SECURE FAST HANDOFF IN IEEE 802.11-BASED WIRELESS MESH NETWORKS

GANG YAO

PhD

The Hong Kong Polytechnic University

2018

The Hong Kong Polytechnic University

Department of Computing

Secure Fast Handoff in IEEE 802.11-based Wireless Mesh Networks

Gang Yao

A thesis submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

November 2016

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____________________ (Signature)

______Gang Yao_______ (Name of Student)

Abstract

IEEE 802.11-based Wireless Mesh Networks (WMNs) have become a de facto alternative network solution for the widespread and rapid deployments of wireless metropolitan area networks in recent decade, in addition to traditional mobile cellular networks e.g. 3G/4G. It has drawn a huge amount of attention in both research and industry communities to explore its pros and cons and improve its performance. We embrace this future and envision that WMNs will keep up evolving and finally dominate the mainstream wireless wide area networks rather than 5G networks to serve mobile applications on pervasive smartphones in coming future, similar to what had happened between IP and ATM technology in the field of wired networks in late 90s in last century.

802.11-based WMN serves the purpose to provide pervasive mobile Internet access and intranet communications in enormous and sophisticated application scenarios. Despite time-insensitive applications such as web browsing and email, 802.11-based WMN faces a major challenge which is how to fully support time-sensitive mobile applications, e.g. streaming media, Voice over IP (VoIP), video conference, Internet of Things (IoT), and etc. These aforementioned mobile applications all require small delay and low packet loss encountered when there are interruptions of communication which are mainly caused by handoff procedures. Particularly, the challenge aggravates when more and more frequent handoffs occurred in a WMN with dense APs in real-world deployment. For an instance, when a Mobile Client (MC) moves randomly, it has to constantly change its point of association, e.g. Access Point (AP) or mesh router, due to the proximity of small coverage of the radio signal. The performance of streaming media demands such a delay

to be less than 300ms while VoIP demands that less than 50ms, specified by International Telecommunication Union (ITU) standard [6]. When it comes to the newly emerged applications of IoT in recent years, the delay of packets between sensors and actuators and the back-end server needs to below 5ms. Nevertheless, the current legacy 802.11-based WMN handoff procedures could cause an end-to-end delay exceeding 1 second [2] so that it cannot guarantee such Quality of Service (QoS) at all.

In order to respond to the aforementioned challenge, we explore the problem domain to understand the mechanisms behind and aim to improve the mechanisms of handoff procedures on multiple levels so as to support QoS of mobile applications. In general, it mainly covers Layer 2 (L2 i.e. Data link Layer) handoff and Layer 3 (L3, i.e. IP Layer) handoff respectively, and occasionally even involves higher layer. The L2 handoff deals with issues of probing for the most suitable AP, authentication of MC, and association with AP. The L3 handoff involves issues of IP acquaintance, IP mobility, authentication and security of IP connectivity. In this work, we address the problem of Secure Fast Handoff (SFH) which is to tackle the delay and packet loss issues while maintaining secure communication on L2 and L3 as well.

In our research, we systematically investigate the related issues, generalize problems, and propose novel and effective mechanisms to solve the aforementioned challenge. We propose a theoretic framework to facilitate fast and secure handoff for all time-sensitive applications in 802.11-based WMNs on the basis of procedure parameter optimization, network-based proactive AP-probing schemes, and improved authentication protocols. Moreover, the mechanisms are implemented in a real-world testbed, evaluated and developed to improve the performance for mobile applications. The contributions of us are summarized as follows:

First, we address the fast handoff problem with an empirical study on an 802.11-based WMN testbed HAWK (Heterogeneous Advanced Wireless Mesh Networks), which incorporates our reactive L2 and L3 fast handoff schemes: Background Selective Channel Scanning and Location Management-based Routing Update, on top of off-the-shelf hardware. It is a tradeoff between optimal performance and low practical cost so that it can apply to real-world scenarios as much as possible. A series of field tests are conducted to investigate and fine-tune key parameters of the above reactive schemes in order to evaluate the performance.

Second and mainly, to address secure fast handoff problem in an 802.11-based WMN, a novel total solution comprised of Network-assisted Radio Signature (NRS) and Dual Re-Authentication (DRA) has been proposed. In particular, the NRS scheme is proposed to proactively obtain neighboring AP knowledge by measuring and profiling the radio characteristics and to determine the most suitable AP to associate prior to actual handoff. The Dual Re-authentication mechanism is proposed to enable fast handoff by granting the MC an immediate access based on a lightweight authentication as long as the associate AP is determined, while a strong authentication is executed within a period of timeslot. The proposed solution is optimized in terms of fast handoff at cost of a prior training stage in the deployment of the network and actualized authentication to secure communication afterwards.

Furthermore, to address the handoff problem in an AP-dense 802.11-based WMN environment which has been more prominent, we advance our proposed technique NRS and further develop Temporal-NRS (T-NRS) scheme, which leverages historical knowledge of APs associated in time series to assist in handoff decision in addition to NRS technique based on spatial knowledge. The enhanced scheme improves the performance whilst greatly eliminates the inflexibility of the original approach.

At last, to continue addressing the handoff problem in an AP-dense WMN environment, a novel handoff scheme called OppoScan (Opportunistic Scanning) supported by virtual radio is proposed. OppoScan opportunistically leverages nearby MCs and APs to produce the required information of neighboring AP for handoff, thus significantly decrease the number of switching channel of APs. Our evaluation based on experiments indicates that OppoScan can efficiently achieve low delay while maintaining handoff in more practical scenarios for 802.11-based WMN.

Publications

Journal Papers

1. G. Yao, J.N. Cao, Y. Yan, Y.S. Ji, Secured Fast Handoff in 802.11-based Wireless Mesh Networks for Pervasive Internet Access, IEICE Transactions on Information and Systems. vol.E93-D, no.3, pp. 411-420, Mar. 2010.

2. G. Yao, J.N. Cao, X.F. Liu, J. Siebert, Enabling Fast Handoff in Dense 802.11 WMNs via Temporal-spatial Knowledge and Opportunistic Probing. Journal of Reliable Intelligent Environments. 10.1007/s40860-018-0057-2. Feb. 2018

Conference Papers

1. G. Yao, J.N. Cao, X.F. Liu, et al. OppoScan: Enabling Fast Handoff in Dense 802.11 WMNs via Opportunistic Probing with Virtual Radio, IEEE 14th Mobile Ad Hoc and Sensor Systems (MASS), pp. 198-205, Oct. 2017.

2. G. Yao, J.N. Cao, X.F. Liu, et al. Fast Handoff based on Enhancement of Network-assisted Radio Signature in 802.11 Dense WMNs, ACM International Conference on Distributed Computing and Networking (ICDCN 2018), Jan. 2018.

3. C.S. Zhang, J.N. Cao, G. Yao, CoDA: Connectivity-Oriented Data Dissemination Algorithm for Vehicular Internet Access Networks (MSN 2015). pp.186-193, Nov. 2015.

4. J.N. Cao, K. Xie, W.G Wu, G. Yao, W. Feng, et al. HAWK: Real-world Implementation of High-performance Heterogeneous Wireless Network for Internet Access, The 1st ICDCS International Workshop on Next Generation Network Architectures (NGNA2009). pp. 214-220, June 2009.

5. J.N. Cao, C.S. Zhang, J. Zhang, G. Yao, et al., SHAWK: Platform for Secure Integration of Heterogeneous Advanced Wireless Networks. AINA Workshops 2012.

