Fast scanning schemes for IEEE 802.11 WLANs in virtual AP environments

Sunggeun Jin\textsuperscript{a,}* , Munhwan Choi\textsuperscript{b} , Lei Wang\textsuperscript{c} , Sunghyun Choi\textsuperscript{b}

\textsuperscript{a} ETRI, 218 Gajeongno, Yuseong-gu, Daejeon, Republic of Korea
\textsuperscript{b} School of Electrical Engineering and INMC, Seoul National University, Republic of Korea
\textsuperscript{c} School of Software of Dalian University of Technology, Dalian, China

**Article info**

**Article history:**
Received 3 October 2010
Received in revised form 17 February 2011
Accepted 20 April 2011
Available online 30 April 2011
Responsible Editor: C. Westphal

**Keywords:**
IEEE 802.11 WLAN
Handoff
Scanning

**Abstract**

Recently, provisioning a fast handoff in IEEE 802.11 Wireless Local Area Network (WLAN) has been an attractive research issue since the incurring latencies are unsatisfactory to support today's fast emerging realtime services with stringent Quality of Service (QoS) requirements. The 802.11 handoff consists of (1) scanning, (2) authentication, and (3) reassociation. Particularly, the scanning has been studied intensively since it accounts for a major portion of the handoff latency due to potential overheads in the scanning operation. That is, (1) a scanning STA should stay in a scanned channel waiting for Access Points (APs)' responses without any assurances of APs' existence after request frame transmissions; moreover, (2) per-channel waiting time is not explicitly defined in the 802.11 standard. In order to reduce the overheads, we propose novel scanning schemes composed of two phases: (1) channel selection phase; and (2) AP search phase. In the channel selection phase, a scanning STA identifies and selects the channels suitable for the 802.11 scanning by assessing all employed channels with Request-to-Send/Clear-to-Send (RTS/CTS) handshaking in a virtual AP environment. Then, in the AP search phase, it performs a unicast-based scanning in order to search for an AP suitable for its handoff in a selected channel. For a proper handoff decision depending on STA's requirement, we build two algorithms referred to as first-fit and near best-fit algorithms, respectively. We demonstrate the superiority of our proposed schemes to existing approaches via simulations with realistic time-varying channel models in various simulation environments. The simulation results demonstrate that the proposed scanning schemes provide relatively shorter handoff latencies as (1) a scanning STA has shorter channel switching time; (2) an 802.11 WLAN employs more channels; and (3) the APs are deployed more densely.

© 2011 Elsevier B.V. All rights reserved.

**1. Introduction**

IEEE 802.11 handoff consists of three components, namely scanning, authentication, and reassociation: (1) a scanning is a process to find Access Points (APs) suitable for handoff and acquire the information of the found APs; (2) an authentication is a process to identify an individual STA (STA) authorized to associate; and (3) an association is a process to establish a connection between an AP and a STA while a reassociation occurs for an already-associated STA to make a new connection with another AP. The previous studies [6,7] show that scanning is the most time-consuming job among these three components. The 802.11 scanning accounts for more than 90% of the overall handoff latency. For this reason, reducing scanning time is expected to improve the overall handoff latency more than anything else [8].

The 802.11 scanning takes more than 300 ms to scan all the channels in typical 802.11 WLANs [6,7]. Accordingly, such a high latency can incur severe service disruption. Therefore, many studies have been made in order to improve the handoff latency by reducing the scanning time...
The recently-developed 802.11k standard also defines some functions for this purpose [4].

The reason for high scanning latency is rooted in non-optimized request-and-response operations relying on obscure waiting time in order to find neighboring APs. In other words, a scanning STA broadcasts Probe Request frames, and then, waits for Probe Response frames from APs in each frequency channel. More in detail, the 802.11 standard specifies that a STA should stay in a scanned channel for MaxChannelTime when the scanned channel is sensed busy for MinChannelTime after a Probe Request transmission. However, either the scanned channel could be sensed busy due to the frame transmissions of neighboring STAs without reachable APs or, even if there exist APs nearby, they might not be suitable for a handoff. In either case, a scanning STA should spend MaxChannelTime, which is not specified in the 802.11 standard but typically longer than 10 ms, in waiting. It implies that a scanning STA may waste scanning time as much as MaxChannelTime for only waiting in each busy channel irrespective of the existence of a satisfactory AP. The value of MinChannelTime is not specified in the standard either. However, it is a minimum time required to determine if AP does not exist in a scanned channel. Therefore, a properly chosen MinChannelTime can contribute to the improvement of the scanning latency.

In summary, we discover that we can improve the scanning latency if we consider two cases: (1) if a scanning STA can recognize a priori if there is no AP satisfying STA’s requirement without waiting for MaxChannelTime in a busy channel, it is possible to save scanning time as much as MaxChannelTime by skipping a scanning operation in the busy channel; (2) if a STA is able to find out that no AP exists in a scanned channel without sensing idle for MinChannelTime, it is also possible to reduce scanning latency as much as MinChannelTime in the channel.

From this discovery, we propose new scanning schemes with two phases, i.e., channel selection phase and AP search phase. First, we design a new procedure to find channels, where there may exist an AP providing satisfactory Quality-of-Service (QoS) for handoff delay and service data rate. The procedure can be performed by assessing each channel with Request-to-Send/Clear-to-Send (RTS/CTS) handshaking in the virtual AP environment, which will be described in detail later. Second, we propose a unicasting-based scanning scheme for the AP search phase. A scanning STA performs the unicasting-based scanning in order to search for an AP suitable for its handoff in the channels identified via the preceding channel selection phase. Third, we combine our proposed schemes into two algorithms, i.e., first-fit and near best-fit. A STA adopting first-fit initiates a handoff to the first-found satisfactory AP in the first-found channel. On the other hand, a STA employing near best-fit algorithm chooses the channel estimated to be the most appropriate by comparing all candidate channels.

