which is equivalent to $eTMR$. We denote this decision metric as $eTMR$ in subsequent sections.

3.2.2 eTP_n

In this case we denote t_{alloc} as t^n_{alloc}. Let $N_{\text{assoc}}(i)$ be the number of STAs that are already associated with AP i, which can be broadcasted by the AP along with the BI signals. Assuming that APs can allocate equal amount of channel time to each associated STA, t^n_{alloc} can be computed as:

$$t^n_{\text{alloc}}(i) = \frac{1}{N_{\text{assoc}}(i) + 1} \tag{3.24}$$

Substituting Equation (3.24) into Equation (3.3), we obtain:

$$eTP_n(i) = \frac{1}{N_{\text{assoc}}(i) + 1} \times eTMR(i) \tag{3.25}$$

3.2.3 eTP_r

In this case we denote t_{alloc} as t^r_{alloc}. Let $R(i)$ be the MAC rate from AP i to the joining STA, and $R_i(k)$ be the MAC rate from AP i to the STA k. Assuming that APs can transmit equal amount of data in bits to each associated STA, t^r_{alloc} can be estimated as:

$$t^r_{\text{alloc}}(i) = \frac{1/R(i)}{1/R(i) + \sum_{k\in\mathcal{S}_i} 1/R_i(k)} \tag{3.26}$$

where $\mathcal{S}_i$ is the set of STAs that are already associated with AP i. Intuitively speaking, the higher MAC rate one STA can get from AP i, the less time AP i spends to transmit packets to the STA, and hence the less channel time AP i would allocate to the STA. Specifically for the joining STA, the higher MAC rate other associated STAs can get from AP i, the more channel time AP i can allocate to the joining STA.

Substituting Equation (3.26) into Equation (3.3), we obtain:

$$eTP_r(i) = \frac{1/R(i)}{1/R(i) + \sum_{k\in\mathcal{S}_i} 1/R_i(k)} \times eTMR(i) \tag{3.27}$$

Note the quantity $\sum_{k\in\mathcal{S}_i} 1/R_i(k)$ can be computed by AP i and broadcast along with BI signals.

3.2.4 eTP_t

In this case, we denote eTP as eTP_t and t_{alloc} as t^t_{alloc}. In practice, $t_{\text{alloc}}(i)$ is not only determined by the number of associated STAs and the MAC rate of those STAs, but also determined by the success probability of packet transmission to each STA. For example, if one STA has excessively low packet success probability, the AP would have to keep retransmitting to it and hence spend less time on other STAs. $t_{\text{alloc}}(i)$ is also affected by how often AP i has to wait for the transmissions of other APs and STAs due to CSMA.

In order to estimate $t^t_{\text{alloc}}(i)$, we introduce the concept of the expected transmission time $t_u(i)$ to deliver one unique MSDU from AP i to the joining STA, and the expected waiting time $t^{\text{after}}_w(i)$ between two consecutive unique MSDUs for traffic from AP i to the joining STA if the joining STA selects AP i as its serving AP. By definition, $t^t_{\text{alloc}}(i)$ is the ratio between these two quantities:

$$t^t_{\text{alloc}}(i) = \frac{t_u(i)}{t^{\text{after}}_w(i)} \tag{3.28}$$

We now describe how to estimate $t_u(i)$ and $t^{\text{after}}_w(i)$.

$t_u(i)$ is defined to be the average time spent by AP i to deliver one unique MSDU to the joining STA, which includes backoff time, possible retransmissions, and other protocol related overheads. To estimate $t_u(i)$, we need to determine how many retries X_i are required to deliver one unique MSDU from AP i to the joining STA.

The distribution of X_i is the stationary distribution π of the MC model shown in Figure 3.1.

Let $T_u(i)$ be the random variable for the time spent by AP i to send one MSDU to the joining STA until the MSDU is successful after consecutive retries or until the MSDU is dropped due to exceeding retransmission limit. $T_u(i)$ can be written as:

$$T_u(i) = \sum_{x=0}^{X_i} t(i) \tag{3.29}$$

where $t(i)$ is the average time that AP i allocates to the joining STA for one PPDU transmission given by Equation (3.7). Then $t_u(i)$ is the expected value of $T_u(i)$:

$$t_u(i) = E[T_u(i)] = E[\sum_{x=0}^{X_i} t(i)] \tag{3.30}$$

Expanding Equation (3.30), we obtain:

$$t_u(i) = E[X_i]t(i) \tag{3.31}$$

where

$$E[X_i] = \sum_{n=0}^{N-1} (n+1)\pi(n) \tag{3.32}$$

Next we describe how to estimate $t_w^{\text{after}}(i)$. $t_w^{\text{after}}(i)$ depends on the amount of channel time AP i needs to allocate to the joining STA and other associated STAs, as well as the amount of time AP i has to wait when other nodes are active. This quantity cannot be measured before AP i starts transmitting to the joining STA. However, assuming only one new STA joins the network, we can estimate $t_w^{\text{after}}(i)$ as:

$$t_w^{\text{after}}(i) = \begin{cases} t_w^{\text{before}}(i) & \text{if } t_{\text{idle}}(i) > t_u(i) \\ t_w^{\text{before}}(i) + t_u(i) - t_{\text{idle}}(i) & \text{if } t_{\text{idle}}(i) < t_u(i) \end{cases} \tag{3.33}$$

