Performance Enhancement for Mobility Management in Wireless LANs

Li Jun ZHANG and Samuel PIERRE
Mobile Computing and Networking Research Laboratory (LARIM)
Department of Computer Engineering, Ecole Polytechnique de Montreal
Montreal, Canada
Email: {lijun.zhang, samuel.pierre}@polymtl.ca

Abstract—Due to the small radio area covered by an Access Point (AP), handoffs frequently occur in Wireless Local Area Networks (WLANs) as Mobile Stations (MSs) move their associations from one AP to another. This paper presents a new handoff management scheme to support ongoing real-time applications while MSs change their network point of attachment. This approach consists of minimizing the total number of scanned channels, as well as the probe-waiting time for each examined channel. Performance is evaluated through simulations whose results show that our proposal delivers better performance, compared to the IEEE 802.11b Standard, the IEEE 802.11b Standard with Min in which an MS only waits for MinChannelTime on each examined channel and two other well-documented solutions in the literature: Selective Scanning plus Caching and Neighbor Graphs.

Keywords—fast handoff, mobility management, performance, VoIP, wireless LAN

I. INTRODUCTION

Nowadays, with the rapid advancement of wireless technology, contemporary Internet Service Providers (ISPs) can deliver real-time multimedia services, such as audio streaming and video conferencing to mobile/wireless subscribers. Nevertheless, such ‘infotainment’ media streaming applications impose strict Quality of Service (QoS) requirements on wireless networks. On the other hand, using spread-spectrum technology, Wireless Local Area Networks (WLANs) provide stations with free mobility within the radio coverage of their associated Access Point (AP), while they are still connected to the network. However, the IEEE 802.11 specification family (IEEE 802.11a/b/g) legacy does not provide fast handoff support for mobile hosts as they move from the radio coverage area of one AP to another. Hence, as the handoff process turns into inevitably long latencies and packet loss when transferring ongoing calls or data sessions, it is critical to devise more efficient mobility management schemes for stations with real-time applications in progress, roaming in WLANs.

The IEEE 802.11 Standard [1] defines two modes of operation: infrastructure and ad hoc. In the infrastructure mode, an AP comprises a Basic Service Set (BSS) and provides network connectivity to its associated Mobile Stations (MSs). One or more APs comprise an Extended Service Set (ESS) that covers a larger service area. In the ad hoc mode, two or more MSs form a Peer-to-Peer (P2P) wireless network without deploying any APs. This paper is concerned only with the infrastructure mode.

An ideal WLAN can provide successive radio signal coverage for MSs in its service area. An MS may decide to handoff from one AP to another for mobility reasons, AP load balancing state or signal fading. The MAC layer (L2) legacy handoff process specified in the IEEE 802.11 Standard [1] comprises three phases: scanning, authentication and reassociation.

The objective of scanning is to determine the characteristics of the available BSSs within the radio range of the MS. Two scanning modes are defined by the IEEE 802.11 Standard: Passive and Active Scanning [1]. The former allows an MS to listen on each existing channel, waiting for beacon frames periodically sent by adjacent APs, while the latter involves the generation of probe request frames by MSs and the subsequent processing of received probe responses from nearby APs.

After discovering accessible APs, the MS then selects one AP as its next associated AP according to certain preferences, such as the received signal strength indicator (RSSI), the support data rate, the number of frame retransmission [2], etc.

Then the MS launches the authentication procedure with the New AP (NAP). Generally, authentication strives to identify the MS as a member of the specified BSS and to authorize it to communicate with other stations within the same BSS [1]. Two authentication methods have been specified in the IEEE 802.11 Standard: Open System and Shared Key Authentication [1].

Open System Authentication involves the exchange of a pair of frames between the MS and the NAP: an authentication request as well as an authentication response. Usually, all stations can be authenticated.

Shared Key Authentication is an optional four-step process that uses a Wired Equivalent Privacy (WEP) key. An MS launches authentication by transmitting an authentication request to the NAP. Upon receiving this request, the NAP generates a challenge text using a WEP key and sends an
authentication response with such challenge text as a reply. The MS then encrypts the received challenge text with a shared WEP key and returns an authentication request with the encrypted challenge text to the NAP. The NAP subsequently decrypts this request using the shared key and compares the decrypted and the original challenge texts. If they are identical, the NAP transmits an authentication response to confirm a successful authentication.