The proposed scanning schemes are distinguished from the existing scanning schemes in three aspects: (1) a prior channel selection based on the channel assessment employing RTS/CTS handshaking in each channel; (2) a unicasting-based Probe Request/Response exchange, i.e., unicasting-based scanning, performed in the chosen channels; and (3) the virtual AP environment supporting the prior channel selection and unicasting-based scanning.

The rest of this paper is organized as follows: in Section 2, we introduce the previous work for scanning improvement and discuss the related issues. In Section 3, we propose new scanning schemes employing virtual AP concept. In Section 4, we demonstrate the superiority of the proposed scanning schemes through extensive simulations with various practical scenarios. Finally, Section 5 concludes the paper.

2. Related work

An 802.11 handoff consists of scanning, authentication, and (re) association. A scanning is a procedure to find the best available AP(s) for either handoff or system startup. In the 802.11 standard, both active scanning and passive scanning are defined [1]. For an active scanning, a STA broadcasts one or more Probe Request frames in each channel, and then waits for Probe Response frames from neighboring APs. When a STA conducts a passive scanning, it attempts to listen to beacon frames, which APs periodically transmit.

Fig. 1 shows a typical example of the 802.11 handoff when a STA conducts active scanning. Specifically the zoomed-in part of this figure illustrates the 802.11 active scanning procedure. If the channel is sensed to be idle for MinChannelTime after broadcasting a Probe Request frame, the STA concludes that there is no AP in the scanned channel. Accordingly, it continues to broadcast a Probe Request in the next channel. When the channel becomes busy within MinChannelTime, it extends its stay at the channel until MaxChannelTime in order to receive Probe Responses from AP(s) since a busy channel indicates the possibility that AP(s) could exist in the scanned channel.

In this figure, ChannelSwitchingTime reflects the latency required to switch the scanned channel. This value is an implementation-dependent factor influencing the scanning latency. Today’s high-tech commercial transceivers already support 100 µs [25]. Intel and Netgear provide practical devices supporting 200 µs and 6 ms, respectively [5], and hence we consider both 200 µs and 6 ms for our evaluations later. The authentication or (re)association is simply performed by exchanging Authentication or (Re)association Request and Response messages.

The authors of [19] propose an efficient passive scanning scheme assuming that ChannelSwitchingTime is long. However, such a passive scanning is not likely to be very attractive for today’s WLAN devices with small ChannelSwitchingTime. Moreover, it might be costly to maintain synchronization consistently among deployed APs. Therefore, we in this paper concentrate only on active scanning. It is a faster scanning method in general.

Practically, when there exist contending STAs in a scanned channel, a scanning STA would sense the scanned channel to be busy due to the contending STAs’ frame...
transmissions. Accordingly, it is possible that a scanning STA should wait until $\text{MaxChannelTime}$ rather than $\text{MinChannelTime}$ in vain without detecting any AP. For an easy explanation about such a case, we consider an extreme scenario depicted in Fig. 2, where we assume that the AP is too far to serve STA$_a$ and it senses its channel busy when STA$_b$ transmits a frame since it is situated within Clear Channel Assessment (CCA) range of STA$_b$. In contrast, STA$_b$ is sufficiently close to the AP for high frame transmission rate (e.g., 5.5 Mbps of IEEE 802.11b). In addition, STA$_b$ is loaded highly enough to transmit frames within $\text{MinChannelTime}$. In this example, STA$_a$ has to wait for $\text{MaxChannelTime}$ in vain after broadcasting a Probe Request since it senses the channel busy due to STA$_b$'s frequent frame transmissions. Apparently, the more neighboring STAs exist, the more likely a scanning STA waits until $\text{MaxChannelTime}$. Later, our simulation results actually show that unnecessary $\text{MaxChannelTime}$-waiting is more common than expected.

For a given frequency band, many legitimate channels might be allowed for 802.11 WLANs, and hence, it could be time-consuming to probe all the defined channels. For this reason, vendors typically implement their own proprietary algorithms for scanning in a reduced set of channels since IEEE 802.11 standard defines a scanning procedure without a specific algorithm [8,17]. For example, under Observed-Scanning scheme [21], scanned channels are limited to a set of non-overlapping independent channels observed during previous scannings.

However, the authors of [19] realize that Observed-Scanning does not work well when the number of non-overlapping channels is high as in the 802.11a PHY, which defines 12 (or even more depending on the regions) non-overlapping channels [2]. Therefore, several schemes are proposed to reduce scanning time by using neighbor graph [15–17]. According to their scheme, an AP can provide STAs with candidate channels suitable for handoff among all possible channels, and hence, STAs can attempt to scan with a reduced number of candidate channels. The 802.11k standard reflects their contribution as discussed later in detail. Although the authors of [15–17] provide efficient schemes to reduce candidate channels for scanning, the scanning latency remains relatively long due to large values of $\text{MaxChannelTime}$ and $\text{MinChannelTime}$.

The authors of [18] propose aggressive values for both $\text{MaxChannelTime}$ and $\text{MinChannelTime}$ to reduce the scanning latency. They argue that 10 Time Units (TUs) (~10.24 ms) is reasonable for $\text{MaxChannelTime}$ while one TU (~1.024 ms) for $\text{MinChannelTime}$ in 802.11b WLANs. However, as the authors addressed, various parameters such as channel conditions, the number of contending

---

2 In the case of 2.4 GHz, 13 channels are allowed in Korea and many other countries while 11 channels are in the US.
3. Fast scanning schemes

3.1. Basic concept of virtual AP

The 802.11 standard states that an AP addressed by an RTS frame transmits the corresponding CTS frame after a Short InterFrame Space (SIFS) period if the channel is determined idle. This implies two aspects. First, a STA can transmit an RTS only if the destination Medium Access Control (MAC) address is known. Second, if the channel remains idle until Point coordination function InterFrame Space (PIFS) time, i.e., a SIFS plus one slot time, after an RTS transmission, then there exists no AP responding to the RTS. Now, since no CTS corresponding to an RTS is possibly an indication of the absence of AP, an RTS/CTS exchange in a particular channel can be used to determine whether there is an AP or not.