where $t_u(i)$ is given by Equation (3.31), $t_w^{\text{before}}(i)$ is the minimum over all associated STAs' average waiting time to serve two consecutive unique MSDUs to the same STA from AP i before the joining STA selects any AP, and $t_{\text{idle}}(i)$ is the minimum over all associated STAs' average idle time at AP i for a duration of $t_w^{\text{before}}(i)$ before the joining STA selects any AP. Both $t_w^{\text{before}}(i)$ and $t_{\text{idle}}(i)$ can be computed at AP i, and be transmitted to the joining STA as additional fields in BI signals. The intuition in computing $t_w^{\text{after}}(i)$ is as follows: if AP i is not too busy and has enough idle time to serve packets to the joining STA without sacrificing the TXOPs of other already associated STAs, the average waiting times $t_w^{\text{before}}(i)$ and $t_w^{\text{after}}(i)$ should be the same before and after the joining STA enters the network; otherwise, all STAs served by AP i experience longer waiting times.

$t_w^{\text{before}}(i)$ depends on a number of factors such as the number of STAs associated with AP i, duration of interference that can pause backoff due to CSMA, as well as the

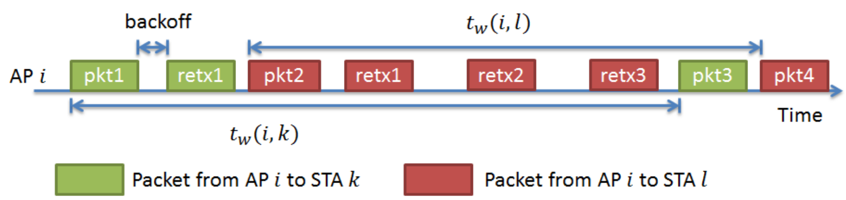


Figure 3.2. Illustration of $t_w(i,k)$.

MAC rate and the packet loss rate to each associated STA. Rather than estimating all these quantities, we simply require AP i to keep track of $t_w(i,k)$ which is the average time duration between two consecutive unique MSDUs transmitted by AP i to STA k. Then:

$$t_w^{\text{before}}(i) = \min_{k \in \mathcal{S}_i} t_w(i,k) \tag{3.34}$$

where $\mathcal{S}_i$ is the set of STAs associated with AP i before the joining STA selects any AP. Intuitively $t_w(i,k)$ can be described as follows: Assume upper layer applications are sending saturated traffic, all traffic streams have the same priority, and an AP does not switch to a new MSDU until the previous MSDU is successful or dropped due to its retransmission limit being exceeded. Then the average behavior of AP i can be modeled as transmitting MSDUs to each of its associated STA in a round-Robin fashion. The intuition for $t_w(i,k)$ is that on average AP i will be served one unique MSDU to STA k within a time period of $t_w(i,k)$, as shown in Figure 3.2.

$t_{\text{idle}}(i)$ can be computed at AP i as follows: Define the total idle time $t_{\text{idle}}^{\text{total}}(i)$ at AP i to be the period of time when AP i is ready to transmit but there is no packet in the queue during a certain observation period denoted by t_{count}. We opt to choose t_{count} to be the same amount of time to collect BI signals, which is $3sec$ in this work. Let $N_{\text{MSDU}}(i,k)$ be the number of MSDUs transmitted from AP i to STA j during t_{count}. Then $t_{\text{idle}}(i)$ is given by:

$$t_{\text{idle}}(i) = \frac{t_{\text{idle}}^{\text{total}}(i, t_{\text{count}})}{\max_{k \in \mathcal{S}_i} N_{\text{MSDU}}(i,k,t_{\text{count}})} \tag{3.35}$$

In estimating $t_{\text{idle}}(i)$ in the above equation, we assume t_{count} to be large enough for the estimate to be independent of t_{count}.

Once the joining STA receives estimated $t_w^{\text{after}}(i)$ and $t_{\text{idle}}(i)$ from AP i, given by Equations (3.33) and (3.35) respectively, $t_w^{\text{after}}(i)$ can be computed as in Equation (3.33). Combing with $t_u(i)$ as in Equation (3.31), we can compute the decision metric in this case as:

$$eTP_t(i) = \frac{t_u(i)}{t_w^{\text{after}}(i)} \times eTMR(i) \tag{3.36}$$

3.3 AP Selection Algorithm

We have described four AP decision metrics, namely $eTMR$, eTP_n, eTP_r, and eTP_t. In practice, one of the four metrics is used to estimate $eTP(i)$. The proposed AP selection algorithm is summarized in Algorithm 1.

Algorithm 1 AP Selection Algorithm

The joining STA scans for APs and obtains candidate AP set $\mathcal{A}$
Collect local BI for 3 seconds
Receive BI signal and other fields from all APs in $\mathcal{A}$
for each AP $i \in \mathcal{A}$ **do**
    compute $eTP(i)$
end for
The joining STA selects AP with largest eTP

Chapter 4

Performance Evaluation

4.1 Simulation Methodology

We use NS-2.31 to simulate 802.11b networks in infrastructure mode. This can be easily extended to other standards, such as 802.11a, 802.11g and 802.11n. The NS simulator has been modified to compute collision probability as described in [20]. The transmission range of nodes is about 32m, and no fading or shadowing is used in our path loss model. Each STA receives DL traffic from its serving AP. Each DL stream consists of traffic from a Constant Bit Rate (CBR) application which is generating packets at a rate that saturates the network. UDP is used as the transport layer protocol. MAC retry limit is set to be 10. Both the transmit and receive antennae have 0 dB gain.