Regardless of the authentication method used, the IEEE 802.11 Standard requires mutually acceptable responses for a successful authentication [1]. Moreover, the authentication procedure always happens before the association (or reassociation) procedure. Due to security flaws in open system and shared key authentications, the authentication methods specified in IEEE 802.11 have been replaced by IEEE 802.11i [3]. However, to maintain backward compatibility, IEEE 802.11i allows Open System Authentication and exchanging authentication messages after the reassociation phase [3] [4].

Reassociation is performed after a successful authentication. Furthermore, the IEEE 802.11 Standard specifies that each MS must be associated with a single AP at any given time [1] and an MS must issue a reassociation request frame to the NAP after a successful authentication. The reassociation request contains the concerning MS’s MAC address, its previous Service Set Identifier (SSID), the MAC address of the Previous AP (PAP). Upon receipt of this request, the NAP starts to use the Inter-Access Point Protocol (IAPP) to deliver the MS-related security context information from the PAP (or old AP) to the NAP [5]. In doing so, the NAP sends an Access-Request message to the RADIUS (Remote Authentication Dial-In User Service [6]) server, which then looks up the IP address of the PAP and verifies the SSID, before returning an Access-Accept message to the NAP. The Access-Accept message contains the IP address of the PAP and security block items required to establish a secure communication channel between APs. This process is called address mapping that matches the PAP’s MAC address to its IP address.

After exchanging security elements through Send-Security-Block and ACK-Security-Block packets, both APs own sufficient information to encrypt all further packets. Afterwards, the NAP sends an encrypted MOVE-notify packet to the PAP inquiring about the concerned MS’s context. Upon verifying the MS’s association, the PAP removes the MS from its association table and returns an encrypted MOVE-response packet to the NAP, including the pertaining Context Block. Then, the NAP adds the MS into its association table and broadcasts a Layer 2 Update frame to inform all layer 2 devices, such as bridges and switches, that they must update their forwarding table for a given MS.

Briefly, the IAPP allows an AP to communicate with other APs in the same ESS, while minimizing opportunities for transmitting MSs’ security contexts over the air. However, context transfer using IAPP results in additional delays during handoff. Finally, the NAP sends a Reassociation Response to the MS [5], [7], [8]. Once the MS receives this response, the overall handoff process is completed. Fig. 1 illustrates the handoff process in WLANs.

Probe delays represent over 90% of overall L2 handoff latencies [9]. This fact motivated the development of an effective fast scanning scheme for mobile hosts roaming with ongoing real-time applications.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of fast scanning methods in WLANs. Section 3 describes the proposed fast scanning scheme. Basically, this proposal allows an MS to launch authentication upon receiving the first probe response on a scanned channel while the MS performs intra-subnet and inter-subnet handoff with ongoing voice over IP (VoIP) session with a Correspondent Node (CN). Section 4 covers performance evaluations that were conducted through simulations. The experimental results and analyses are also presented in detail. Finally, Section 5 concludes the paper with possible future work.

II. RELATED WORK

This section describes several typical fast scanning methods in recent literature, such as the Selective Scanning plus Caching methods, the Neighbor Graphs approach, the Synchronized Scanning approach and the MultiScan approach.

Selective Scanning plus Caching methods (also called Channel Mask schemes [10]) are designed to reduce MAC layer handoff latency to a level where VoIP communication becomes seamless [11]. Moreover, they focus on reducing the
probing time of non-existing channels via selective scanning, as well as the scanning frequency using caching techniques.

Selective Scanning method enables an MS to probe only a well-selected subset of available channels during its consecutive handoffs while the Caching method resorts to bypassing scanning using an AP cache table which is built through previous selective scanning. AP caching techniques enable MSs to reassociate with an AP without performing scanning during handoff. This results in enhanced handoff performance. However, as caching tables are built from previous scanning results, MSs are more likely to select an inappropriate AP during handoff, thus triggering false handoffs. Moreover, cache misses (no entry found in the cache) hinder network performance.