Patents

1. Gang Yao, Jiannong Cao, Chisheng Zhang, Chuda Liu. A Scheduling Method for Fast Handoff Using Dual Radio Interfaces. C.N. Patent. CN101873673B. Aug, 15, 2012.

2. Gang Yao, Jiannong Cao, Chisheng Zhang. A Smart Handoff Decision Method and System Using Fingerprint Based on Fuzzy Logic. C.N. Patent. CN101998381B. May 2013.

3 Gang Yao, Chisheng Zhang. A Method, System and Client for Message Dissemination Based on 802.11 Wireless Network SSID. C.N. Patent. 102413429A, April 2012.

4 Gang Yao, Chisheng Zhang, Libin Yang. A Method and System for Message Transfer and Distribution. C.N. Patent. 104168297A, Nov. 2014.

Acknowledgements

I would like to express my deepest gratitude to my supervisor Prof. Jian-nong Cao for his continuous encouragement, support, guidance, and everything. His enthusiasm on science and technology, his vision on academic research, his broad knowledge, and his kindness on his students have always inspired me, not only on my research but also on my personality and life. Without his help and support, this body of work would not have been possible and I would never be able to go this far.

In addition to that, I thank Ronnie Cheung and Dr. Alvin Chan for their support and help during my graduate studies at PolyU. I also thank for the patience and insightful comments to members of my committee Prof. Zili Shao, Prof. Yi Pan and Prof. Xiaohua Jia. I thank all other faculty members who have been my teachers and mentors over these years.

I would like to express my thanks to Xuefeng Liu, Ye Yan, Jiaxing Shen, Xin Xiao, Chisheng Zhang and Siunam Cheung for their collaboration in my research for their constructive suggestions and a lot of support.

At last, I would like to give my special thanks to my wife, mother and father for their unconditional love and support for so many years. I owe them so much.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

1.1.1 Rise of 802.11-based Wireless Networks

The convergence of traditional mobile cellular communication systems and the emerging wireless access networks has been a prevailing phenomenon in recent decades with the evolution and proliferation of Wireless Local Area Network (WLAN) and Wireless Metropolitan Area Network (WMAN) technologies. The popularity of WLAN networks is indicated by Wi-Fi hotspots being deployed worldwide and approximately 3 billion Wi-Fi enabled devices being sold as of second quarter, 2016 [36]. IEEE 802.11 standard is one of the most dominant technologies for wireless access networks. Portable 802.11 enabled devices, such as smartphones, notebooks and personal multimedia devices are becoming increasingly popular. As many mobile applications for 802.11 enabled devices require higher Quality of Service (QoS) and better mobility support, there is an increasing demand for improving the performance of 802.11-based networks. Deployment of multiple APs (Access Points) is an effective way to improve the capacity and performance of 802.11-based networks. Nowadays, more and more large-scale 802.11-based networks with multiple APs have been deployed in public venues and even city-wide in order to provide broadband wireless networking services to a larger area.

One of important goals of deploying such networks is to support seamless services roaming when mobile users are on the go. For example, a user making VoIP (Voice over

IP) calls from an 802.11 enabled smartphone should not experience service disruption or quality degradation to the calls when the user moves around in the network. In practice, this is a challenging task mainly due to the long delay incurred in the 802.11 handoff procedures, during which the data services are disrupted.

Previously, WLAN is used as wireless access network to provide Internet access. The major applications are data services. The popularity of 802.11b/a/g/n WLAN as the last-mile wireless access networks for Internet, instead of its original role of data communication for local area networks, has opened a new window of academic and industrial researchers. As a wireless access network for Internet, all kinds of broadly-used data services and applications are supposed to be enabled and provided in such an 802.11 WLAN environment. More than that, in the aforementioned merged heterogeneous wireless network scenario which lays the foundation for the next-generation mobile communication system in the coming future, the voice applications supported imposes more critical challenges upon the 802.11 WLAN, which requires the same QoS for end users as it does in traditional mobile cellular communication world.

1.1.2 Applications Requirements upon Wireless Networks

Traditionally, 802.11 WLAN complements conventional mobile cellular networks by providing higher bandwidth with cheaper even totally zero expenditure only in small coverage. For example, the present cellular networks, e.g. 4G network offers potential throughputs of 20Mbps and 802.11a/g/n WLAN offers potential throughputs of around 300~500Mbps. In the meanwhile, the cellular mobile networks dominate the market of access networks with much larger coverage, usually nation-wise, and the commercial services of carrier-grade QoS.

802.11 standards can potentially offer higher bandwidths than cellular networks. However, qualitative comparisons of 802.11-based networks and current 4G networks are unjust because 4G networks are partially circuit-switched whereas 802.11-based networks are packet-switched. The profound difference of 802.11 WLAN compared with a cellular network is its shared medium access and hence bandwidths and delays are influenced by surrounding mobile users. 4G networks provide a guaranteed connection and delays and bandwidth are only affected by the distance to the cell tower. Subsequently, the application scenarios have been radically changing since 802.11 WLAN becomes more popular, covers larger areas, and connects with other 802.11 WLAN. How to achieve similar application QoS, particularly VoIP and IoT, in the connected 802.11 WLANs as that in cellular networks imposes the largest challenge which lie in front of 802.11-based WLAN.

It should be noted that nowadays we practically work on 802.11-based Wireless Mesh Networks (WMNs) so that 802.11-based WLAN and 802.11-based WMN are exchangeable and alternatively used hereafter in this thesis. The referred WMN is based on 802.11 WLAN whereas further supported by IP Layer multiple relay communication via wired or wireless backbone connections to extend the coverage of single WLAN.

IEEE 802.11-based WMN has been a promising candidate for rapid and widespread deployment of WMAN in recent years. It merges multiple WLANs and serves as a community network, campus network, or corporate network to offer pervasive Internet access and intra-network communications in larger area. Despite the time-insensitive applications, e.g. web browsing, email, the biggest challenge it faces is to enable time-sensitive applications such as streaming media and VoIP which requires low delay and packet loss during the interruption of communication.

The most common interruption of communication in 802.11-based WMN occurs in the event of handoff, which is that a MC moves and changes its point of attachment, e.g. Access Point (AP), from one to the other in the network. Under this circumstance, streaming media applications demand a delay to be less than 300ms while VoIP demands that less than 50ms, specified by the long accredited International Telecommunication Union (ITU) standard [6]. When it comes to the newly emerged applications of IoT in recent years, the delay of packets between sensors and actuators and the back-end server needs to below 5ms. Nevertheless, the current legacy 802.11-based WMN handoff procedures could cause an end-to-end delay exceeding 1 second [2] so that it cannot guarantee such QoS at all. During the handoff process, MC is unable to send or receive data. If the handoff takes longer, packets could be missed and the session could be dropped.

In addition to that, the communication needs to be secured so that authentication mechanisms are enabled during the handoff. The authentication procedures which involve many rounds of handshaking and exchange of authentication packets will introduce more interruptions of communication. Obviously, the longer interruption caused by a secured handoff will further deteriorate the QoS performance.

Therefore in this research, the key issue to achieve desirable QoS of low delay and packet loss is about to reduce the communication interruptions introduced by handoff procedures.

1.2 Problem Domain and Challenges

1.2.1Handoff Procedures in 802.11-based WMN

The major difficulty of communications in WMN is the issue of mobility. Although wireless networks were designed to support roaming, the early design considerations were that MCs would be stationary whilst using the network, otherwise it was known as ‘nomadic’ operation. Advances in miniaturization of portable devices have led to the introduction of 802.11 enabled handheld’s such as PDA’s and smartphones. These devices demand fast, seamless roaming and mobile operation when they are on the move.