In all simulations, both APs and STAs are randomly placed in a 110 m $\times$ 110 m region mimicking typical hyper dense scenarios observed in practice. The number of simulated APs and STAs are shown in Table 4.1, and selected example topologies are shown in Figure 4.1. To reduce simulation overhead, we pre-generate three different sets of random locations for a given number of APs with the following constraints: 1) the APs cover at least 95% of simulation area, 2) the APs have a minimum separation distance depending on number of APs as shown in Table 4.1. Each simulation trial places all APs according to one of the three sets of pre-specified random AP locations. The STAs are placed at random according to a spatial Poisson process.

To determine the ground truth, for each simulation trial we fix AP and STA locations and run the simulations to compute the throughput of the joining STA under

Table 4.1. Simulated Topology Parameters

# AP	8	8	16	16	16	24	24	32
# STA	20	40	20	40	60	40	60	40
Minimum AP distance [m]	30	30	20	20	20	10	10	10

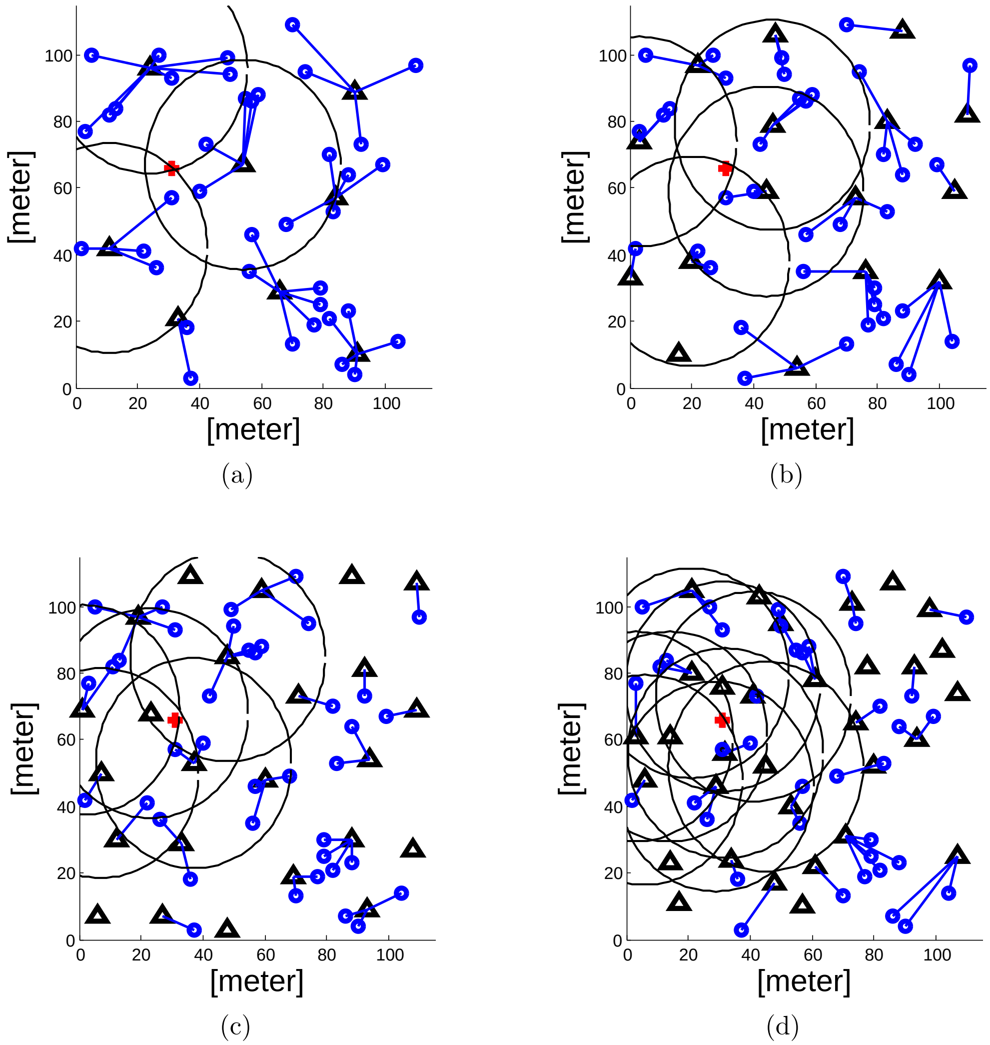


Figure 4.1. Example Topologies. Small blue circles: STAs. Big black circles: AP's range. Black triangles: APs. Red star: the joining STA. A line connecting an STA to an AP: the STA is associated with the AP. (a) 8 APs and 40 STAs; (b) 16 APs and 40 STAs; (c) 24 APs and 40 STAs; (d) 32 APs and 40 STAs.

the exact same conditions except that the joining STA associates with different APs, in order to determine the highest throughput AP which we call "optimum". Next we run the AP selection algorithms with proposed metrics and compare their selections with those obtained from the traditional strongest received power ($rxpwr$) algorithm in which the AP with the strongest received power is chosen. For each pair of number of APs and STAs shown in Table 4.1, we run 900 trials and discard the trials whose optimal throughput is less than 1kbps. We call these discarded trials "invalid". We refer to a non-optimal AP selection for a given algorithm as a valid trial in which the algorithm does not result in the same AP as the optimum.

