Traffic-aware Decentralized AP Selection for Multi-Rate in WLANs

Jae-wook Jang*, Yeon-sup Lim**, Chong-kwon Kim*

*School of Computer Science and Engineering, Seoul National University, Korea
**Department of Computer Science University of Massachusetts Amherst

Abstract— AP selection problem is one of the major issues in infrastructure WLANs. Recently, many authors in the literatures have proposed a novel AP selection scheme which can provide better performance (e.g. aggregated throughput, load balancing) than RSS-based legacy scheme. However they have presented the schemes with non-practical assumptions, e.g. they have assumed that adjacent APs are configured with orthogonal channels and each node transmits the data frame using a single data rate. As we have studied, adjacent BSSs’ transmission and rate-anomaly impact the network’s performance. In this paper, we propose a practical traffic-aware AP selection considering the factors previously mentioned. By exploiting the Retry field in the MAC header as feedback about channel conditions, we can infer the network conditions, and each client can select the ‘best’ AP in terms of expected throughput. We demonstrate the effectiveness of our solution by comparing with existing approaches through ns-2 simulation.

Keywords— IEEE 802.11 WLANs; AP selection; RetryRatio; Multiple transmission rate; adjacent BSSs’ transmission;

I. INTRODUCTION

Recently, as the deployment of IEEE 802.11 WLANs is going rapidly, we can use wireless equipments ubiquitously. They provide mobility, convenience, flexibility, and tolerable throughput compared to wired equipments.

In infrastructure WLANs, a STA associates with a single AP that coordinates all traffic in a downlink or uplink manner. Then AP selection is an important issue, which prominently impacts the system performance including throughput, fairness, QoS (Quality of Service). However, how to select an appropriate AP is not specified in current IEEE 802.11 standard, so manufacture vendors usually adapt AP selection scheme based on the RSSI (Received Signal Strength Indication) [1]. After scanning the channels, a STA selects the AP from which it receives frames with the strongest signal strength. As we have studied, such a RSS-based AP selection scheme cannot support the best throughput performance as shown in [4]-[14]. This legacy scheme leads to imbalance of system performance. The legacy scheme results in concentration on specific APs, i.e. many STAs may associate with a few APs, because they have received the frames with the strongest signal strength from those APs. But many works [4]-[11] have proposed AP selection schemes with non-practical assumptions: authors have assumed that each STA uses same transmission rate and adjacent APs are configured with different non-overlapping channels.

Instability of medium due to fading, multi-path and the limited number of orthogonal channels is inherent to WLANs. To cope with the variation of wireless channels and achieve higher spectral efficiency, the current 802.11 PHY provides users with data transmission rates of up to 54Mbps in IEEE 802.11a/g and 11Mbps in IEEE 802.11b [1]. Our interest in this work is not Rate-Adaptation, but the impact of Rate-Adaptation on the overall performance of multi-rate WLANs. As pointed out in [3], when some STAs use a lower data transmission rate than that of the others, the performance of all STAs is considerably degraded. When a slower STA captures the channel for a long time, it penalizes faster STAs; the faster STAs’ throughput is down-equaled to the slowest STA’s throughput. This phenomenon is defined as rate-anomaly. Moreover, the limited number of orthogonal channels (e.g. 3 channels in the 802.11 b/g, 12 channels in the 802.11a) results in Co-Channel Interference among the adjacent BSSs (Basic Service Set). These adjacent BSSs’ transmission that is called inter-BSS interference can result in performance degradation.

Nevertheless, most previous works [4], [7], [8], [10], [11] did not consider multiple transmission rates and adjacent BSSs’ transmission. Only a few papers [12], [13] have considered multiple transmission rates, but still without adjacent BSSs’ transmission. Authors of [14] have considered both of them, but did not reflect these factors perfectly.