Current research in mobility can be divided into Layer 2 (Data link Layer, L2), Layer 3 (IP Layer, L3), inter-domain mobility and cell mobility. Layer 2 mobility is the capability to move between APs within the same subnet or routed domain. Layer 3 mobility is the capability to move between APs located in different subnets, introducing IP addressing issues. Inter-domain mobility occurs between differently governed networks such as a free hotspot and a carrier’s wireless network. Cellular mobility occurs when a dual band device roams outside the WMN network and a cell call must be made to maintain the connection. These handoffs are shown as the below in Figure 1.

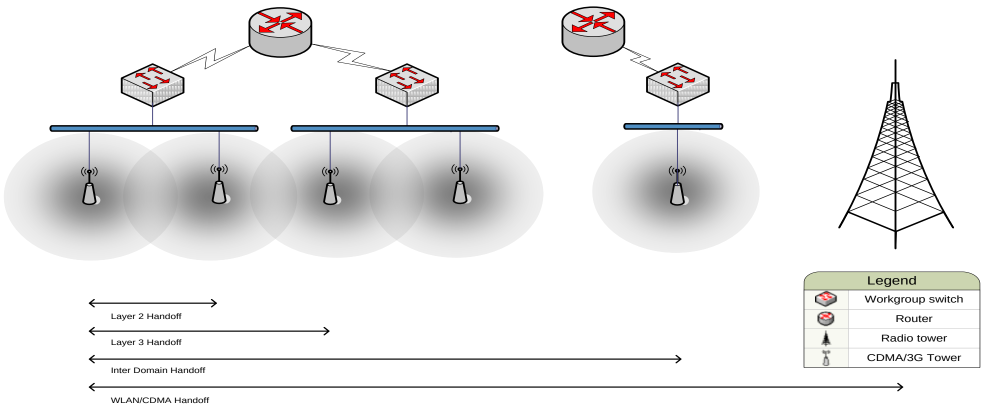

Figure 1 Handoff problems on different level of 802.11-based WMN / Cellular Network

In the case of Layer 2, Layer 3 and inter-domain handoff, the delays are cumulative. For example, inter-domain mobility will incur the delays of Layer 2 and Layer 3 mobility. However, the scope of this research is confined to Layer 2 as well as Layer 3 mobility in homogeneous networks, often referred to as handoff. In 802.11-based networks, a handoff occurs when a mobile client moves out from the radio coverage area of its actual serving AP to a new target AP.

A complete process of handoff in 802.11-based WMN covers aspects of establishing wireless link and maintaining IP connectivity which are respectively referred to Layer 2 handoff and Layer 3 handoff. Issues in handoff process are shown in Figure 2. The L2 handoff involves issues of probing for the most suitable AP, authentication of the MC and association with the AP, while L3 handoff involves issues of IP acquaintance and location management.

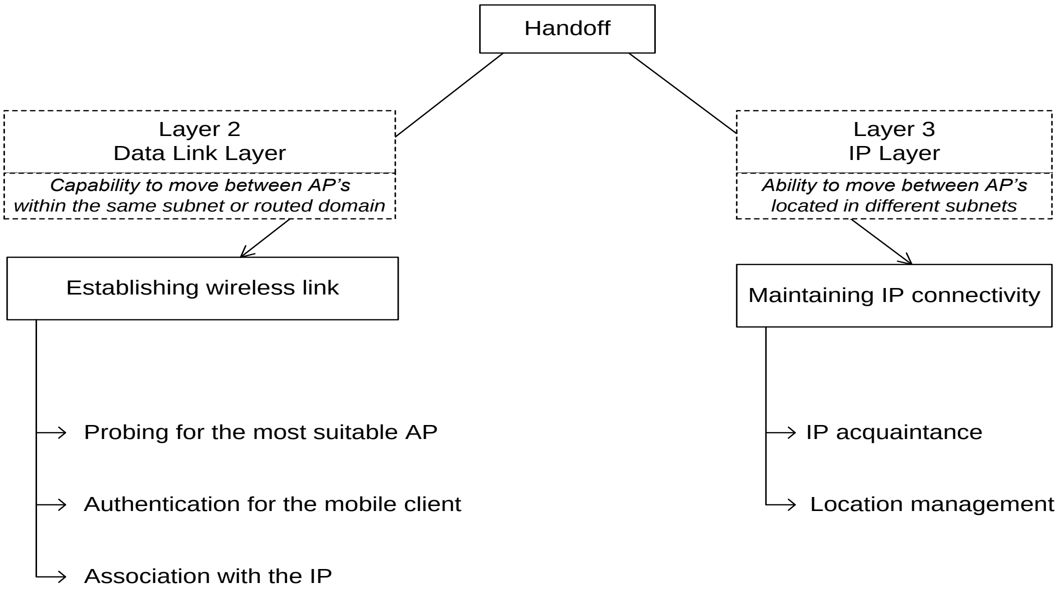
Handoff
Layer 2
Data Link Layer
Capability to move between AP's within the same subnet or routed domain
Layer 3
IP Layer
Ability to move between AP's located in different subnets
Establishing wireless link
Maintaining IP connectivity
Probing for the most suitable AP
Authentication for the mobile client
Association with the IP
IP acquaintance
Location management

Figure 2 Handoff issues in details in 802.11-based WMN

1.2.2 Challenges of Secure Fast Handoff

As illustrated in Figure 3 below, the handoff issue involves interruptions of multiple dimensional operations of the network from physical layer L1 on the bottom to application layer L4 on the top. The standard 802.11 handoff operations on L2 are comprised of three procedures: namely scanning, authentication and re-association. The scanning is that a MC collects information about neighboring APs and chooses the most appropriate AP to handoff. The authentication procedure is to authenticate and authorize the access to the AP. The association procedure is to establish the association with AP.

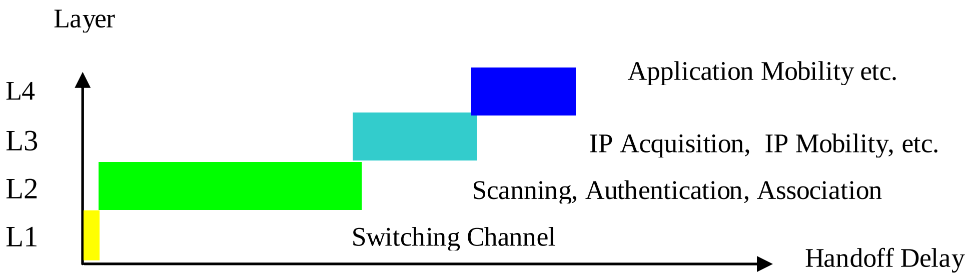
Layer
L4
L3
L2
L1
Application Mobility etc.
IP Acquisition, IP Mobility, etc.
Scanning, Authentication, Association
Switching Channel
Handoff Delay

Figure 3 Handoff delay accumulated in 802.11-based WMN

Therefore Layer 2 handoff delays are composed of scanning, re-authentication and re-association delays. The fundamentals of this research aim to reduce such delays. We mainly address on Layer 2 handoff issue in this thesis due to it contributes the majority of the delay.