In order to utilize the RTS/CTS exchange, MAC address of a target AP should be known in advance. For this purpose, we propose to reserve a specific MAC address, which is shared among APs and STAs, only for the RTS/CTS exchange purpose in a network. As proved experimentally later, it is a feasible assumption that all APs in a network share a reserved MAC address called shared address, which is known to all STAs in the network in advance. Since the shared address is a globally unique MAC address, it is impossible to make STAs confused with other MAC addresses for normal operations.

Additionally, we configure each of the APs to operate with its own unique MAC address for a normal operation simultaneously. That is, the WNIC of each AP is configured to operate with both its own unique MAC address for normal operation and the shared address only for performing the proposed scanning schemes. This environment is very practical because most of today’s WNICs (e.g., Atheros chipset-based ones) already support multiple MAC addresses concurrently on a single physical WNIC and the APs employing such a WNIC are referred to as virtual APs [23].

3.2. Experiment with virtual APs

In order to demonstrate that virtual AP operation is feasible, we conduct an experiment by establishing two APs operating in different channels as shown in Fig. 4. Each AP has a single Atheros chipset based WNIC configured to serve two MAC addresses simultaneously. In this configuration, two APs share the same MAC address, namely, “02:30:0D:1D:38:29,” but maintain their own unique MAC address simultaneously, i.e. “08:30:0D:1D:38:29” for AP1 and “0C:30:0D:1D:38:29” for AP2, respectively.

Fig. 3. Scanning failure probability when a STA attempts to hand off.

3.1. Basic concept of virtual AP

The 802.11 standard states that an AP addressed by an RTS frame transmits the corresponding CTS frame after a

---

3. 2 Mbps is used for Probe Request transmissions when the number of contending STAs is 10 and the normalized offered load is 0.6.

---
For the configuration, we follow the steps described in [26]. When the STA transmits an RTS frame with each of AP’s shared address and own unique MAC address, we observe using a packet sniffer software, called Airopeek [27], that the AP sends CTS frames with a proper source MAC address each time.

3.3. Channel assessment with virtual APs

As discussed earlier, a scanning STA assesses employed channels with RTS/CTS handshaking in the channel selection phase. For the purpose, we design a new channel assessment in the proposed virtual AP environment. During the channel selection phase, a STA concentrates on determining which channels include AP(s) suitable for a handoff without Probe Request/Response exchange. Fig. 4 shows an example of a channel assessment utilizing virtual APs. A single WNIC of each AP holds two different MAC addresses, where “02:30:0D:1D:38:29” is the shared address. A STA will transmit an RTS with the destination address of “02:30:0D:1D:38:29” for the channel assessment when it needs to handoff.

When a STA transmits an RTS addressed at the shared address to virtual APs, the virtual APs receiving the RTS should respond to the STA with a corresponding CTS. Note that the RTS frame is a class-1 frame so that an AP upon reception of an RTS responds with a CTS even if the RTS-transmitting STA is not authenticated or associated. There might not be a successful CTS reception after an RTS transmission due to the following four reasons: (1) the absence of APs; (2) a collision or a channel error of the RTS; (3) a channel error of the CTS; and (4) a collision of CTSs from multiple APs with the shared address.

In the proposed scheme, all the virtual APs share one MAC address. Accordingly, if multiple virtual APs respond to the RTS after SIFS time with the exact synchronization, the STA performing channel assessment can successfully receive the CTS. It is because its receiver treats synchronously transmitted CTS frames as multiple copies of a single CTS, which traveled through multiple radio paths. The 802.11 receivers are commonly designed to handle such multipath transmissions.

However, the 802.11 standard allows ±10% margin of one slot time for the SIFS time interval implementation. It implies the SIFS time is allowed to vary from 8 μs (15.1 μs) to 12 μs (16.9 μs) in the 802.11b/g (the 802.11a/g) channel. Since the cyclic prefix duration for the 802.11a/g is defined as 800 ns, the STA may experience CTS collision. Therefore, in practical 802.11 WLANs, a STA performing channel assessment either receives the corresponding CTS successfully or experiences CTS collision.

Upon CTS collision, a STA performing channel assessment senses the channel to be busy although it may fail to receive the CTS frame correctly. For this reason, the busy-channel status is used to indicate the existence of AP(s) in the channel. It implies that the proposed scheme works well even in the case that the responding CTS frames collide since the channel becomes busy for the period corresponding to the CTS reception time after a RTS transmission irrespective of whether a CTS is successfully received or not.

Accordingly, when the STA senses an idle channel after an RTS transmission, the STA assumes that there is no AP nearby, and will proceed with a channel assessment in the next channel. Since a PIFS time is long enough for an RTS timeout, which is the waiting time for receiving a CTS frame, to determine whether there is an AP replying to the RTS or not in the channel, the waiting time is reduced to a PIFS time when the channel is idle. However, in order to reduce the negative effect of collision and channel error cases for RTS, a STA retransmits an RTS once again if it is not able to sense busy channel following the first RTS transmission. That is, an RTS is transmitted twice in the absence of a CTS response in each channel.

3.4. Maximum achievable data rate under virtual APs

When a STA performs scanning, it should search the most appropriate AP, which may provide the highest data rate among detected APs. Typically, a STA estimates it from the Signal to Noise Ratio (SNR) measured during a scanning procedure. However, in the proposed scheme, the measurement might not yield a reliable result since it could be from multiple CTS frames transmitted by multiple APs. Therefore, we develop a scheme for the channel selection phase to find the channels in which there exists an AP expected to provide the highest data rate among detected APs by utilizing the fact that transmission ranges depend on the transmission rates as depicted in Fig. 5(a) and (b).