The Neighbor Graphs (NGs) approach is designed to reduce the total number of probed channels as well as the probe-waiting time on each channel [12]-[14]. The neighbor graphs and non-overlapping graphs are utilized to dynamically capture the mobility topology of wireless networks and to assist MSs in making decisions regarding whether a channel needs to be scanned and whether to wait further for probe responses on each examined channel before the MaxChannelTime expires. The drawback of this approach is that MSs must have a global knowledge of the wireless environments before actual handoff. Furthermore, the time required to build these graphs is rather lengthy and the performance of this NG approach can be degraded due to significant topological changes [12], e.g., adding new APs or removing existing APs, thus rendering the maintenance of such mobility graphs very difficult.

Synchronized Scanning (SyncScan) approach allows stations to perform passive scanning and to switch channels at the exact moment when a beacon frame is about to arrive [15]. It provides a technique to continuously track nearby APs by synchronizing short listening periods. In doing so, a staggered periodic schedule of beacon periods is spread across channels and created for all APs. For example, all APs operating on Channel 1 are forced to broadcast beacons at time $T$, APs on Channel 2 broadcast beacons at time $(T + d)$, and APs on Channel 3 broadcast beacons at time $(T + 2d)$, and so on. Hence, if an MS connected to an AP on Channel $c$ receives a beacon from Channel $c$ at time $T$, it can receive beacons from APs operating on Channel $(c + 1)$ at time $(T + d)$, and so forth.

Briefly, SyncScan lowers the cost of continuous scanning and yields better handoff decisions. However, time synchronization is a critical issue amongst all adjacent APs. In addition, to facilitate time synchronization, these adjacent APs should belong to a common network administration; this further leads to the scalability issue. On the other hand, multiple APs operating on the same channel attempt to generate and broadcast beacons simultaneously, thus bringing about beacon conflicts. Therefore, a greater number of collisions take place on wireless mediums. This side effect reduces the productivity on wireless links and system throughput. Moreover, more packets are lost while stations explore other channels.

The MultiScan approach exploits multiple radios on the MS side in order to eliminate handoff latency [16]. Two radio interfaces are implemented for each MS: one is used to communicate with the PAP, the other to scan channels. During handoff, the second interface performs reassociation with the NAP while the first one continues data transfer with the PAP. Once the layer-2 handoff is completed through the second interface, the primary and secondary interfaces switch roles [16]. The drawback is that performance may degrade significantly due to radio cross-interference, since a single device uses two simultaneous radio interfaces [16]. In addition, handoff between interfaces should be taken into further consideration.

III. THE PROPOSED FAST HANDOFF SCHEME

Our research objective is to provide fast handoff support for mobile hosts roaming with ongoing real-time applications in wireless LANs. Hence, the main research motivation consists of minimizing handoff delays and packet loss during handoff. The assumption of stations can completely skip the handoff detection phase using any triggers provided by the physical layer is confirmed through the experimental study conducted in [17]. The proposed fast scanning scheme is described as follows:

When the received signal strength is below a pre-defined threshold, the physical layer of a handoff MS sends a physical layer event notification (also called trigger) to the MS’s MAC layer. However, it is a challenging issue to define this threshold as this depends on the real-life practical circumstance and affected by the surrounding interferences.

Upon receiving a trigger from the physical layer, the MS with ongoing real-time communication with a CN, launches a fast scanning procedure by analyzing its currently associated channel. In our algorithm, a channel analyzer module is defined and designed to store the current associated channel information.

Then, the MS switches to the next channel, achieves wireless medium access control using normal channel access procedure, e.g. Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). The MS broadcasts a probe request on the examined channel and starts a probe timer. Then, it listens to the channel, waiting for probe responses sent by APs within radio range. If no response is received before the MinChannelTime is reached, the MS switches to next channel and performs an active scanning.

Once the first probe response is received, the MS immediately begins authentication with the AP that sent the response. This optimally minimizes the probe-waiting time on an examined channel. Upon successful authentication and reassociation, the MS completes the L2 handoff.