Scanning delay: The scanning delay is the time taken to scan every channel for APs. It involves determining the characteristics of channels of surrounding APs. There are two scanning mechanisms are defined, namely passive and active scanning. Using default mechanisms in 802.11 networks, the scanning delays vary between 80 and 1100ms [30] [34]. Scanning delays are the focus of this research and are subsequently discussed in greater details**.**

Re-authentication delay**:** Re-authentication delay is the time taken to securely provide credentials to the next AP in order to authenticate and authorize the access to the new AP. Re-association delay: Since association procedure is to establish the association with AP in the first time, the re-association procedure is to associate with the next AP after handoff.

An end-to-end delay caused by the complete handoff process in a standard 802.11-based WMN could be accumulated and exceed more than several seconds, which results in an unacceptable degraded quality of time-sensitive applications, e.g. VoIP, location management, etc. The ITU specifies that total end-to-end delay for VoIP must not exceed 50ms. It should be noted it is not an arbitrary objective but provides a target to aim for.

1.2.3 Objectives of Secure Fast Handoff

Since the communication of low delay and packet loss is referred as seamless communication, a handoff which facilitates seamless communication is called as seamless handoff. Accordingly, fast handoff deals with low delay issues and smooth handoff deals with low packet loss. The major efforts of the research community are initially focused on the fast handoff because there are more severe constraints on solving this problem.

In practice, this is a challenging task due mainly to the long delay incurred in the 802.11 handoff procedures, during which the data services are disrupted. The time-sensitive applications such as streaming multimedia demand seamless and continuous network connectivity whose delay are less than 150ms. A mobile user making VoIP calls from an 802.11 enabled smartphone should not experience service disruption greater than 50ms otherwise the call would be dropped. The delay for such applications as IoT typically ranges from 5ms to around 50ms. Network connectivity is lost or the session severely degraded if the MC fails to associate with a new AP within the above mentioned time constraints as it moves out from the vicinity of the currently associated AP. Since application layer performance is what ultimately matters, any successful handoff mechanism must be able to honor such tight timing constraints.

1.3 Main Contributions of the Thesis

In our research, we systematically investigate the related issues, generalize problems, and propose novel and effective mechanisms to solve the aforementioned challenge and achieve the above objectives. We have proposed a theoretic framework to facilitate fast and secure handoff for all time-sensitive applications in 802.11-based WMNs on the basis of procedure parameter optimization, network-based proactive AP-probing schemes, and improved authentication protocols. Moreover, the mechanisms are implemented in a real-world testbed, evaluated and developed to improve the performance of the network for mobile applications.

As in the below Figure 4, we classify our proposed solutions and define a framework to carry out our studies on secure fast handoff mechanisms. The details of our work are summarized as follows:

Framework to facilitate secure fast handoff in 802.11-based WMNs					
Fast Handoff				Fast Authentication	
Terminal-based	Network-assisted			HAWK: Pre-authentication by Re-using Credentials	Dual Re-Authentication by Using Light Authentication Prior to Strong Authentication
HAWK Tesbed: Selective Probing and L2&L3 Cross-layer Optimization	NRS: Proactive Probing by Utilizing Radio Signatures	T-NRS: Proactive Probing by Utilizing Temporal Radio Signatures	OppoScan: Opportunistic Probing by Utilizing Virtual Radios		

Figure 4 Framework of proposed secure fast handoff approaches

First, we address the fast handoff problem with an empirical study on an 802.11-based WMN testbed HAWK (Heterogeneous Advanced Wireless Mesh Networks), which incorporates our reactive L2 and L3 fast handoff schemes: Background Selective Channel Scanning and Location Management-based Routing Update, on top of off-the-shelf hardware. It is a tradeoff between optimal performance and low practical cost so that it can apply to real-world scenarios as much as possible. A series of field tests are conducted to investigate and fine-tune key parameters of the above reactive schemes in order to evaluate the performance.

Second and mainly, to address secure fast handoff problem in an 802.11-based WMN, a novel solution comprised of Network-assisted Radio Signature (NRS) and Dual Re-

Authentication (DRA) has been proposed. In particular, the NRS scheme is proposed to proactively obtain neighboring AP knowledge by measuring and profiling the radio characteristics and to determine the most suitable AP to associate prior to actual handoff. The Dual Re-authentication mechanism is proposed to grant the MC an immediate access based on a lightweight authentication as long as the associate AP is determined, while a strong authentication is executed within a period of timeslot. The proposed solution is optimal in terms of QoS performance at a cost of prior training stage in the deployment and a cost of actualize authentication to secure communication afterwards.

Furthermore, to address the handoff problem in an AP-dense WMN environment and forward the above technique Network-assisted Radio Signature, we further propose Temporal-NRS (T-NRS) scheme, which leverages historical knowledge of APs associated in time series to assist in handoff decision in addition to NRS technique based on spatial knowledge. The enhanced scheme improves the performance whilst greatly eliminates the inflexibility of the original approach.

At last, to continue addressing the handoff problem in an AP-dense WMN environment, a collaborative scheme OppoScan (Opportunistic Scanning) supported by virtual radio is proposed. OppoScan opportunistically leverages nearby MCs and APs to produce the required information of neighboring AP for handoff, thus significantly decrease the number of switching channel of APs. Our evaluation based on experiments indicates that OppoScan can efficiently achieve low delay while maintaining handoff in more practical scenarios for 802.11-based WMN.

1.4 Outline of the Thesis

The structure of the thesis is shown as follows. Chapter 1 is the introduction to the thesis. Chapter 2 introduces the background knowledge of secure handoff issues in 802.11-based WMN and reviews related works in the literature.

The main body of the thesis is divided into three parts from Chapter 3 to Chapter 6. Chapter 3 proposes an empirical approach of Selective Channel Scanning and Location Management-based Routing Update to improve handoff delay with a tradeoff of performance in a real-world scenario, which is implemented on our testbed HAWK.

Chapter 4 presents a total solution of secure fast handoff of a proactive AP-probing scheme, namely Network-assisted Radio Signature (NRS) and a proactive authentication approach in order to achieve optimal performance at certain practical cost.

Chapter 5 addresses the handoff problem in an AP-dense WMN environment by forwarding the above technique Network-assisted Radio Signature to an enhancement of Temporal-NRS (T-NRS) scheme, which leverages historical knowledge of APs associated in time series to assist in handoff decision in addition to NRS technique based on spatial knowledge.

Chapter 6 further presents an opportunistic cooperative probing scheme OppoScan, by utilizing the neighboring APs and MCs' resources to achieve better performance.

Finally, Chapter 7 concludes the thesis and points out our future research directions.

Chapter 2 Background and Literature Review

2.1 General Background of Legacy Handoff

2.1.1 802.11-based Wireless Networks

IEEE 802.11 series protocols were standardized in the late 1990s. Since ratified in 1999 and became the first Physical Layer standard of WLAN, it takes no more than 10 years to witness that IEEE802.11b wireless network experienced exponential growth to provide cheap and pervasive Internet access in 2000's. Especially after the industry alliance Wireless-Fidelity (Wi-Fi) steps in, the Wi-Fi hotspots are now found everywhere including homes, workplaces, governmental buildings, universities, airports and central businesses districts.

It is noticeable that different regulatory bodies in different countries may have set certain principle-insignificant different rulings concerning spectrum allocation, e.g. European versions of 802.11, security mechanism, e.g. China's WAPI, etc, currently our research in this research only focuses on the typical IEEE standardized specifications.