For the purpose, we utilize multiple RTS transmissions with incremental transmission rates as follows: a scanning STA transmits an RTS frame at the lowest transmission rate first. If the STA senses the channel to be busy due to the corresponding CTS, it then transmits another RTS at the next higher data rate, with which the RTS transmission covers a shorter range. In this procedure, an AP receiving an RTS should reply with a CTS at the lowest rate irrespective of the RTS rate. If a STA can sense the channel to be busy due to CTS(s) corresponding to an RTS transmitted at a given data rate, it can expect that at least one AP exists within the transmission range of the data rate. Accordingly, it continues to increase the RTS transmission rate until it fails to sense the channel busy in order to determine the maximum achievable data rate in the channel.

This procedure is apparently not for link adaptation. In the 802.11 WLANs, link adaptation typically attempts to find the most suitable data rate over a link by estimating the link quality. However, while a STA performs scanning as a part of a handoff, it should determine proper AP(s) urgently, and hence, it is impossible to measure link quality provided by all the detected APs for the handoff just as for link adaptation. In the proposed procedure, the finally obtained data rate indicates the maximum data rate, which an AP possibly provides, for handoff decision.

Fig. 6 shows an example assuming the 802.11b [3], where there exists a single neighboring AP. In this figure, a STA initially transmits the first RTS addressed at the shared address at the lowest rate, i.e., 1 Mbps. After a SIFS time, the corresponding CTS arrives at the STA, which implies the existence of an AP as well as the reachability via 1 Mbps. Now, the STA continues to transmit an RTS at the next higher data rate, i.e., 2 Mbps. Then, the STA again
successfully receives a CTS transmitted at 1 Mbps. However, if there is no CTS from any AP after transmitting an RTS at 5.5 Mbps, the STA concludes that the maximum achievable data rate in this channel is 2 Mbps.

In the 802.11a [2], eight data rates are defined. As shown in Fig. 5(a), the transmission ranges for 9 Mbps and 12 Mbps are nearly the same, and similarly, 48 Mbps and 54 Mbps are nearly the same. For this reason, we adopt data rates of 6, 12, 18, 24, 36 and 54 Mbps for our channel assessment in the case of the 802.11a. Note that the maximum achievable data rate among what have been found as the previously-scanned channels, which provide the data rate equal to \( \tau \), the STA determines whether it needs to update \( \tau \) and/or set \( \Gamma \) containing pairs of the SNR value and the corresponding channel number for the previously-scanned channels, which provide the data rate equal to \( \tau \).

3.5. Fast scanning algorithms

Typically, a handoff triggered when the SNR of a received frame drops below a predefined scan trigger threshold. In order to complete a handoff to an AP providing a satisfactory performance, we propose two algorithms with the maximum achievable data rate search procedure for the scanning, namely, near best-fit and first-fit algorithms, respectively. Table 1 shows a list of the notations used for the algorithms.

**Algorithm 1: Near best-fit algorithm**

1. let \( \gamma = 0 \), \( \tau = 0 \), \( \Gamma = \emptyset \);
2. for all \( \psi_i \in \Psi \) do
3. perform a channel assessment to find \( \rho_i \) and measure \( \gamma_i \) in channel \( \psi_i \);
4. if \( \tau > \rho_i \) then
5. continue;
6. else if \( \tau = \rho_i \) then
7. let \( \Gamma \leftarrow \Gamma \cup \{(\gamma_i, \psi_i)\} \);
8. else if \( \tau < \rho_i \) then
9. let \( \Gamma \leftarrow \{(\gamma_i, \psi_i)\}, \tau \leftarrow \rho_i \);
10. end if
11. end for
12. if \( \Gamma = \emptyset \) then
13. return false;
14. else
15. let \( M \leftarrow \arg\max_{i} \{(\psi_i, \gamma_i) \in \Gamma \} \);
16. return \( \psi_M \);
17. end if

The near best-fit algorithm, shown in Algorithm 1, provides a procedure to find the most desirable AP for a handoff in a channel providing the highest and satisfactory data rate with a significantly reduced time overhead. In the channel selection phase, a scanning STA adopting this algorithm tries a channel assessment in each channel \( \psi_i \) contained in set \( \Psi \), and then, it compares the maximum achievable data rate \( \rho_i \) obtained in channel \( \psi_i \) with the value of a temporary variable \( \tau \) containing the highest achievable data rate among what have been found as the maximum achievable data rate before the trial in \( \psi_i \). From the comparison, the STA determines whether it needs to update \( \tau \) and/or set \( \Gamma \) containing pairs of the SNR value and the corresponding channel number for the previously-scanned channels, which provide the data rate equal to \( \tau \).

If \( \tau > \rho_i \), it determines that \( \psi_i \) is not appropriate for handoff, and hence, continues the iteration for the next channel. It is possible that the STA obtains \( \rho_i \) identical to
the value of $\tau$. It then appends a pair of the measured SNR $\gamma_i$ from the channel and the corresponding channel number $\psi_i$ to set $I'$ for a future decision irrespective of whether a CTS is correctly received or not. For $\tau < \rho_i$, it flushes and updates $\tau$ and $I'$ with a newly obtained pair. By completing the iteration, the STA has the candidate channels for a satisfactory handoff. If there exist more than one element in set $I'$, it chooses the one with the highest SNR among the candidates.

However, it is possible that the AP providing the highest SNR value does not exist in the chosen channel but in another channel in set $I'$ due to simultaneous CTS transmissions from multiple virtual APs. Nevertheless, the proposed algorithm is valid for two reasons: (1) even in the worst case that the chosen channel does not include the AP with the highest SNR, the chosen channel, namely, near best channel, is guaranteed to include an AP providing the maximum data rate identical to that of the AP with the highest SNR since the channel is selected out of the channels in the set $I'$; (2) despite the potential sub-optimal operation, the proposed scheme reduces the scanning overhead while selecting a channel providing the highest data rate. As a result, it mitigates the realtime service disruption with a satisfactory data rate.