The advantage of the proposed approach is that probe delays can be reduced significantly, while only a subset of the allowed channels is scanned aside from the minimal probe-waiting time on each examined channel. As a result, handoff latencies and packet losses are reduced for mobile hosts roaming with real-time applications in progress.

This proposal applies to fast movement cases as well as those where MSs need to handoff as quickly as possible. In addition, it requires neither AP modifications (such as
SyncScan) nor pre-knowledge of the wireless network topology (such as the Neighbor Graphs approach and the Selective Scanning plus Caching schemes). Moreover, the addition of a second radio interface for each MS is unnecessary (unlike MultiScan). Furthermore, simulation results for handoff latencies and packet losses are obtained from the same test bed (unlike Selective Scanning plus Caching), guaranteeing the consistency and credibility of results. However, this proposal also includes certain limitations, such as the possibility that the mobile does not select the best AP at the moment of handoff.

IV. PERFORMANCE EVALUATION

To evaluate the performance of our proposal, simulations were conducted with SimulX [18], a C++ simulator developed at Louis-Pasteur University in France. Especially designed for IEEE 802.11 networks, SimulX also provides mobility support in IPv6 networks. The IEEE 802.11b Standard [20] with 14 channels, Mobile IPv6 protocol [19] and Selective Scanning plus Caching schemes were already implemented. Based on these codes, we implement the IEEE 802.11b Standard with 11 channels, The IEEE 802.11b Standard with Min (an MS only waits for MinChannelTime on each examined channel), the Neighbor Graphs approach and our proposal. Fig. 2 illustrates the simulation scenario.

The investigated scenario consists of a Mobile Node (MN) moving inside a building at an average speed of 1 m/s, communicating with a CN that sends UDP packets every 20 ms to emulate 64 kbps pulse code modulated voice stream packetized into 160 bytes. The radio range of each network entity (including MN, CN and AP) equals 12 m. The MinChannelTime is set to 17 ms and the MaxChannelTime is 38 ms corresponding to Cisco devices [9]. AP1 operates on the channel 1, AP2 on channel 6, AP3 on channel 11 and AP4 on channel 6. The default beacon interval for each AP is 100 ms. Five LANs with a 100 Mbps capacity are present, and four WLANs of which the transmission rate ranging from 2 to 11 Mbps.

The MN performs three movements: from AP1 to AP2, then to AP3, before returning to AP1. However, the following performance analysis is based on the simulation results of the last two movements as the MN performs a full scan for Selective Scanning plus Caching schemes and constructs neighbor graphs and non-overlapping graphs for the Neighbor Graphs approach during the first movement. Thus, the performance evaluation represents a fair comparison, as the first movement of the MN is excluded from the analysis.

Fig. 3 shows the relationship between probe delay and AP’s capacity (transmission rate). Our proposed scheme outperforms the other four handoff solutions: the IEEE 802.11b Standard, the IEEE 802.11b Standard with Min, the Selective Scanning plus Caching and the Neighbor Graphs approaches. This is because our proposed scheme enables an MS to quickly terminate the scanning procedure once it finds an available AP to associate with. The average probe delay of the proposed scheme equals 35.70 ms, compared to 210.20 ms for the IEEE 802.11b Standard, the performance gain is 83.02%; compared to 189.51 ms for the Standard with Min, the gain is 81.16%; compared to 55.51 ms for the Selective Scanning plus Caching, the performance gain is 35.69%; compared to 55.51 ms for the Neighbor Graphs, the performance gain is 35.69%. Note that in our simulated scenario, the Selective Scanning plus Caching methods almost have the same network performance as the Neighbor Graphs approach. We also find that the increasing of AP’s capability has little impact on the probe delay. Further study will be done to evaluate and confirm this observation.
Fig. 4 illustrates the relationship between L2 handoff latency and AP’s capacity. The increasing of AP’s capacity leads to shorter L2 handover latencies. This is because the time taken for exchanging frames between the MS and the involved APs becomes lower due to higher transmission rate of the AP. Our proposed scheme delivers better performance than the other four handoff schemes. This is because L2 handoff latency comprises probe delay, authentication delay and reassociation delay, and our proposed scheme requires less probe delay than other approaches. The average L2 handover delay of our proposal equals 38.92 ms, compared to 213.34 ms for the IEEE 802.11b Standard, the performance gain is 81.76%; compared to 192.41 ms for the Standard with Min, the gain is 79.77%; compared to 58.46 ms for the Selective Scanning plus Caching, the gain is 33.42%; compared to 58.38 ms for the Neighbor Graphs approach, the optimization is 33.33%.