Such a merged heterogeneous wireless network lays the foundation for the next-generation mobile communication system of the coming future. The result has to be a network ensuring seamless handovers from one technology to another while providing a continuous service to mobile users.

Meanwhile, it imposes a critical challenge upon the wireless access networks which is to provide the same quality services and applications for end users as in traditional cellular communication systems. Wireless networks, such as IEEE 802.11, have been designed to offer wireless connectivity with a moderate mobility. To be efficiently integrated, in the same global network with cellular networks, WLANs must ensure a better mobility support.

2.1.2 Overview of Layer 2 Handoff

The Layer 2 handoff procedure in 802.11-based WMN introduces different delays at multiple stages. As shown below in Figure 5, this section of the thesis outlines the key details regarding the 802.11a/b/g standards and how they impact the handoff. A common assumption about the handoff delay is that the wireless radio interfaces currently equipped with MC can only receive and send packets on one channel at a given time. A MC is unable to communicate with any AP when the handoff operation is performing. It is just this characteristic of wireless interface of 802.11 that causes the delay while probing for potential new APs, authenticating with new AP and associating with the new AP.

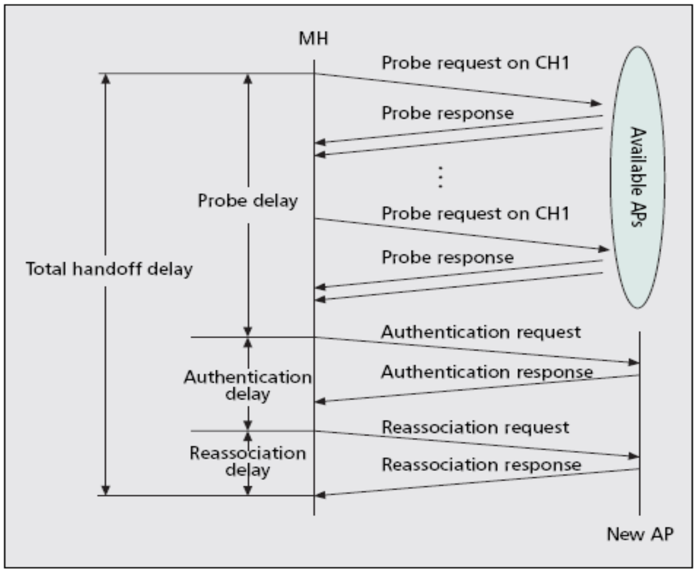

Figure 5 A complete procedure of Layer 2 handoff in 802.11-based WMN

2.1.3 Probing for APs

Probing is the mechanism which a MC performs to search for available new APs as the potential candidates in case that handoff occurs. It is the major source of handoff delay. The probing could be further divided into handoff initiation, probing and handoff decision which are elaborated here below.

2.1.3.1 Initiation of Probing

The initiation of probing decides how and when to initiate the probing process. Two kinds of initiation mechanisms exist, namely frame loss rate and signal strength:

i) The frame loss rate technique is straightforward. The handoff is initiated by counting unacknowledged frames. The number of unacknowledged frames before triggering the handoff is predefined by wireless radio interface, i.e. Network Interface Card (NIC), vendor. The frame loss rate probing initiation is not widely adopted even before 802.11 WLAN becomes popular. The mechanism implies that heavy frame loss has happened due to collision or interference on Physical Layer before initiating the probing, not mentioning triggering the handoff. The lost frames and the delay introduced by counting frame loss which is NIC vendor-specific both are intolerable to applications and user experience.

ii) Comparatively, signal strength technique does not result in the loss of frames prior to probing. Instead of that, signal strength technique initiates probing process before the signal deteriorates too much so at to lose frames. The initiation triggers in practice include Received Signal Strength Indicator (RSSI), Signal-to-Noise Ratio (SNR), Signal-to-Interference (SIR) and etc. Taking RSSI [2] as an example, it measures the signal strength of the associated AP. When the RSSI of a MC drops below a predefined threshold, the probing process is initiated. The RSSI threshold value is configurable and predefined before the deployment of APs. For dense AP deployment the threshold will higher whereas more sporadic AP deployment uses lower thresholds.

2.1.3.2 Classification of Probing

Probing is the process of determining the potential available APs in case that handoff occurs. When the probing process in undergoing, the MC cannot send or receive frames. The original IEEE 802.11 standards present two optional probing mechanisms: namely passive probing and active probing which allows flexibility for lower power consumption or faster probing.

Passive Probing

The protocol of passive probing is straightforward. The passive probing requires an AP to broadcast specific beacon frames to announce its presence. And a MC is required to continuously listen for beacon frames that APs broadcast. The MC has no prior knowledge about the last beacon transmission and consequently has to scan for the period of the beacon interval.

Given that the default beacon interval is usually 100ms and a MC has to scan each channel, the total probing delays will vary based on the number of channels. In general, 801.11b/g/n takes 300ms to do probing for there are 3 non-overlapping channels while 802.11a takes 800ms to do probing for there are 8 non-overlapping channels. The numeric results are in line with the empirical study result carried out in [2].

Although the passive probing is unsuitable for time-sensitive voice applications, it requires minimal power consumption and bandwidth usage comparing with active probing.

Active Probing

The protocol of active probing is a bit complicated than passive probing. The active probing process is a mechanism whereby a MC aggressively probes for APs. The MC sends out a probe request frame, APs receive the probe request and sends back a probe response. The MC retrieves the AP information in the probe response.

The duration on each channel for probing is determined by two variables, namely *MinChannelTime* and *MaxChannelTime*. The *MinChannelTime* variable is the minimum time in TU (Time Units, one TU is equal to 1024 µs.) to stay on each channel while probing. While the *MaxChannelTime* variable is defined as the maximum time in TU to stay on each channel while probing [2].

When the MC actively scans a channel, it broadcasts the probe request and begins the *MinChannelTime* timer. If no response is heard within the *MinChannelTime*, the MC assumes the channel empty and switch to the next channel. If response is received within the *MinChannelTime*, the MC awaits for the *MaxChannelTime* to collect as many as possible AP information as candidate new AP.

The active probing delay TA is bounded by the formula:

$$\mathrm{N} \times \mathit{MinChannelTime} \leq \mathrm{TA} \leq \mathrm{N} \times \mathit{MaxChannelTime}. \tag{2.1}$$

Many studies have attempted to research on the associated delays with active probing. Empirical studies have shown that it vary between 50 and 550ms [2]. As claimed previously, the probing process is responsible for majority of the total handoff delay. Probing delays are the focus of this research and are subsequently discussed in greater detail later.

2.1.4 Authentication

Basically the authentication delay is incurred by the exchange of the authentication frames. In general, two authentication approaches are widely accepted in 802.11 WLAN. One is the open system authentication, in which the AP always accepts a MC without real authentication procedure. Optionally, MAC address filtering can be employed with the open system authentication. The other is with real authentication process. The authentication methods could include Wired Equivalent Privacy (WEP), Wi-Fi Protected Access (WPA) and WPA2 etc., which requires that both the AP and MC implement the corresponding authentication protocols.

Taken WEP as an example, the authentication process takes four message exchanges as follows: i) The MC requests authentication to the AP by sending a challenge request message; ii) The AP sends a random number to the MC through a challenge-response message; iii) The MC signs this random number using WEP, which is a pre-shared secret key, and sends a response message back to the AP; iv) The AP verifies that the random number has been signed by the correct key, by calculating the signature itself and comparing the computed and the received values. Once the key has been verified, the AP authenticates the MC by sending an approval frame.