In the AP search phase, the STA transmits a unicast Probe Request addressed at the shared address in channel $\psi_M$, which is expected to provide the highest data rate. If a virtual AP receives a Probe Request, it responds with an ACK. Accordingly, it can recognize whether the AP receives the Probe Request without waiting for Probe Response. It is possible that multiple virtual APs reply with their ACK simultaneously so that there might be a collision. However, as in the case of a CTS collision, a busy-channel status is a sufficient condition for the successful Probe Request reception. In case that the channel is idle until PIFS time after the Probe Request transmission, the STA retransmits the Probe Request once more after another channel contention assuming that the first transmission was involved with either a collision or a channel error. Once a virtual AP receives the Probe Request, it accesses the channel to transmit Probe Response frame, whose source address is set to its own unique MAC address, not the shared address.

In this algorithm, we can set a very long MaxChannelTime compared with that proposed in [18] since a STA expects the existence of an AP with high probability as confirmed by both the RTS/CTS-based channel assessment and an ACK following a unicast Probe Request. Nevertheless, if a STA fails to receive a Probe Response after a long MaxChannelTime, it will switch to the next candidate channel, which is determined by the next largest value in set $I'$.

Algorithm 2: First-fit algorithm

1: \textbf{let} $\gamma \leftarrow 0$, $\tau \leftarrow r$, $I \leftarrow \emptyset$;
2: \textbf{for all} $\psi_i \in \Psi'$ \textbf{do}
3: \hspace{1em} \{perform a channel assessment to find $\rho_i$ in channel $\psi_i$\};
4: \hspace{1em} \textbf{if} $\tau > \rho_i$ \textbf{then}
5: \hspace{2em} \textbf{continue};
6: \hspace{1em} \textbf{else}
7: \hspace{2em} \textbf{return} $\psi_i$;
8: \hspace{1em} \textbf{end if}
9: \hspace{1em} \textbf{end for}
10: \textbf{return} false;

In the first-fit algorithm, shown in Algorithm 2, if a STA detects an AP which meets the minimum data rate requirement during a channel assessment, it immediately stops the channel assessment process, and then continues to transmit a Probe Request. As shown in Fig. 7, a STA performs the channel assessment successfully concluding that 54 Mbps is achievable in the scanned channel, and then it immediately transmits a unicast Probe Request after a SIFS time.

In order to reduce the unnecessary overheads upon performing near best-fit or first-fit algorithm, we employ a two-step approach for MaxChannelTime. A STA waits until 10 ms after successfully (i.e., with an ACK response) transmitting a unicast Probe Request. For the first 10 ms, it might receive multiple Probe Responses. If it does not receive any Probe Response, it waits for an extra 30 ms. In this case, if it receives a single Probe Response providing satisfactory SNR requirement for a handoff, it will immediately proceed to the handoff procedure.

3.6. Discussions

3.6.1. Channel assessment with AP channel report

According to the 802.11k standard [4], the AP channel report conveyed via periodic beacon frames from APs contains the information about the channels where neighboring APs operate. Since the goal of this scheme is to reduce the number of channels for scanning, this AP channel report scheme can be employed for our proposed schemes. However, the MAC addresses of neighboring APs are not conveyed as part of the AP channel report [4]. By combining the AP channel report with the proposed virtual APs, we can attain a fast scanning without incurring extra signaling overheads.

3.6.2. Coexistence with existing scanning schemes

A STA without the proposed scanning schemes should conduct an existing scanning scheme to search APs even in a proposed virtual AP environment. As explained earlier, the STA following the existing scanning schemes should broadcast Probe Request frames, and then wait for Probe Responses in a channel. Upon the reception of a Probe Request frame, a virtual AP replies with a Probe Response frame, of which the source and destination addresses are the unique MAC address of the virtual AP for its normal operation and the MAC address of the scanning STA, respectively, via...
channel contention. We already experimentally showed that virtual AP works well with both MAC addresses simultaneously in Section 3, and hence, the coexistence is feasible. It implies that the STA conducts the existing scanning schemes correctly with the ignorance of the proposed virtual AP environments.

4. Performance evaluation

We have developed a simulation model using ns2 [24] to evaluate our proposed scanning schemes. We consider the PHY characteristics of both 802.11a and 802.11b. Table 2 shows the parameters used for our simulations. Since the 802.11g adopts the same Orthogonal Frequency Division Multiplexing (OFDM) PHY as the 802.11a and shares the same timing parameters including SIFS, PIFS, and DIFS as the 802.11b, we do not evaluate 802.11g separately considering the evaluations for both 802.11a and 802.11b are sufficient.

For more realistic evaluations, we adopt a time-varying 802.11a/b channel model, which has been verified through the evaluation of a link adaptation algorithm [22]. In our simulations, APs periodically broadcast beacons and STAs measure the SNR of the broadcast beacons. If the measured SNR drops below a threshold, the corresponding STA triggers scanning. The threshold is set to 18.07 dB as shown in Table 2.

Table 2
Parameters for scanning simulations

<table>
<thead>
<tr>
<th>System dependent parameters.</th>
<th>IEEE 802.11b</th>
<th>IEEE 802.11a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slot</td>
<td>20 μs</td>
<td>9 μs</td>
</tr>
<tr>
<td>SIFS</td>
<td>10 μs</td>
<td>16 μs</td>
</tr>
<tr>
<td>CWmin</td>
<td>31</td>
<td>15</td>
</tr>
<tr>
<td>PHY header</td>
<td>192 μs</td>
<td>20 μs</td>
</tr>
<tr>
<td>Rates used for max. achievable rate detection</td>
<td>1, 2, 5, 5.11</td>
<td>6, 9, 18, 24, 36, 54</td>
</tr>
<tr>
<td>Probe Req./Resp. tx rate</td>
<td>1 (Mbps)</td>
<td>6 (Mbps)</td>
</tr>
</tbody>
</table>