In this paper, we present a more realistic approach to AP selection, i.e. we propose a new traffic-aware AP selection scheme considering multiple transmission rates as well as adjacent BSSs’ transmission. By exploiting the Retry field in the MAC header, we can estimate the number of active STAs in each BSS implicitly as shown in [15], and we derive the expected throughput at client side, i.e. the maximum achievable throughput when associating with a target AP. We demonstrate the effectiveness of our solution by comparing with existing approaches through ns-2 simulation [17].
The rest of the paper is organized as follows. Section 2 reviews the related work. Formulation of the proposed algorithm is presented in Section 3. Section 4 shows the performance evaluation through ns-2 simulation, and finally the paper concludes with Section 5.

II. RELATED WORK

According to IEEE 802.11 standard, AP association procedure is composed of passive scanning, active scanning, and association phase [1], [4]. The purpose of scanning is to find the appropriate AP which can provide the best performance among candidate APs. There are two types of scanning in the infrastructure WLANs: passive scanning, active scanning.

In passive mode, an AP periodically broadcasts the beacon frames including BSS-specific information such as Timestamp and SSID. A STA listens to the beacon frames on each candidate channel and selects a proper AP according to signal strength. In active scanning, each STA broadcasts the probe request frames, while the AP receiving these frames immediately acknowledges the probe request frame. After receiving multiple probe response frames, each STA chooses a proper AP according to signal strength. After any type of scanning, a STA attempts to get associated with the target AP by exchanging in series authentication and association request/response frames.

While 802.11 DCF ( Distributed Coordination Function) gives same access opportunities on long term period, the legacy scheme does not achieve good performance efficiency from client/network standpoint, since higher RSSI does not mean higher throughput at all. RSSI depends on the distance between a client and its AP as well as on transmission power of each AP, so RSSI does not provide any information about current each AP’s load. Moreover, throughput depends on the number of clients associated with each AP [8], since the wireless channel is a shared medium.

Many works [4]-[14] have shown that the legacy AP selection scheme leads to poor performance in terms of achievable throughput and load balancing; they have proposed a novel scheme considering traffic-aware in the BSS. There are two kinds of approaches about AP selection: centralized and decentralized approaches. In the former, wired equipment such as an AP or an intelligent management system connected to the WLANs controls communication between APs and STAs, and collects information such as the number of STAs, channel conditions. Such centralized architectures have been proposed in [5], [6]. However, when the wired equipment is broken down, the system cannot provide service at all. Moreover, the link between APs and the wired equipment might be bottle-neck. On the other hand, all STAs using decentralized approach [4], [7]-[14] select the AP based on various information piggybacked in the management frame or in self-recognized manner, instead of centralized help [5],[6].

In [7], two new dynamic association schemes have been proposed. The first scheme considered the channel conditions in both uplink and downlink to each AP as well as load at each AP. The second scheme combined this information with the routing information of packets from a candidate AP to the destination. Although the scheme did exploit an airtime cost metric, these schemes did not consider the multiple transmission rates and the adjacent BSSs’ transmission.

Authors of [8] have proposed an AP selection metric based on both the number of STAs and wireless channel conditions rather than RSSI. However, they have assumed same data transmission rate for all STAs: they did not consider multiple transmission rates. Authors of [9] have proposed AP selection scheme considering hidden terminal effect, by exploiting the QBSS Load information in IEEE 802.11e standard. But authors did not consider the adjacent BSSs’ transmission. Reference [10] described the methodology by estimating probe delay time in active scanning. Reference [11] described the methodology by estimating the potential bandwidth based on the delays experienced by beacon frames. However, AP selection schemes in [10], [11] have considered neither multiple transmission rates nor adjacent BSSs’ transmission. In [12], authors have proposed an AP selection scheme which considered the theoretical throughput as well as its impact on already associated STAs. But authors of [12] have assumed single-hop environment adopted by Bianchi’s model, so they did not consider the adjacent BSSs’ transmission. In [13], authors have proposed an AP selection only considering the multiple transmission rates. Reference [14] has proposed the metric of “expected throughput” that combined AP capacity in the presence of interference, the aggregated transmission delay of all existing clients, and transmission rate of a new client. They considered the two factors previously mentioned. However, they assumed the traffic was fully saturated downlink; in case of uplink traffic, their proposed metric is not working properly, since the measurement value of Aggregated Transmission Delay (ATD) has large variations. The ATD is directly related to the number of active STAs.