The authentication delay is proportional to the number of messages exchanged between the AP and MC. If an IEEE 802.11 network utilizes more complicated and enhanced authentication schemes e.g. IEEE 802.1X and EAP-TLS, more authentication delays are introduced. Therefore, the secure handoff with authentication becomes an even more challenging issue than fast handoff. In practice, simple WEP authentication introduces extra delay less than tens of milliseconds, while 802.1X authentication can consume around 1 second.

Regarding the enhancement to achieve secure but faster handoff, a prominent contribution has been made by IEEE 802.11r working group regarding authentication delay. The Pair-wise Master Key (PMK) caching can significantly reduce authentication delay by allowing PMK to be retained in memory. Using Neighbor Graph to proactively distribute PMK to the possible next AP, authentication delay can be reduced to around 20ms. This 20ms delay is a result of the 4-way handshake required to derive fresh PTK (Pairwise Transient Keys) from the PMK. The JIT-TAP (Just-In-Time Transition Acceleration Proposal) which became the basis for the 802.11r draft includes mechanisms to derive PTK’s prior to handoff. It is estimated that the JIT-TAP mechanism will further reduce authentication delays below 5ms.

2.1.5 Association/Re-association

The association is the process of associate with an AP within an Extended Service Set (ESS). An ESS is a set of one or more interconnected Basic Service Sets (BSSs), where a BSS is the service coverage of an AP.

The association delay is incurred due to the exchange of the association frames. Upon the successful completion of the authentication process, a MC sends an association request frame to the AP and receives an association response frame and completes the handoff. In addition, the next AP and the old AP may interact with each other to deliver frames related to the association. The delay caused by association is relative small. In practice, it is at the order of a few milliseconds.

In the future, the association delay may increase due to more packets to exchange when it comes to association. The Inter Access Point Protocol (IAPP) defines more complicated association and re-association process known as a fully ratified IEEE standard 802.11f. IAPP specifies the context information to be transferred between APs when a MC roams. Standard context transfer takes approximately less than 20ms. In the context of roaming, IAPP specifies the information or context to be transferred between APs when MC roams [18]. Standard context transfer takes approximately 15ms [26]. Similar to the 802.11i specification, mechanisms for proactive context transfer using neighbor graphs are an optional feature of the standard. Proactive context transfer reduces re-association delay from 15ms to 1.5ms [26].

Regarding reduce association delay, similar to the 802.11r specification, mechanisms of proactive context transfer using neighbor graphs are an effective and option for the standard. The proactive context transfer reduces association or re-association delay to the order of tens of milliseconds as well.

The mechanism to reduce association delay is very similar to the efforts on reducing authentication delay. In addition to that, context transferring related to association has not been mainly adopted rather than simple association. Therefore the research community usually does not inject special efforts on this topic.

2.1.6 Triggers of Handoff

Research on the triggers for handoff has been an attractive topic on handoff management for decades. In spite of conventional triggers on Physical Layer which includes the aforementioned RSSI, SNR and SIR, etc. to provide basic handoff initiation, the research community also extend the horizon of the problem further to other higher layer handoff triggers. For example, proposed by the 802.11e (QoS) working group, 802.11e mechanisms induce handoff when other thresholds such as load balance of APs are breached. As a matter of fact, multiple QoS performance indicators can be utilized as handoff triggers which include delay, jitter, etc. on lower layers, application QoS on MC side, and system-wise QoS, e.g. load balance, fairness. The further work on this topic does not fall into the scope of this thesis and will be reserved for continuing research.

2.1.7 Handoff Decision

The handoff decision phase occurs after the handoff is initiated and is the point at which the MC decides whether to stay connected to the current AP or associate to a new AP. An important aim of the research on handoff decision is to achieve association without ping-pong effect, which means the MC performs handoff back and forth between APs continuously. The classic solution to solve ping-pong effect is the heuristics algorithm as in the Figure 6. As discussed in the detection phase, MC begins scanning when the SNR drops below the *CellSearchThreshold* [24]. Only when the difference between the SNR of the old AP and a newly scanned AP is greater than the delta SNR, the MC will handoff to the new AP.

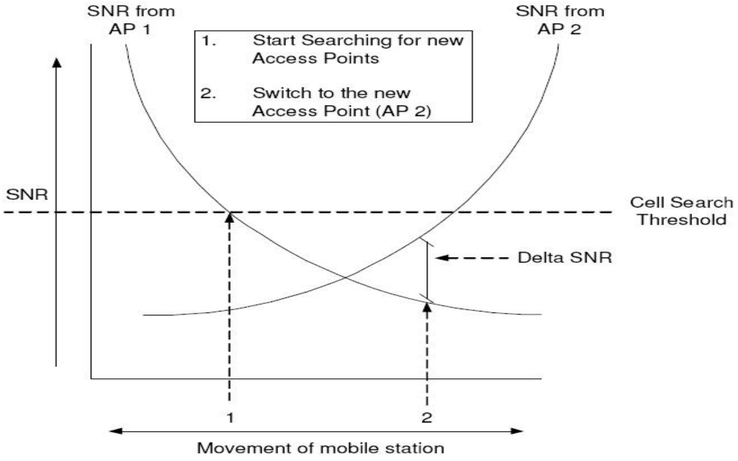

Figure 6 Handoff decision in 802.11-based WMN

The delta SNR specifies the minimum difference between the old AP and the new AP. It is important because SNR are never static hence the delta SNR prevents MCs from ping-pong between APs. Both the cell search threshold and the delta SNR are assigned based on the density of AP deployment.

2.2 Fast Handoff Mechanisms

2.2.1 Introduction

Although many studies back since 2004 have made solid progress in fast and secure handoff and shown that time-sensitive application QoS is achievable in 802.11-based wireless networks, there are less research activities or proposals of specific improvements on fast and secure handoff in 802.11-based WMN after 2007 until 2010. In this chapter, we aim to present a broad survey on the previous research progress as the basis and the starting point for our further enhancements.

As stated earlier, the probing phase makes the most significant contribution to the total handoff delays, we only focus on optimized mechanisms how to reduce probe delay despite association, handoff trigger and handoff decision, etc. in the following sections. Regarding optimization mechanisms on probing, it involves certain number of methods and techniques which are summarized as follows:

Reactive mechanisms

i) Optimistic probing channels and reducing channels to probe

ii) Reduce probing time on each channel

Proactive mechanisms

i) Neighbor graph-based technique

ii) Proactive probing by time-division

iii) Fixed time slot, e.g. SyncScan

iv) Dynamic time division, e.g. ProactiveScan

v) Proactive probing by dual radios, e.g. MultiScan

2.2.2Reactive Mechanisms of Fast Handoff

2.2.2.1 Optimistic Channel Probing

Optimistic channel probing [5] is an intuitive method which is to probe selected channels in terms of number of channels and sequence of channel. The MC can use optimistic probing in combination with passive or active probing. NIC vendors commonly implement optimistic probing in 802.11a/b/g/n mobile clients.

Instead of probing all 11 channels in 802.11 b/g/n or all 32 channels in 802.11a, the MC merely scans the non-overlapping channels, namely 1, 6 and 11, or 1, 5, 9…etc.

In case overlapping channels are also necessarily scanned, the above channels, or certain specific channels according to onsite measurements, will be set top priority to scan. The optimistic probing has the potential to greatly reduce probing delays while it is a not standardized method adopted by all 802.11 mobile clients.