Common parameters (in bytes)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTS size</td>
<td>20</td>
</tr>
<tr>
<td>CTS size</td>
<td>14</td>
</tr>
<tr>
<td>ACK size</td>
<td>14</td>
</tr>
<tr>
<td>Data frame size</td>
<td>500</td>
</tr>
</tbody>
</table>

Typical commercial WLAN devices are known to use 20 ms for MinChannelTime and 40 ms for MaxChannelTime, respectively [8]. However, as discussed in [18], 1 ms and 10 ms would be reasonably aggressive values for MinChannelTime and MaxChannelTime, respectively, and we also apply these values to existing scanning schemes in our simulations. For convenience, we refer to the 802.11 scanning with MinChannelTime of 20 ms and MaxChannelTime of 40 ms as the legacy 802.11 scanning. The 802.11 scanning with MinChannelTime of 1 ms and MaxChannelTime of 10 ms is referred to as the improved 802.11 scanning. Additionally, we refer to the proposed scanning schemes based on near best-fit and first-fit algorithms to near best-fit scanning and first-fit scanning, respectively. We set minimum requirement for the first-fit scanning to 5.5 and 36 Mbps for the 802.11b and the 802.11a WLAN, respectively. For simple notations inside the following figures, we adopt the short terms 'legacy,' 'improved,' 'best,' and 'first' to represent 'legacy 802.11 scanning,' 'improved 802.11 scanning,' 'near best-fit scanning,' and 'first-fit scanning,' respectively. In our simulations, a STA scans only the reported channels if the 802.11k AP Channel Report (ACR) is adopted. Otherwise, it scans all the employed channels by default.

In order to observe the 802.11 scanning behaviors in a highly-loaded environment, we vary the number of STAs per AP with the saturated channel load. Since scanning failures influence the scanning latency, we observe the relationship between the failure probability and the latency of each 802.11 scanning scheme for our evaluations. In Figs. 8 and 9, One Probe Request (One PR) represents the scanning scheme by which a STA broadcasts a Probe Request frame only once, and similarly, Two Probe Request (Two PR) represents the case that two consecutive Probe Request frames are broadcasted. It is well-known that two consecutive broadcasts of Probe Request frames can increase the success rate of a Probe Request reception [18].

The figures show that the consecutive broadcasts cause higher latencies, but reduce scanning failures since two consecutive Probe Request frames increase the successful transmission probability by overcoming collisions and channel errors. Fig. 8(a) shows interesting results for the 802.11a. The scanning latencies of the 802.11a increase less compared with the results for 802.11b shown in Fig. 9(a) as the number of contending STAs increases. There are two reasons for this: (1) in the 802.11a, transmission rates are higher than the 802.11b, and hence, it takes much

![Fig. 8. Scanning performance in an 802.11a WLAN.](image-url)
less time to transmit frames. Accordingly, a scanning STA wins an opportunity to transmit a *Probe Request* in the 802.11a WLAN in a shorter time than in the 802.11b WLAN; and, (2) since the *Probe Request* is a broadcast frame, the minimum contention window size is always used, where the minimum contention window size of the 802.11a and the 802.11b are 15 and 31, respectively. In a competitive channel, the priority achieved by small contention window size for *Probe Request* dominates the channel access delay caused by contentions.

In our simulations, Two *Probe Requests* are used. However, scanning failures still occur due to the channel access delay, channel error, and collisions of the *Probe Response* transmission. In order to cope with such scanning failures, for the rest of our simulations, we assume that the scanning process repeats until the scanning success probability becomes 99%.

In order to observe the effect as the number of scanned channels increases, we consider a scenario where a STA scans $n$ channels for its full scanning under a simple topology where two APs exist. Fig. 10(a) shows the topology, in which two APs are separated by 60 m and a STA moves back and forth between the APs at 3 km/h. In this scenario, we assume that there is no traffic while 11 and 12 channels are employed for the 802.11b WLAN and the 802.11a WLAN, respectively. A STA has to scan the entire employed channels. However, since two APs occupy different channels, the rest of the available channels are sensed to be idle all the time.

Fig. 12(a) and (b) show scanning latencies for both 802.11a and 802.11b under the assumption that the channel switching time is either 200 $\mu$s or 6 ms. From now on, we use the term ‘CST’ to stand for Channel Switching Time. Note that the results should be compared under the same condition depending on whether the ACR is employed or not. The figures show that the ACR helps eliminating unnecessary scannings in the channels where APs do not occupy. Accordingly, it keeps scanning latency constant irrespective of the number of the employed channels. Without it, scanning latency increases in proportion to the number of employed channels. However, when improved 802.11 scanning and legacy 802.11 scanning are utilized without the ACR, scanning latency increases more rapidly. This is because a STA utilizing our proposed schemes requires PIFS time to determine that there is no AP in an idle channel while it requires at least 1 ms, i.e., the MinChannelTime, otherwise.

In particular, Fig. 12(b) shows insignificant superiority of the proposed schemes compared with Fig. 12(a) when the channel switching time is 6 ms. However, the simulation scenario is an extreme case in order to observe how the employed channel number influences the scanning time. As detailed later, the proposed schemes enable scanning STA to avoid unnecessary waiting until MaxChannelTime in busy channels, where APs would provide an unsatisfactory performance. Herein, we exclude the effect caused by the unnecessary waiting. Even in this case, the proposed scanning schemes outperform the 802.11 scanning schemes. From Fig. 12(a) and (b), we observe that the proposed schemes are more beneficial with the shorter channel switching time.

In the next scenario, we observe the scanning behavior when there exist multiple APs at two positions as shown in Fig. 10(b). We assume that 12 and 11 channels are employed for the 802.11a WLAN and the 802.11b WLAN, respectively. Each of the APs situated at each position...
occupies a channel, which is different from that of the APs at the same position. However, it shares the channel with its neighboring AP at the other position. Therefore, the number of occupied channels are the same as the number of APs in each position. In fact, this scenario might be unrealistic. However, it is beneficial to observe how busy channel influences the scanning latency since a busy channel forces a scanning STA to wait until $\text{MaxChannelTime}$.