Table 4.2. Simulation Results: 24 APs and 60 STAs

	% non-optimal selection	Average throughput [kbps]	Throughput gain vs. rxpwr
rxpwr	54%	551.58	0
eTMR	31%	721.27	31%
eTPn	28%	771.10	40%
eTPr	26%	782.61	42%
eTPt	21%	838.92	52%
optimal	0%	865.92	57%

Table 4.3. Reduction in Non-Optimal Selections By eTP_t Compared to *rxpwr*

Reduction in non-optimal selection		# AP			
		8	16	24	32
# STA	20	-57%	-54%	n/a	n/a
	40	-51%	-59%	-57%	-68%
	60	n/a	-52%	-62%	n/a

4.2 Comparing Algorithms

Using extensive simulations, we have found that all proposed metrics perform better than *rxpwr* for all topologies listed in Table 4.1, with eTP_t being the best. As an example, Table 4.2 summarizes the results for a topology with 24 APs and 60 STAs. As seen, our proposed algorithms all outperform *rxpwr* in terms of both the percentage of non-optimal selections and the average throughput of the joining STA; furthermore, eTP_t attains the best performance among all algorithms, achieving 97% of the average throughput obtained by optimal selection, with a 52% improvement in average throughput as compared to *rxpwr*. We focus on analyzing eTP_t in the remainder of the thesis.

Table 4.3 summarizes the reduction in non-optimal AP selections made by eTP_t as compared to *rxpwr* for all simulated topologies. As seen, eTP_t can reduce the non-optimal selections by more than 50% in all cases. Table 4.4 shows the average throughput gain of the joining STA obtained by eTP_t compared to *rxpwr*. In general, eTP_t achieves positive throughput gain over *rxpwr* in every simulated topology, with up to 72% improvement. Table 4.5 shows the percentage of achieved throughput by eTP_t compared to optimal, which is at least 94% in all topologies.

Table 4.4. Average Throughput Gain By eTP_t Compared to $rxpwr$

Throughput gain vs. rxpwr		# AP			
		8	16	24	32
# STA	20	4%	22%	n/a	n/a
	40	9%	29%	38%	72%
	60	n/a	42%	52%	n/a

Table 4.5. Percentage of Optimal Throughput Achieved By eTP_t

Percenage of optimal TP achieved		# AP			
		8	16	24	32
# STA	20	99%	96%	n/a	n/a
	40	99%	96%	96%	94%
	60	n/a	97%	97%	n/a

4.3 Increasing Node density

With the increasing popularity of WiFi enabled devices, for a fixed number of APs in a given region, the throughput of each STA drops as the STA density increases. To alleviate this problem, more APs are typically deployed in order to provide better throughput to each STA. This creates hyper dense WLANs where AP selection based on $rxpwr$ results in sub-optimal performance. Even for scenarios with a fixed number of users in a given area, more APs are typically added in order to achieve higher throughput per STA. An inherent problem with WLANs is that there is no closed-loop power control mechanism to adjust cell size of APs. Therefore, as the AP density increases, the joining STA can select from an increasing number of candidate APs, making the AP selection problem even more important.

Figure 4.2 shows the NS simulation results for the distribution of the number of candidate APs for the joining STA, in a number of scenarios as the number of APs in a 110 m $\times$ 110 m area increases from 8 to 32. As expected, the joining STA can select from more candidate APs as the number of APs increases. For topologies with 32 APs, in 96% of simulation trials there are four or more candidate APs to choose from, and as such, AP selection becomes a significant issue. On the other hand, for topologies with 8 APs, the joining STA has one or two candidate APs in more than 85% of simulated trials, and therefore the AP selection problem is not as severe, so we do not expect much throughput gain in this case.

To evaluate the throughput performance of eTP_t compared to $rxpwr$ and optimal as a function of node density, we plot the average throughput of the joining STA over all valid simulation trials obtained by different selection algorithms in Figure 4.3.