2.2.2.2 Beacon Optimization for Passive Probing

While passive probing under default conditions was shown to be unacceptable for time-sensitive communication, modifications can improve the time delay. The passive probing has the advantage of low power consumption as stated previously. This is because messages are not transmitted by the MC but instead passively received while receiving packets uses comparatively minimal power. Passive probing delays are roughly equal to the number of channels multiplied by the beacon interval. By default, APs transmit beacons every 100ms. Consequently the default handoff delay in 802.11-based WMN can be around 1000ms. 300ms delay is expected when using passive and optimistic probing technique.

However, beacon intervals are configurable on APs and a beacon interval of 100ms represents a mere 4% drain on 802.11b throughputs [31]. In [31] it is showed that the beacon intervals can be reduced to 10ms, while the tradeoff is that over 30% of the throughput is wasted. Beacon intervals of 10ms would result in probing delays of 110ms and 30ms for passive probing and optimistic probing techniques. By optimizing the beacon intervals, it enhances the support for time-sensitive applications, e.g. VoIP. In the meantime, the loss of throughput in 802.11-based networks is an acceptable tradeoff as long as the bandwidth is enough to support other data applications.

2.2.2.3 Timer Optimization for Active Probing

Active probing is faster than passive probing and involves aggressively probing channels for new APs. The active probing delay is determined by the minimum and maximum channel timers. The bounds for active probing delays are defined below where *Num_Channels* is the number of channels and t is the total scan time:

$$Num\_Channels \times MinChannelTime \leq TA \leq Num\_Channels \times MaxChannelTime \qquad (2.2)$$

Optimizing the Timer *MinChannelTime*

By lowering the min and max channel times, the total scan time can be significantly reduced. However, if they are inappropriately set, the MC can miss potential APs. Velayos et al [34] consider the optimization of the *MinChannelTime* variable for faster handoff using the formula shown below: *MinChannelTime* ≥ DIFS + (*aCWmin* × *aSlotTime*). DIFS (DCF Interframe Space) is the time interval between frames to provide contention free access. The *aCWmin* is the number of slots in the minimum contention window and *aSlotTime* is the length of a slot. These values can be collectively thought of as the maximum time an AP would need to answer a probe request, given the AP is idle. If the AP is busy, the MC will detect the radio activity and start the max channel timer. Each physical layer 802.11 standard will contain different timers associated with their respective modulation. Inserting the values into the formula returns a value of 670 µs. However, the exact figure is somewhat superfluous as the standard stipulates that all timers are expressed in TU's (Time Units) denoted as 1024 µs or 1.024ms. Subsequently a MinChannelTime of 1 TU is the proposed min channel time for devices requiring fast handoff.

Optimizing the Timer *MaxChannelTime*

The max channel time is somewhat more difficult to quantify than the min channel time. This is because the number of responses received from the probe request is unknown. The timer must be suitably large to receive and acknowledge all AP responses. Consequently the more APs sharing the same channel in range of the client, the longer the *MaxChannelTime* Timer must be set or dynamically extended. Based on active probing measurements, max channel times are approximately 15ms, however this value varies between vendors and chipsets. Arbaugh claims that 6ms is the optimal max channel time in 802.11b networks with one reachable AP per channel.

2.2.3 Proactive Mechanisms of Fast Handoff

2.2.3.1 Neighbor Graph-based Techniques

Neighbor graphs are the key to building intelligence for fast handoff into 802.11-based networks. They provide a way to dynamically map wireless network topologies and inform MCs of surrounding nodes. Featuring in 802.11f, 802.11r, 802.11e and 802.11k proposals, neighbor graphs are at the forefront of the latest wireless research. Neighbor graphs turn physical topological maps into data structures displaying potential associations.

Neighborhood graphs are dynamically learned through disassociation messages. When a client transitions to leave an AP, it sends a disassociation message to the old AP [5] containing the details of the new AP. The AP specified in the disassociation message can now be recorded as a neighbor. In reference [5] it presented a novel discovery method using a neighbor graph (NG) and a non-overlap graph (NOG). This scheme (referred to as the NG-pruning scheme) focuses on reducing both the total number of channels to be probed and the waiting time on each channel. They suggested two algorithms: the NG and NG-pruning algorithms. The rationale behind these algorithms is to ascertain whether or not a channel needs to be probed (by the NG algorithm) and whether the MC has to wait more probe response messages on a specific channel before the expiration of MaxChannelTime (by the NG-pruning algorithm). The NG abstracts the handoff relationship between adjacent APs. Using the NG, the set of channels on which neighboring APs are currently operating and the set of neighbor APs on each channel can be learned. Based on this information, an MC can determine whether or not a channel needs to be probed. On the other hand, the NOG abstracts the non-overlapping relation among the APs. Two APs are considered to be non-overlapping if and only if the MC cannot communicate with both of them simultaneously with acceptable link quality. For instance, if the distance between AP_i and AP_j is far, an MC can associate with only one of them. In this case, AP_i and AP_j are non-overlapping each other. Therefore, if the MC has received a probe response frame from AP_i, this implies that the MC cannot receive a response frame from AP_j by the principle of non-overlapping. By means of the NOG, the

MC can prune some of the APs which are non-overlapping with the current AP group that has already responded. Figure 7 illustrates the operation of the NG-pruning scheme. The un-bracketed and bracketed numbers represent the AP identifier and channel number used by the AP, respectively.

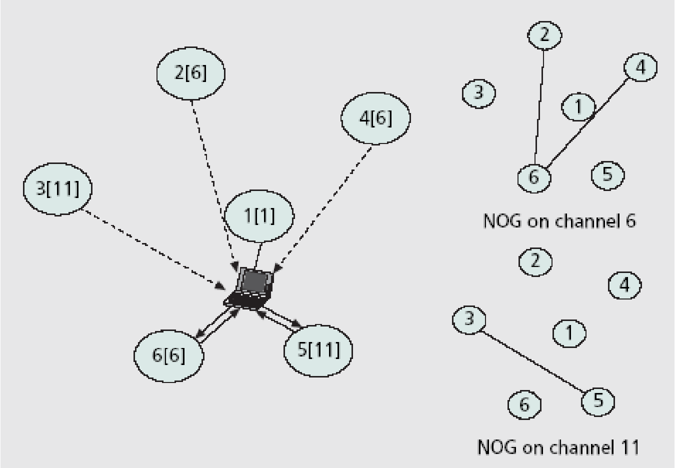

Figure 7 Neighbor Graph-based handoff scheme

In this example, only three channels (i.e., 1, 6, and 11) are used and the current AP (AP_1) has five neighboring APs (AP_2 to AP_6). The neighbor information can be learned by the construction of the NG. By using this neighbor information, the MC knows that the number of channels it has to probe is just two (i.e., channels 6 and 11). On the other hand, individual NOGs are constructed on each channel (i.e. one NOG on channel 6 and the other NOG on channel 11). First, suppose that the MC is probing on channel 6. When it receives a probe response message from AP_6, the MC decides that it is unnecessary to wait for additional probe response messages on channel 6. This is because AP_6 is non-overlapping with AP_2.

2.2.3.2 Fixed Time-Division Scan: SyncScan

The original reactive approaches perform handoff whereby the probing begins following the loss of connectivity or the reduction of SNR. SyncScan introduces the concept of proactively probing surrounding channels at synchronized intervals regardless of the SNR [31]. The SyncScan design requires that all APs' broadcast beacons simultaneously. Clock synchronization between every AP and every client is achieved with NTP (Network Time Protocol). Assuming synchronization between MCs and APs, MCs can simultaneously switch to a different channel just in time to receive beacons. After the beacons have been received, MCs switch back and resume normal operation.