Fig. 5 shows packet loss rate versus AP’s capacity. Packet loss rate is defined as a ratio of the number of lost packets over the total number of transmitted packets at the application layer. Again, our proposed solution yields better performance than other schemes. This is because longer handoff latency leads to more packet loss. The average packet loss rate for the proposed scheme equals 1.39%, compared to 2.55% for the IEEE 802.11b Standard, the gain is 45.49%; compared to 2.29% for the Standard with Min, the gain is 39.30%; compared to 1.94% for the Selective Scanning plus Caching schemes, the gain is 28.35%; compared to 1.72% for the Neighbor Graphs approach, the performance gain is 19.19%. To maintain VoIP quality, the packet loss rate should be at or below 3% [21], thus our proposed fast scanning solution can meet this requirement.

Fig. 6 shows authentication delay versus AP’s capacity. Authentication delay decreases rapidly as AP’s capacity increases. Neighbor Graphs has better performance amongst all solutions. The average authentication delay of the proposed scheme is 1.57 ms, compared to 1.34 ms for the IEEE 802.11b Standard, the performance decrease is 0.23 ms; compared to 1.26 ms for the Standard with Min, the decrease is 0.21 ms; compared to 1.32 ms for selective scanning and caching, the decrease is 0.25 ms; compared to 1.24 ms for neighbor graphs, the decrease is 0.33 ms. All the differences are less than 0.4 ms. This is because the processing time for executing our proposed scheme is a little bit longer than other approaches as our MN needs to quit the channel analyzer module before launching authentication, this makes authentication delay sounds a little bit longer than other solutions. We also find that authentication delay is quite shorter than probe delay for all the schemes.

Fig. 7 shows reassociation delay versus AP’s capacity. From the figure, we find that reassociation delay decreases rapidly as AP’s capacity increases for both the IEEE 802.11b Standard with Min and the proposed scheme. The Selective Scanning plus Caching schemes have better performance amongst all the handoff solutions. The average reassociation delay of the proposed scheme is 1.65 ms, compared to 1.80 ms for the Standard IEEE 802.11b, the performance gain is 8.29%; compared to 1.65 ms for the Standard with Min scheme, no gain no loss; compared to 1.63 ms for the Selective Scanning plus Caching, the performance decrease is 0.02 ms; compared to 1.75 ms for the Neighbor Graphs approach, the performance gain is 5.61%. Additionally, reassociation delay is quite shorter than probe delay for all the schemes.
V. CONCLUSION

This paper proposes a fast L2 handoff scheme to enhance the performance of mobile nodes roaming in WLANs with ongoing real-time applications. Our proposal allows mobile stations to actively scan only a subset of all accessible channels without pre-knowledge of the wireless environment and it also decreases the probe-waiting time to an optimal minimum on each examined channel. As a result, handoff latency is reduced significantly, making the support of real-time ongoing services in WLANs possible.

Simulations results show that our proposal delivers better performance than the IEEE 802.11b Standard, the Standard with \( \text{MinChannelTime} \), the Selective Scanning plus Caching schemes and the Neighbor Graphs approach. As the average L2 handoff latency of the proposed fast handoff scheme is about 39 ms, the support of multimedia applications such as VoIP for mobile hosts roaming in WLANs seems promising.

In the near future, large-scale simulations will be conducted for performance analyses and novel effective mobility management schemes will be proposed in which cross-layer design will be taken into account.

ACKNOWLEDGMENT
The authors would like to thank Julien Lee Kien On, an intern student from the Institut National des Sciences Appliquées (INSA) de Toulouse, for his assistance in the implementation of our proposal.

REFERENCES