III. PROPOSED AP SELECTION SCHEME

A. System model

We consider large-scale WLANs densely deployed which consists of many APs and clients. All APs operate on infrastructure mode and are connected to wired networks.

For the medium access, we consider only DCF ( Distributed Coordination Function); all STAs access the channel based on CSMA/CA (Carrier Sense Multiple Access / Collision Avoidance) protocol. All transmissions in each BSS are made by an AP, i.e. either downlink or uplink. The traffic is generated only from STAs to its AP. Transmission rate of client is decided only by the distance between a client and its AP.

We assume perfect channel conditions, i.e. no packet loss, and also assume that no STA resorts to RTS/CTS mechanism due to high overhead.

B. Metric Formulation

The expected throughput of each client is defined as the inverse of average transmission time including all retransmissions until one is successful. We define the concept
of successful transmission time along with the frame exchange sequence of DCF mode in 802.11 WLANs [2], [4].

When a data frame arrives at the head of the queue but the channel is busy, the MAC waits until the medium becomes idle. If the channel becomes idle during the DIFS duration, the MAC starts the backoff mechanism. As long as the medium stays idle, a random backoff counter is decremented. When the backoff counter reaches zero, sender tries to transmit the data frame. For each successful reception of a data frame, the receiver immediately acknowledges the data frame reception by sending an ACK frame; the ACK frame is transmitted after SIFS duration. If the ACK frame is successfully transmitted, the procedure of data transmission is over [1], [2], [15].

Keeping the concept mentioned above in mind, we define the expected throughput as follows. The average transmission time until a client k transmits a data frame successfully is referred to as \( T_{k,i} \); the expected throughput is then given by:

\[
\text{expected throughput}(\theta_{k,i}) = \frac{l}{T_{k,i}} \quad (1)
\]

The average transmission time until a data frame is transmitted successfully can be derived easily by exploiting analysis of [12]. Above (1) does not consider retransmissions due to collisions between STAs. \( T_{k,i} \) denotes the average transmission time until client k transmits a data frame of length L to AP i successfully, after \( j \)th retransmission due to collision, and is given by:

\[
T_{k,j}(i) = tDIFS + tbackoff(j) + tDATA + tSIFS + tACK
\]

\[
tDATA = tPreamble + tHeader + \frac{l_{macheader} + l_{payload}}{rate(k)}
\]

\[
tACK = tPreamble + tHeader + \frac{l_{ack}}{rate(k)} \quad (2)
\]

where \( rate(k) \) is the data transmission rate of client k; \( tDATA \) depends on data transmission rate of client k. \( l_{macheader}, l_{payload} \) and \( l_{ack} \) denote length of Mac-header, payload, ACK frame respectively. All the control frames such as ACK and RTS and CTS are transmitted at the basic rate according to 802.11 standard; in case of 802.11b, the basic rate \( rate_{basic} \) is 1Mbps. \( tACK \) is the duration of the ACK frame, \( tDIFS \) is the DCF InterFrame Space and \( tSIFS \) is the Short Inter Frame Space.

The contention window takes an initial value of \( CW_{\min} \). If a data transmission attempt fails, the value of \( CW \) is doubled until it reaches \( CW_{\max} \). Once it reaches \( CW_{\max} \), the value of \( CW \) remains \( CW_{\max} \) until the transmission successfully goes to an end or the frame is discarded due to Retry limit, and \( CW \) is reset. This improves the stability of the access protocol and the performance under congestion conditions. The backoff interval randomly draws an integer number from a uniform distribution over the interval \([0, CW]\) [1], [2], [4].