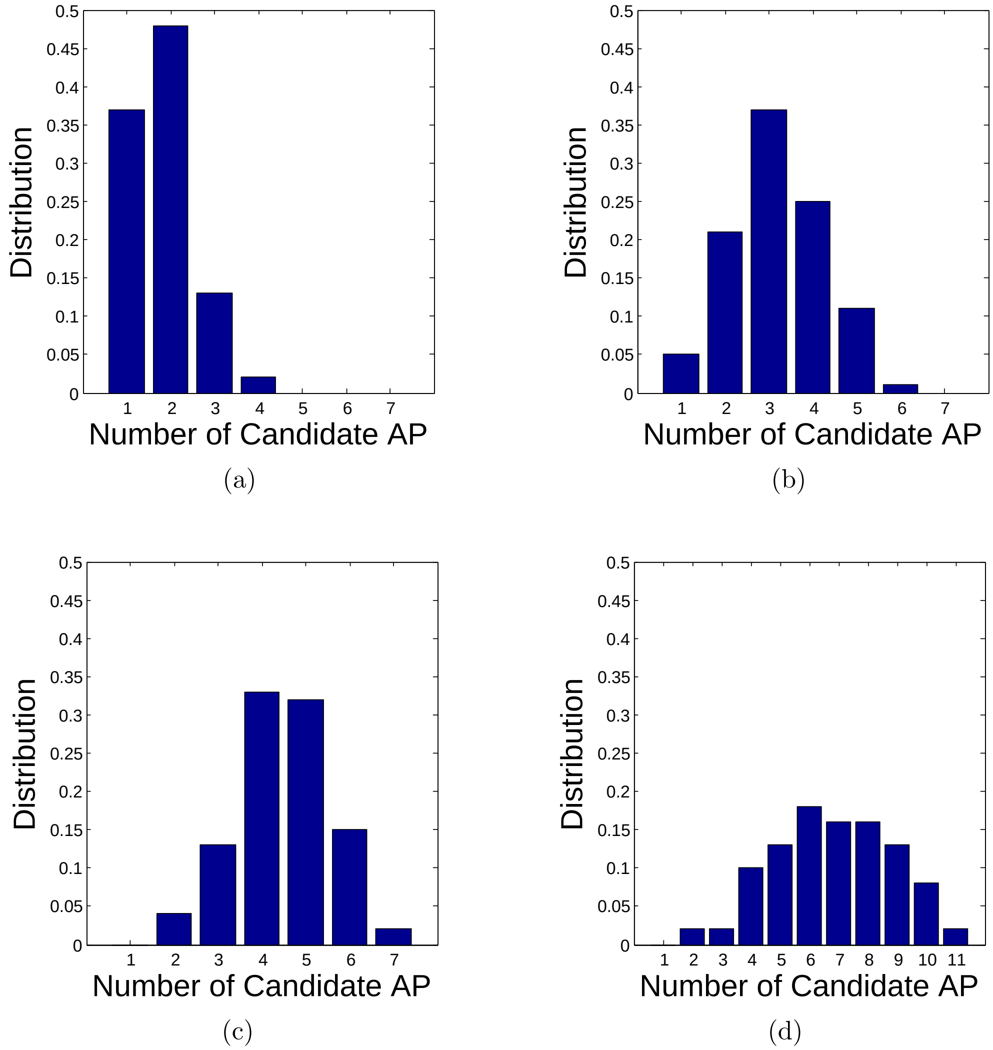
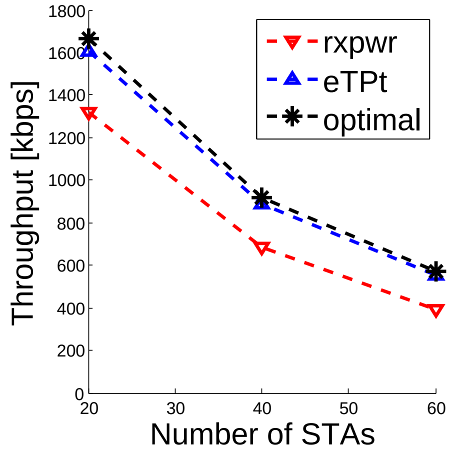
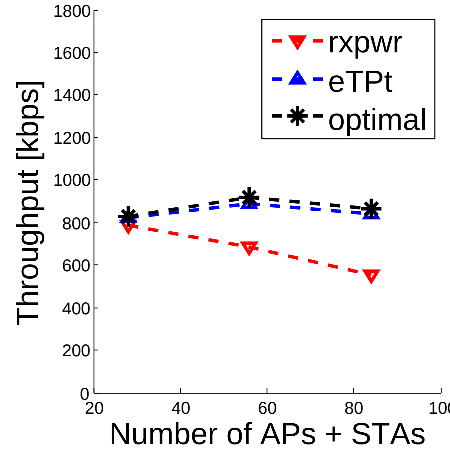


Figure 4.2. Distribution of the number of candidate APs for the joining STA: (a) 8 APs; (b) 16 APs; (c) 24 APs; (d) 32 APs.

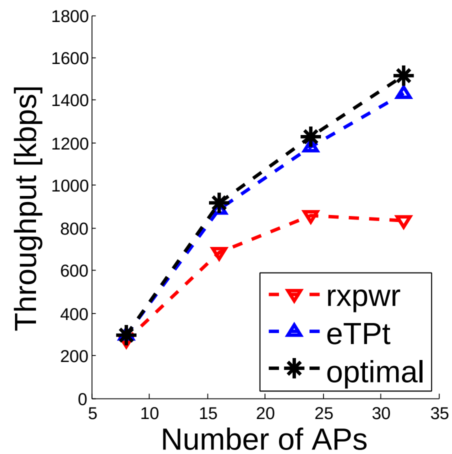
The results in Figure 4.3 are the same as those shown in Tables 4.3, 4.4, and 4.5. As expected, the average throughput for all AP selection algorithms drops with number of STAs, or equivalently the STA density, as shown in Figure 4.3(a). If more APs are placed in a WLAN to accommodate the increase in the number of STAs, the *rxpwr* throughput drops with total node density, while the eTP_t throughput stays almost invariant and is very close to the optimal throughput, as shown in Figure 4.3(b). Specifically in Figure 4.3(b), the throughput gain of eTP_t over *rxpwr* increases with total node density, growing from 4% for 28 nodes to 52% for 84 nodes. As seen in Figure 4.3(c), if the AP density is increased while STA density is fixed, the *rxpwr* throughput indicates a diminishing return, while the optimal and eTP_t throughput



(a)



(b)



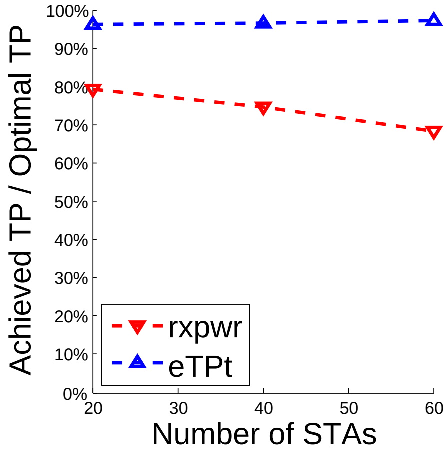
(c)

Figure 4.3. Average throughput of the joining STA achieved by different AP selection algorithms as a function of: (a) STA density with 16 APs; (b) total node density with STA to AP ratio fixed at 5:2; (c) AP density with 40 STAs.