Fig. 13(a) and (b) show that the scanning latencies of both legacy 802.11 scanning and improved 802.11 scanning increase remarkably according to the number of occupied channels while first-fit’s and near best-fit’s do not. The reason is that a STA utilizing the legacy 802.11 scanning or the improved 802.11 scanning has to wait until $\text{MaxChannelTime}$ in each busy channel, where an AP replies with a Probe Response. On the other hand, when utilizing the proposed schemes, a STA does not wait for $\text{MaxChannelTime}$ in all busy channels, thus the reducing scanning time. This result implies that the proposed scanning schemes are better as more channels are occupied.

We perform our simulations by incorporating a grid topology, in which 16 APs are located by forming a 4-by-4 grid as shown in Fig. 11(a). In this topology, adjacent APs are separated by either 15 m or 30 m. We additionally consider the case that the APs are randomly scattered inside a square as depicted in Fig. 11(b). The length of a side of the square is set to 90 m. In both topologies, each AP occupies a single arbitrarily selected channel, and a STA follows the random waypoint mobility model [9]. In this model, a STA begins moving to a randomly chosen position. Once it starts moving, it keeps its speed as well as the direction. When it arrives at the destination, it repeats its movement toward a newly chosen destination.

The following simulation results show the performances in grid and random topologies. For convenience, we use the term ‘NoCH’ to represent the number of employed channels. Fig. 14 (a) and (b) show scanning latencies for both the 802.11a WLAN and the 802.11b WLAN depending on the distance between adjacent APs in a grid topology. Each of them is assumed to employ 3 or 12 channels. In the figures, it is observed that the scanning STA suffers from a bit higher latency in the case of 15 m than 30 m. It represents that the scanning latency becomes higher as APs are more densely deployed. In a densely deployed WLAN, it is highly possible that a scanned channel would be busy. It implies that the scanning STA waits for $\text{MaxChannelTime}$ with high probability. Accordingly, it incurs a long scanning latency.

As observed in the figures, the ACR helps a STA avoid unnecessary scannings, thus reducing the scanning latencies when many channels are employed. Even so, the scanning latencies of the 802.11 scanning becomes longer compared with the case when the number of channels is 3. As neighboring APs occupy more channels, the STAs performing the 802.11 scanning should wait for $\text{MaxChannelTime}$ in more channels, thus resulting in a longer scanning latency. However, the scanning latencies are basically the same irrespective of utilizing the ACR since the number of employed channels is small enough for neighboring APs to occupy every possible channel. It implies that the ACR is not useful when the number of employed channels is small.
On the other hand, the first-fit scanning’s latency becomes slightly longer in a sparsely deployed WLAN. The longer distance APs are separated at, the higher probability that it takes more time to find a satisfactory AP since the signal strength basically decreases according to the distance. Fig. 15(a) and (b) show the scanning latencies in the case that the channel switching time is 6 ms. The figures show that the proposed scanning schemes still outperform the 802.11 scanning schemes in a similar manner to the case that the channel switching time is 200 μs.

Fig. 16(a) and (b) depict the scanning latencies under the random topology assuming that the channel switching time is 200 μs. Each figure illustrates the obtained results in cases that 10 and 20 APs are deployed in the same sized square and the length of the square’s lateral is 90 m. Similar to the grid topology case, a more densely deployed network drives a STA performing the 802.11 scanning to spend its time for waiting in more channels. Nevertheless, first-fit enables STAs to find a suitable AP with less time and the near best-fit makes the scanning latency nearly unchanged.

Fig. 17(a) and (b) show the simulation results when channel switching time is 6 ms. In this case, the proposed scanning schemes outperform the 802.11 scanning schemes. The superiority is highlighted in densely deployed network environments for the same reason explained earlier. Consequently, the simulation results prove the proposed scanning schemes always outperform the existing 802.11 scanning.
From Figs. 12 and 17, when a STA adopts the near best-fit scanning scheme without the ACR, in many cases, we observe that the STA experiences shorter scanning latencies compared with the case that it conducts the 802.11 scanning with the ACR. Additionally, the figures show that the first-fit scanning outperforms the near best-fit scanning all the time because a STA stops scanning immediately after it finds an AP.

There exists an optimal AP providing the best signal quality at a STA’s decision time instant. Even if the proposed near best-fit and first-fit algorithms are not designed to find such optimal AP, it will be still interesting to evaluate the chance for a handing off STA to find optimal APs with the proposed algorithms. For this purpose, we measure a ratio between the number of optimal AP selections and the total number of AP selections when the proposed algorithms are utilized for the 802.11 handoff. The ratio is called hit ratio. The simulations are conducted in the grid topology.

Fig. 18 shows the hit ratio for both 802.11a and 802.11b WLANs with the grid topology. In this figure, D is the distance between APs in the grid topology. Near best-fit’s hit ratios are slightly influenced by the ACR and the distance between APs. In the 802.11a WLAN, near best-fit’s hit ratios are around 95% while first-fit’s hit ratios are around 40% or 60% depending on the ACR. First-fit’s hit ratios are more sensitive to the ACR compared with near best-fit’s since ACR is helpful to eliminate potential first channels providing satisfactory performance. However, in the 802.11b WLAN, the hit ratios drop sharply. Particularly, first-fit’s hit ratios are around 4% and 8% when D = 15. An 802.11b AP has wider coverage than an 802.11a AP since the 802.11b WLAN’s supported rates are lower than the 802.11a WLAN’s. The proposed algorithms rely on RTS/CTS exchanges for channel assessment and wider coverage provides higher probability for successful RTS/CTS exchanges. Therefore, 802.11b STA can find the first satisfactory AP faster than 802.11a STA. However, such first found satisfactory AP is not likely to be the optimal AP. It implies that 802.11a AP’s small coverage is more helpful to find an optimal AP.