\( t_{\text{backoff}}(j) \) denotes the average backoff time during consecutive \( j \)th transmission attempts as follows:

\[
t_{\text{backoff}}(j) = \frac{2^j \cdot (CW_{\min} + 1) - 1}{2} \cdot E[\text{slot time}] \quad (0 \leq j < 6)
\]

\[
= \frac{CW_{\max} \cdot tslot}{2} \quad (j \geq 6)
\]

When \( t_{\text{backoff}}(j) \) is calculated, we should consider the following: the IEEE 802.11 standard [1], [2] depicts, in its section 9.2.5.2, how the backoff counter is decremented. If the medium is busy at any time during a backoff slot, then the backoff procedure is suspended. We assume that a STA has a backoff counter equal to a value \( b \). If the current medium slot-time is idle, at the end of slot-time the backoff counter is decremented, and the STA will start the next slot-time with backoff counter \( b-1 \). On the other hand, if the current medium slot-time is busy, the STA will freeze the backoff counter at \( b \); the backoff counter is decremented only during idle slots. When we calculate \( t_{\text{backoff}}(j) \), slot time does not denote physical slot time, but virtual slot time.

\( E[\text{slot time}] \) is the expected length of a slot-time, it is adopted from virtual slot time of Bianchi’s model [16] to be given by:

\[
E[\text{slot time}] = P_{\text{idle}} \cdot \sigma + P_{\text{tr}} \cdot T_c + P_{\text{tr}} \cdot (1 - P_{\text{tr}}) \cdot T_e
\]

\[
P_{\text{idle}} = \frac{\sum_{i=1}^{m} \sigma_i}{T} = 1 - P_{\text{tr}}
\]

\[
T_c = tDATA + tEIFS
\]

\[
T_e = tDATA + tSIFS + tACK + tDIFS \quad (4)
\]

where \( T \) denotes the time measured to estimate the expected length of a slot-time. \( P_{\text{idle}} \) is the idle probability defined as idle ratio during time interval \( T \). \( P_{\text{tr}} \) is the probability that there is at least one transmission during the considered slot time. \( P_{\text{c}} \) denotes conditional collision probability, meaning that the probability of a collision seen by a packet being transmitted on the channel. \( P_{\text{c}} \) has the same meaning as \( p \) that will be mentioned next. \( 1 - P_{\text{tr}} \) is the probability that transmission occurring on the channel is successful, and is given by the probability that exactly one STA transmits on the channel, conditioned on the fact that at least one STA transmits. \( \sigma \) is the duration of an empty slot-time, \( T_e \) is the average time the channel is sensed busy because of a successful transmission, and \( T_c \) is the average time the channel is sensed busy by each station during a collision. \( T_e \) and \( T_c \) are determined by the slowest STA already associated with AP i. Accordingly, estimating \( P_{\text{idle}} \) plays a key role in knowing \( E[\text{slot time}] \). We assume the channel state seen by each STA in BSS is same, and then \( P_{\text{idle}} \) of all STAs in BSS is same, so each AP can estimate \( P_{\text{idle}} \) of its own BSS by measuring the channel status.

Therefore, the average time required for a client k to transmit a data frame successfully is given by:

\[
\text{expected throughput}(\theta_{k,i}) = \frac{l}{T_{k,i}} \quad (1)
\]
\[
T_{k,i} = (1 - p) \cdot T_{k,i}(0) + \sum_{j=1}^{\gamma} (1 - p) \cdot p^j \cdot \sum_{m=0}^{j-1} T_{\text{fail}}(m)
\]

where \( \gamma \) is the Retry limit. In case of two-hand shaking, RetryLimit (LongRetryLimit) is 4 as shown in [1]; \( p \) is the average up-link transmission failure probability of each BSS, the derivation of \( p \) is shown in next sub-section. \( T_{\text{fail}}(m) \) is the time duration corresponding to \( m \) transmission attempt’s failure and is given by:

\[
T_{\text{fail}}(m) = \text{backoff}(m) + t\text{DATA} + t\text{EIFS}
\]

But above (5) does not consider rate-anomaly: if a newly arrived client selects an AP dealing with STAs which use a lower data rate than the newly arrived client’s data rate, the expected throughput of the newly arrived client is down-equalized to that of the slowest STAs, and not to the value derived from (5). We exploit AP’s goodput during a constant interval (5 sec) to consider rate-anomaly. Let \( S \) be the goodput during the constant interval. On long term, the throughput per client in BSS, is given by:

\[
S = \frac{S_i}{N_i} \approx \frac{[\text{throughput per client in BSS}]}{}
\]

where \( S_i \) denotes AP_i’s goodput and \( N_i \) denotes the number of active STAs associated with AP_i.