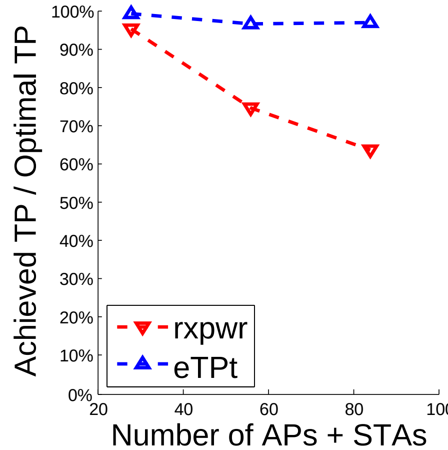
both increase almost linearly. [1] In particular, the throughput gain for eTP_t over $rxpwr$ increases with AP density, growing from 9% for 8 APs to 72% for 32 APs.

In order to examine the relative performance of eTP_t and $rxpwr$ as a function of node density, we plot the percentage of their achieved throughput as compared to

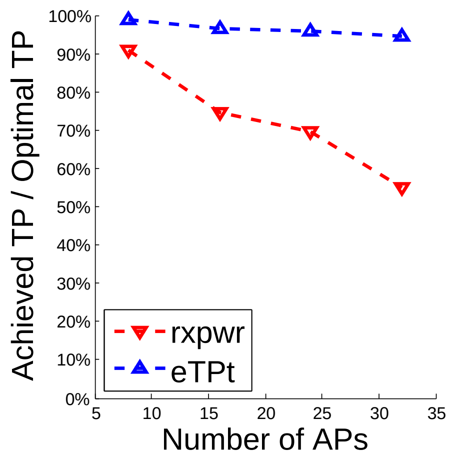
[1] Nonetheless, we speculate the optimal throughput for the joining STA to eventually reach diminishing return when the number of APs exceeds a certain threshold. This is because as the number of APs increases in a fixed area, the average distance between APs decreases, and the spatial frequency reuse of this WLAN decreases, which means more APs have to contend for the same amount of TXOPs. With fewer TXOPs for each AP, the throughput for each STA is likely to decrease. With proper AP selection, the threshold of diminishing return on average throughput for the joining STA is postponed to larger number of APs.



(a)



(b)



(c)

Figure 4.4. Percentage of throughput achieved by different AP selection metrics for the joining STA compared to optimal as a function of: (a) STA density with 16 APs; (b) total node density with STA to AP ratio fixed at 5:2; (c) AP density with 40 STAs.

optimal in Figure 4.4. Specifically, Figures 4.4(a), 4.4(b) and 4.4(c) plot this quantity as a function of STA density, node density and AP density, respectively. The number of APs increases in the latter two cases, with the difference being that in Figure 4.4(b) the STA to AP ratio is maintained as 5:2, while in Figure 4.4(c) the number of STAs is fixed at 40. As seen, eTP_t remains within 6% of the optimal throughput in all three scenarios. In contrast, the relative throughput of $rxpwr$ with respect to optimal decreases with STA density. This is mainly because $rxpwr$ does not take into account collisions and TXOPs, which are important factors in determining throughput as the AP density increases. Overall, the throughput gain achieved by eTP_t over $rxpwr$ increases with node density, be it AP density, STA density, or both.

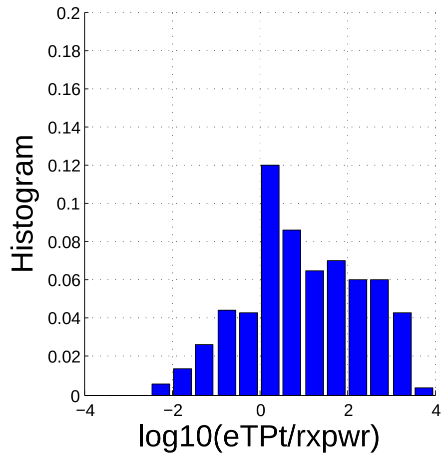


Figure 4.5. Histogram of the logarithm of ratio of eTP_t over $rxpwr$ throughput when decisions are different for 24 APs and 60 STAs.

So far, we have examined average throughput over all valid simulation trials, which is not a representative of the distribution of the throughput among trials. To examine the distribution of throughput gain, in Figure 4.5 we plot the histogram of the logarithm of the ratio of eTP_t over $rxpwr$ throughput of the joining STA for trials in which their selections are different, for the case with 24 APs and 60 STAs. In this case, eTP_t achieves higher (lower) throughput compared to $rxpwr$ in 50% (14%) of the trials. In the remaining 36% of trials they select the same AP and hence have identical throughput. The histograms for other topologies show similar trends. In general for all simulated topologies, the number of trials in which eTP_t achieves throughput gain over $rxpwr$ is larger than those with throughput loss.

Chapter 5

Conclusion

In this paper we proposed a class of AP selection algorithms which take into account TXOPs, inter-BSS interference and collisions. This is achieved by exploiting BI signals both at the AP and at the joining STA, as well as additional information such as the number of associated STAs, sum of inverse of MAC rates to associated STAs, and average waiting and idle times.