5. Conclusion

We propose to divide the scanning process into the channel selection phase and the AP search phase for a fast handoff in order to accelerate the AP-finding process. In the channel selection phase, a scanning STA can find an AP with channel assessment working based on virtual AP. By utilizing the channel assessment scheme, we develop two algorithms to improve the scanning latency significantly, i.e., near best-fit and first-fit algorithms. Near best-fit helps a scanning STA find the AP providing the highest data rate among neighboring APs. On the contrary, first-fit enables a scanning STA to stop its scanning when it finds an AP satisfying its requirement. Through the extensive simulations, we observe our proposed schemes outperform the existing scanning schemes. Our proposed schemes are expected to improve the service quality of emerging delay sensitive applications such as VoWLAN.
Acknowledgements

This research was supported by the KCC (Korea Communications Commission), Korea, under the R&D program supervised by the KCA (Korea Communications Agency) (KCA-2011-08913-04003).

References


Sunggeun Jin is a senior engineer working for ETRI, which he joined in 1998, Korea. Prior to joining ETRI, he received his B.S. and M.S. degrees in School of Electrical Engineering and Computer Science at Kyungpook National University (KNU), Korea, in 1996 and 1998, respectively. He received his Ph.D. at School of Electrical and Computer Engineering, Seoul National University (SNU), Korea, August, 2008. He has participated in standard developments including IEEE 802.11v, IEEE 802.16j, IEEE 802.16m, and IEEE 802.11ad. He has served as a TPC member for IEEE WCNC 2008, ICUFN 2009, ICST BROADNETS 2010, and IEEE GLOBECOM 2011. Also, he co-organized peer reviews for journals and conferences such as IEEE TMC, IEEE INFOCOM, IEEE ICC, IEEE GLOBECOM, and IEEE WCNC.

Munhwan Choi received the B.S. degree in Electrical Engineering and Computer Science from Seoul National University, Seoul, Korea in 2007. He is currently working toward a Ph.D. degree in the Department of Electrical Engineering and Computer Science from Seoul National University, Seoul, Korea. His current research interests include algorithmic design and protocol development for various communication systems such as IEEE 802.11 wireless local area networks and 60 GHz wireless personal area networks. He is a student member of the IEEE.

Lei Wang is currently an associate professor with Dalian University of Technology, and the Deputy Chair of Department of Network Engineering. He received the B.S., M.S. and Ph.D. from Tianjin University, China, in 1995, 1998, and 2001, respectively. He was a Member of Technical Staff with Bell Labs Research China (2001–2004), a senior researcher at Samsung, South Korea (2004–2006), a research scientist in Seoul National University (2006–2007), and a research associate in Washington State University, Vancouver, WA, USA (2007–2008). His research interests include wireless ad hoc networks, sensor network, and embedded systems. He is a member of IEEE, ACM and a senior member of China Computer Federation (CCF).

Sunghyun Choi is an associate professor at the School of Electrical Engineering, Seoul National University (SNU), Seoul, Korea. Before joining SNU in September 2002, he was with Philips Research USA, Briarcliff Manor, New York, USA as a Senior Member Research Staff and a project leader for three years. He was also a visiting associate professor at the Electrical Engineering department, Stanford University, USA from June 2009 to June 2010. He received his B.S. (summa cum laude) and M.S. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology (KAIST) in 1992 and 1994, respectively, and received Ph.D. at the Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor in September, 1999. His current research interests are in the area of wireless/mobile networks with emphasis on wireless LAN/MAN/PAN, next-generation mobile networks, mesh networks, cognitive radios, resource management, data link layer protocols, and cross-layer approaches. He authored/coauthored over 140 technical papers and book chapters in the areas of wireless/mobile networks and communications. He has co-authored (with B. G. Lee) a book "Broadband Wireless Access and Local Networks: Mobile WiMAX and WiFi," Artech House, 2008. He holds 19 US patents, 10 European
patents, and 11 Korea patents, and has tens of patents pending. He has served as a General Co-Chair of COMSWARE 2008, and a Technical Program Committee Co-Chair of ACM Multimedia 2007, IEEE WoWMoM 2007 and IEEE/Create-NetCOMSWARE 2007. He was a Co-Chair of Cross-Layer Designs and Protocols Symposium in IWCMC 2006, 2007, and 2008, the workshop co-chair of WILLOPAN 2006, the General Chair of ACM WMASH 2005, and a Technical Program Co-Chair for ACM WMASH 2004. He has also served on program and organization committees of numerous leading wireless and networking conferences including ACM MobiCom, IEEE INFOCOM, IEEE SECON, IEEE MASS, and IEEE WoWMoM. He is also serving on the editorial boards of IEEE Transactions on Mobile Computing, IEEE Wireless Communications, ACM SIGMOBILE Mobile Computing and Communications Review (MC2R), Journal of Communications and Networks (JCN), Computer Networks, and Computer Communications. He has served as a guest editor for IEEE Journal on Selected Areas in Communications (JSAC), IEEE Wireless Communications, Pervasive and Mobile Computing (PMC), ACM Wireless Networks (WINET), Wireless Personal Communications (WPC), and Wireless Communications and Mobile Computing (WCMC). He gave a tutorial on IEEE 802.11 in ACM MobiCom 2004 and IEEE ICC 2005. From 2000 to 2007, he was a voting member of IEEE 802.11 WLAN Working Group. He has received a number of awards including the Young Scientist Award awarded by the President of Korea (2008); IEEK/IEEE Joint Award for Young IT Engineer (2007); the Outstanding Research Award (2008) and the Best Teaching Award (2006) both from the College of Engineering, Seoul National University; the Best Paper Award from IEEE WoWMoM 2008; and Recognition of Service Award (2005, 2007) from ACM. Dr. Choi was a recipient of the Korea Foundation for Advanced Studies (KFAS) Scholarship and the Korean Government Overseas Scholarship during 1997–1999 and 1994–1997, respectively. He is a senior member of IEEE, and a member of ACM, KICS, IEEK, KIISE.