Therefore, the expected throughput of the client is given by:

\[
\text{expected throughput}(\theta_{k,i}) = \min(S, \frac{1}{T_{k,i}})
\]

C. 802.11 Feedback to infer average transmission failure probability

We exploit a novel technique to estimate the number of active STAs implicitly, by measuring frequency of retransmission as shown in [15]. Fig. 1 shows the format of a general IEEE 802.11 MAC layer frame. Retry field in 802.11 MAC header is a single bit in length and is used to indicate whether a data or a management frame is being transmitted for the first time or it is a retransmission [1].

When this field is set to 0, the frame is being sent for the first time, when this field is set to 1, the frame is a retransmission of an earlier unsuccessful transmission. A receiver uses this indication to aid in the process of eliminating duplicate frames.

The Retry field can be used as a channel feedback to infer the channel conditions, because there is a correlation between \( p \) and pattern of Retry values of arriving frames as shown in [15]. As the channel gets more congested, the number of retransmission increases [1]. By exploiting this indication, a receiver can estimate \( p \) in each BSS, i.e., during measure time \([0, T] \), if the number of Retry field which is set to "1" is increasing, it infers that the transmission attempts more often fail due to collision within BSS as well as with adjacent BSS’s transmissions.

In order to model and analyse the pattern of Retry field, we reuse Bianchi’s Markov chain model [15], [16]. More details can be found in [15]. As explained briefly, Fig. 2 shows a discrete-time Markov-chain model describing the back-off window scheme of 802.11 DCF. \( b(t) \) denotes the stochastic process representing the backoff window size for a given station at time \( t \), and \( s(t) \) denotes the stochastic process representing the backoff stage for a given station at time \( t \), where \( m \) represents the maximum backoff stage. The two-dimensional process \( s(t), b(t) \) is represented by state \( s(t) = i, b(t) = k \) at time \( t \), as \( t \to \infty \), the stationary distribution of the chain is given by:

\[
b_{i,k} = \lim_{t \to \infty} P(s(t) = i, b(t) = k), \quad i \in (0,m), k \in (0,W_i - 1)
\]
\[
\frac{C_i}{C_0 + C_i} = \frac{(1-p)b_{0,0}}{(1-p)b_{0,0} + (1-p)\sum_{k=0}^{m} p^k} = \frac{1 - p}{1 - p^{m+1}}
\]

where the number of \textit{Retry} field which is set is referred to as \(i\) \((i=0, 1)\), and the \(C_i/C_0\) denotes the \textit{RetryRatio}; \(m\) is the maximum number of backoff stage; the value of \(m\) is 4 in case of two hand-shaking in WLANs as shown in [1].

\section{Implementation Issues}

In this sub-section, we will discuss implementation issues related to our proposed scheme, i.e., the modifications required when applying our proposed scheme to current deployed WLANs.

Our proposal needs a client to estimate the \(p\) of each BSS, after carrying out passive or active scanning. However, a newly arrived client cannot derive exactly \(p\) due to hidden nodes; if an AP periodically broadcasts beacon/probe request frames including the \textit{RetryRatio} related to \(p\) directly, each client can derive exactly \(p\) of each BSS by overhearing the beacon/probe request frames. The \textit{RetryRatio} is a good indicator to select the best AP, since \textit{RetryRatio} reflects the traffic tendency which has been experienced 5 sec before. Obviously, this field is only a few bytes, so it does not result in significant overhead.

In addition, if an AP periodically broadcasts beacon/probe request frames including \(P_{idle}\), each client can derive the \textit{Eff} of each BSS by listening to the beacon/probe request frame. The length of this field is only a few bytes, so it does not result in significant overhead, either.