Our proposed AP selection algorithms can reduce the percentage of non-optimal selections and improve the average throughput of the joining STA in all tested scenarios. In particular, eTP_t achieves the best performance among all AP selection algorithms, and performs better than $rxpwr$ as the node density increases, be it AP density, STA density, or both. For a random topology with 24 APs and 60 STAs, the average throughput gain of the joining STA obtained by eTP_t is 52% as compared to the traditional received signal power based method. In this scenario eTP_t achieves as much as 97% of the optimal throughput.

Future work includes extending the current work to AP selection for UL traffic, examining the impact of our AP selection algorithm on aggregate network throughput, and extending current static algorithm to dynamic AP selection in which existing STAs can switch from one AP to another.

References

[1] A. J. Nicholson, Y. Chawathe, M. Y. Chen, B. D. Noble, and D. Wetherall, “Improved Access Point Selection,” in *International Conference on Mobile Systems, Applications and Services*, 2006.

[2] S. Vasudevan, K. Papagiannaki, C. Diot, J. Kurose, and D. Towsley, “Facilitating Access Point Selection in IEEE 802.11 Wireless Networks,” in *ACM SIGCOMM Conference on Internet Measurement*, 2005.

[3] K. Sundaresan and K. Papagiannaki, “The Need for Cross-layer Information in Access Point Selection Algorithms,” in *ACM SIGCOMM Conference on Internet Measurement*, 2006.

[4] M. Abusubaih, J. Gross, S. Wiethoelter, and A. Wolisz, “On Access Point Selection in IEEE 802.11 Wireless Local Area Networks,” in *IEEE Conference on Local Computer Networks*, Nov 2006.

[5] S. Miyata, T. Murase, and K. Yamaoka, “Characteristic analysis of an access-point selection for user throughput and system optimization based on user cooperative moving,” in *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*, Jan 2012.

[6] L. Luo, D. Raychaudhuri, H. Liu, M. Wu, and D. Li, “Improving End-to-End Performance of Wireless Mesh Networks through Smart Association,” in *IEEE WCNC 2008*, Mar 2008.

[7] Y. Fukuda, M. Honjo, and Y. Oie, “Development of Access Point Selection Architecture with Avoiding Interference for WLANs,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Mar 2006.

[8] L. Du, Y. Bai, and L. Chen, “Access Point Selection Strategy for Large-scale Wireless Local Area Networks,” in *IEEE WCNC*, Mar 2007.

[9] M. Abusubaih and A. Wolisz, “Interference-Aware Decentralized Access Point Selection Policy for Multi-Rate IEEE 802.11,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep 2008.

[10] J. W. Jang, Y. S. Lim, and C. K. Kim, “Traffic-aware decentralized AP selection for multi-rate in WLANs,” in *International Conference on Advanced Communication Technology*, Feb 2010.

[11] Y. Bejerano, S. jae Han, and L. Li, “Fairness and Load Balancing in Wireless LANs Using Association Control,” *Networking, IEEE/ACM Transactions on*, vol. 15, no. 3, pp. 560–573, June 2007.

[12] H. Gong, K. Nahm, and J. Kim, “Distributed Fair Access Point Selection for Multi-Rate IEEE 802.11 WLANs,” in *Consumer Communications and Networking Conference, 2008. CCNC 2008. 5th IEEE*, Jan 2008.

[13] W. Zhou and D. Qiao, “Fulfillment-Based Fairness: A New Fairness Notion for Multi-AP Wireless Hotspots,” in *IEEE International Conference on Communications*, June 2007.

[14] G. Judd and P. Steenkiste, “Fixing 802. 11 access point selection,” *ACM SIGCOMM Computer Communication Review*, 2002.

[15] J. C. Chen, T. C. Chen, T. Zhang, and E. van den Berg, “Effective AP Selection and Load Balancing in IEEE 802.11 Wireless LANs,” in *IEEE GLOBECOM*, Dec 2006.

[16] P. Bahl, M. Hajiaghayi, K. Jain, S. Mirrokni, L. Qiu, and A. Saberi, “Cell Breathing in Wireless LANs: Algorithms and Evaluation,” *Mobile Computing, IEEE Transactions on*, Feb 2007.

[17] J. Musacchio and J. Walrand, “WiFi Access Point Pricing As a Dynamic Game,” *IEEE/ACM Trans. Netw.*, vol. 14, no. 2, Apr. 2006.

[18] K. Mittal, E. Belding, and S. Suri, “A game-theoretic analysis of wireless access point selection by mobile users,” *Computer Communications*, 2008.

[19] L. Jiang, S. Parekh, and J. Walrand, “Base Station Association Game in Multi-Cell Wireless Networks (Special Paper),” in *IEEE WCNC 2008*, Mar 2008.

[20] M. N. Krishnan, S. Pollin, and A. Zakhor, "Local Estimation of Probabilities of Direct and Staggered Collisions in 802.11 WLANs," in *IEEE GLOBECOM*, Dec 2009.

[21] M. N. Krishnan and A. Zakhor, "Throughput Improvement in 802.11 WLANs Using Collision Probability Estimates in Link Adaptation," in *IEEE WCNC*, Apr 2010.