Finally, we propose the handoff procedure reflecting traffic dynamics. Each station first independently selects an appropriate AP according to the scheme previously described when joining in WLANs. The AP selected may then become a poor choice, since the number of STAs accommodated by each AP can change due to new STAs arriving, and because client traffic pattern is irregular. Then we propose the Dynamic AP selection to cope with the various changes in WLANs: if the \textit{AD RetryRatio} which a client receives piggybacked in the beacon frame is more than 0.3, each client processes re-scan and find a proper AP.

\section{Performance Evaluation}

In this section, we demonstrate the effectiveness of our proposed scheme compared with existing approaches through ns-2 simulation [17].

\subsection{Simulation Setup}

We have enhanced 802.11 DCF mode in ns-2 simulator (ver. 2.33) to support our proposed scheme by modifying the beacon and probe frame. We simulate the IEEE 802.11b PHY. Carrier Sensing range is set to 550m. Transmission rate of each node depends only on the distance between a client and a target AP, and the correlation between transmission rate and distance refers to the ORiNOCO 11b Card Specification [18].

Path loss of radio signals is modelled by the TwoRayground model of ns-2. We assume that all STAs and APs use the same transmission power. We have simulated a multi-cell network that consists of 9 APs operating on the same channel, and all STAs are randomly distributed under the coverage of at least 1 AP. We have chosen this scenario to demonstrate the proposed scheme effectiveness under adjacent BSSs’ transmission conditions. Rate adaptation is disabled in our scenario, and the arrival time of each station is uniformly distributed over a period of 30sec.

During the simulation, each STA generates traffic to its AP; offered traffic is Exponential On/Off UDP traffic to make it dynamic, and the packet size is 1000 bytes. We also assume that all STAs in the system always have pending some messages for the AP.

We have carried out simulation of experiments for 100 sec, actually going a time interval of 68 sec from 32 to 100 sec to calculate the aggregated throughput. We compare the performance of our proposed scheme with similar approaches in [12], [14]. To simulate the approaches in [12], [14], we use the empirical BER (Bit Error Rate) vs. SNR (Signal-to-Noise Ratio) curves provided by Intersil to estimate the FER (Frame Error Rate) [19]. All results are averaged over 10 runs.

\subsection{Simulation Result}

Fig. 3 presents the aggregated throughput of the proposed scheme and existing approaches, when the number of STAs is increasing from 50 to 80. While the number of STAs is increasing, the aggregated throughput of all schemes grows linearly until it reaches the saturation point at which throughput stops increasing; after that point, the aggregated throughput is deceased sharply though the number of STAs is increasing. It is clear that our proposed scheme can achieve higher throughput than other approaches; we observe that the proposed scheme shows about 10.2% and 33.5% better aggregated throughput than other approaches respectively. The approach’s performance of [12] is less than our scheme’s performance, since this approach does not reflect adjacent BSSs’ transmission. The approach’s performance of [14] is less than our scheme’s performance since the measurement value of
ATD directly related to the number of active STAs has large variations; this factor influences the expected throughput, resulting in wrong AP selection.

V. CONCLUSIONS

In this paper, we have proposed a novel traffic-aware AP selection for multi-rate in WLANs. Unlike previous approaches, we consider multiple transmission rates as well as adjacent BSSs’ transmission, by estimating more accurately and less intrusively: by exploiting the Retry field in the MAC header, we estimate the number of active STAs in the each BSS implicitly. We also compared the proposed scheme with the existing approaches. Through ns-2 simulation, we have demonstrated that the proposed scheme yields the most significant performance enhancement compared with previous work.

As future work, we plan to extend the proposed scheme to Wireless Mesh Networks (WMN).

ACKNOWLEDGMENT

This research was supported by the Ministry of Knowledge Economy, Korea, under the Information Technology Research Center support program supervised by the Institute of Information Technology Advancement.(grant number IITA-2009-C1090-0902 -0006)

REFERENCES

http://www.isi.edu/nsnam/ns/
[16] Qorvo SDR Hardware