Appendix A

Acronyms and Symbols

Table A.1. Acronyms

Acronym	Meaning
AP	access point in WLANs
BEB	binary exponential backoff
BER	bit error rate
BI	busy-idle signal
BSS	basic service set
CBR	constant bit rate
CW	contention window
DC	direct collision
DL	downlink
ESS	extended service set
eTMR	expected true MAC rate
eTP	expected throughput
MAC	medium access control
MPDU	MAC layer Protocol Data Unit
MSDU	MAC layer Service Data Unit
MC	Markov Chain
PHY	physical layer
PPDU	Physical layer Protocol Data Unit
SC1	staggered collision of type 1
SC2	staggered collision of type 2
SNR	signal-to-noise ratio
STA	non-AP station in WLANs
Rx	receive or receiver
rxpwr	strongest received power based AP selection algorithm
TP	throughput
Tx	transmit or transmitter
TXOP	transmit opportunities
UL	uplink
WLAN	wireless local area network

Table A.2. Symbols

Symbol	Meaning
$\mathbb{1}\{\cdot\}$	indicator function
$\mathcal{A}$	candidate AP set
$\mathcal{S}_i$	Set of STAs associated with AP i before the joining STA selects any AP
π	stationary distribution of Markov Chain for backoff modeling
i	index for AP
j, k	index for STA. j is usually used for the joining STA
m	mth PPDU transmission
n	nth MSDU retry state
$A_i(m)$	the event that the mth PPDU sent by AP i to STA j is successful
$CW(n)$	contention window size on retransmission state n
$CW_{\max}$	maximum CW size defined in 802.11
$CW_{\min}$	minimum CW size defined in 802.11
$eTMR(i)$	expected true MAC rate from AP i to the joining STA
$eTP(i)$	expected throughput from AP i to the joining STA
$eTP_n(i)$	expected throughput from AP i to the joining STA estimated using $t^n_{\text{alloc}}(i)$
$eTP_r(i)$	expected throughput from AP i to the joining STA estimated using $t^r_{\text{alloc}}(i)$
$eTP_t(i)$	expected throughput from AP i to the joining STA estimated using $t^t_{\text{alloc}}(i)$
$L_i(m)$	MAC payload (MSDU) size in bits on mth transmission from AP i to STA j
L	maximum MAC payload (MSDU) size
L_{MAC}	MAC frame size in bits
L_{PHY}	preamble and PLCP header size in bits
N	packet retry limit
$N_{\text{assoc}}(i)$	The number of STAs that are already associated with AP i
$P(i)$	transition matrix for modeling BEB from AP i to the joining STA
P_C	total collision probability
P_e	channel error probability
P_L	total packet loss probability
P_S	packet success probability
R_{MAC}	MAC modulation rate
R_{PHY}	PHY modulation rate
$R(i)$	The MAC rate from AP i to the joining STA
$R_i(k)$	The MAC rate from AP i to STA k
T_{ACK}	time to transmit ACK packets
T_{DIFS}	DIFS duration defined in IEEE 802.11
T_{EIFS}	EIFS duration defined in IEEE 802.11
T_{SIFS}	SIFS duration defined in IEEE 802.11
T_{slot}	slot time duration defined in IEEE 802.11

Table A.3. Symbols Continued

Symbol	Meaning
$t(i)$	average time that AP i allocated to STA j for one PPDU transmission
$t_{\text{alloc}}(i)$	The potential percentage of channel time that AP i can allocate to the joining STA
$t^n_{\text{alloc}}(i)$	$t_{\text{alloc}}(i)$ estimated by using number of associated STAs to AP i
$t^r_{\text{alloc}}(i)$	$t_{\text{alloc}}(i)$ estimated by using MAC rates of associated STAs to AP i
$t^t_{\text{alloc}}(i)$	$t_{\text{alloc}}(i)$ estimated by calculating the average waiting time between consecutive unique MSDUs from AP i to the joining STA
$\bar{t_{\text{b}}}(i,n)$	average backoff duration on nth retry state for PPDUs from AP i to STA j
$\bar{t_{\text{b}}}(i)$	average backoff duration for PPDUs from AP i to STA j
$t_h(i)$	time to transmit PHY layer pramble, PLCP header, MAC header and MAC CRC from AP i
$t_i(m)$	the time that AP i takes to transmit mth PPDU to STA j
$t_p(i)$	average MAC payload (MSDU) tx time from AP i to STA j
t_{protocol}	the overhead introduced by the 802.11 protocol for a single PPDU transmission, including inter-frame spacing, backoff time, and possible ACK time
$t_{OH}(i)$	average overhead time associated with one MSDU from AP i to STA j
$t_{OH}(i,n)$	average overhead time from AP i to STA j, when the MSDU is in nth retry state.
$t_h(i)$	time to transmit the PHY layer preamble, the PLCP header, the MAC header, and MAC CRC, from AP i to STA j
$t_w^{\text{after}}(i)$	The average waiting time between consecutive unique MSDUs for traffic from AP i to the joining STA
$t_w(i,k)$	The average waiting time between consecutive unique MSDUs for traffic from AP i to STA k
t_{count}	Time duration to collect BI signals.
$t_w^{\text{before}}(i)$	The minimum average waiting time between consecutive unique MSDUs for traffic from AP i to STAs that are already associated with AP i, before the joining STA selects any AP
$t_{\text{text}}^{\text{total}}(i)$	total idle time at AP i for a duration of t_{count}
$t_{\text{text}}(i)$	minimum average idle time at AP i for a duration of $t_w^{\text{before}}(i)$
$t_u(i)$	average time spend by AP i to send one MSDU to the joining STA, until the MSDU is successful or dropped due to exceeding retransmission limit
$T_u(i)$	random variable for the time spent by AP i to send one MSDU to the joining STA, until the MSDU is successful or dropped due to exceeding retransmission limit
X_i	random variable for the number of retransmissions required to send one unique MSDU from AP i to the joining STA