A major advantage of SyncScan is a more intelligent decision phase [31]. Instead of waiting until the SNR drops below a certain threshold to begin probing, MCs can handoff as soon as they locate an AP with a superior SNR.

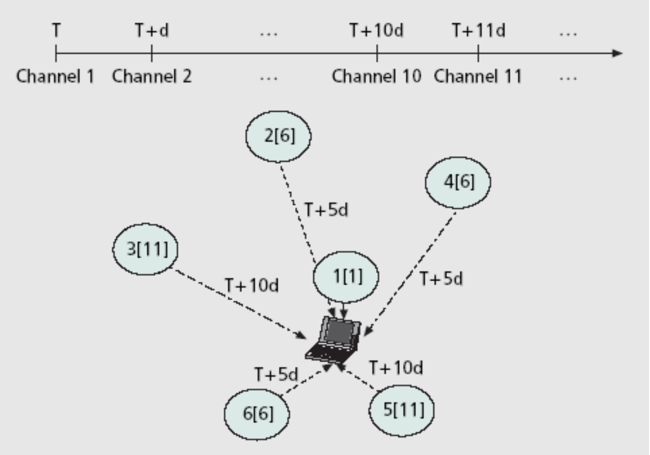

Figure 8 SyncScan handoff scheme

Problems with SyncScan include increased packet loss and the subsequent reduction in bandwidth. Packets are lost as a result of switching to another channel. Though SyncScan channel switches can be performed in less than 40ms, the cost from performing this operation may be too high. As previously discussed packet losses are disastrous for voice communication. While proactive approaches offer ways to fast handoff, SyncScan operation results in many unwanted side effects.

2.2.3.3 Dynamic Time-Division Scan: ProactiveScan

Similar to SyncScan but further optimized in many aspects, Microsoft presents a novel technique to achieve proactive probing which is called ProactiveScan. The difference of ProactiveScan compared with SyncScan is that the latter performs the scan according to fix time allocation while the former performs the scan in rounds.

The ProactiveScan employs two new techniques. The first is to decouple the time-consuming channel scan from the actual handoff, and to eliminate channel scan delay by doing scan early and interleaving it with ongoing traffic in a non-intrusive way. The second technique is a smart trigger that takes into account both uplink and downlink quality and explicitly addresses the link asymmetry which has not yet been touched in previous work.

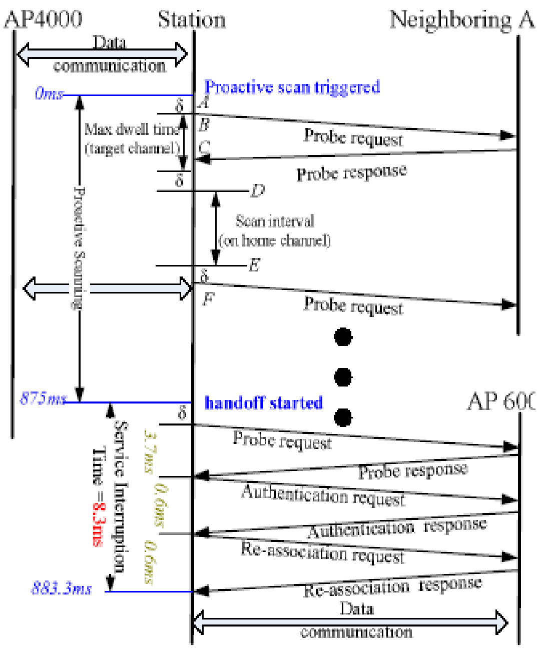

Figure 9 ProactiveScan handoff scheme

2.2.3.4 Dual-Radio Proactive Scan: MultiScan

Brik et al. introduced a handoff scheme utilizing multiple radios called MultiScan [38]. Similar to SyncScan, MultiScan obtains information on neighbor APs by scanning opportunistically. However, MultiScan requires an additional radio interface for the channel scanning. In MultiScan, the primary interface is associated with the current AP and used for data transmission. At the same time, the secondary interface is performing the channel scanning. If a handoff to a new AP is required, the second interface is associated with the new AP while the primary interface is still employed for data transmission. After the completion of a new association by the secondary interface, interface switch from the secondary interface to the primary one is triggered. As a result, the formerly secondary interface becomes primary for data transmission and the formerly primary interface is used for channel scanning. Consequently, MultiScan achieves a make-before-break handoff by using multiple radio interfaces.

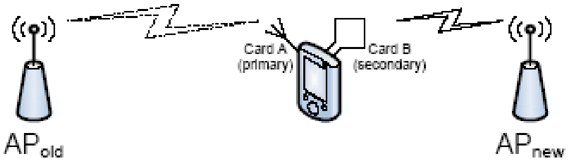

Figure 10 MultiScan handoff scheme

2.3 Secure Fast Handoff

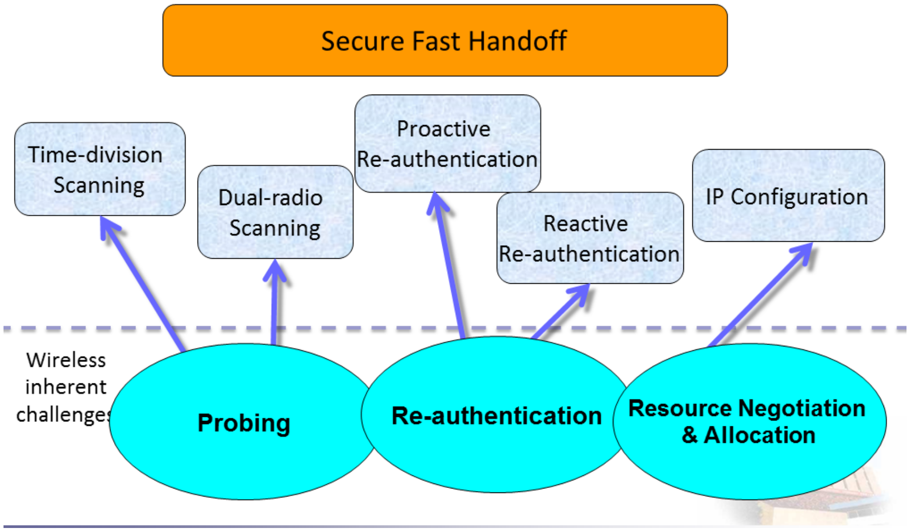
Secure Fast Handoff
Time-division Scanning
Dual-radio Scanning
Proactive Re-authentication
Reactive Re-authentication
IP Configuration
Wireless inherent challenges
Probing
Re-authentication
Resource Negotiation & Allocation

Figure 11 From fast handoff to secure fast handoff

In the previous sections, we have investigated the conventional handoff mechanism and the related works to achieve fast handoff in academic community. Nevertheless, despite the Layer 2 and Layer 3 handoff mechanisms, the upper layer operations of the network could also have adverse effect on handoff, in particular, the security and authentication mechanisms. Therefore, the secure fast handoff take all the above into account in order to maintain better performance. We present the prominent work in literature as follows.

2.3.1 Centralized Authentication in Wireless Networks

Since currently the most popular wireless networks for free access is IEEE 802.11 wireless networks, therefore most of research work on authentication in wireless networks is carried out in this network scenario, although they can also be applied to other wireless access networks.

2.3.1.1 802.1X Authentication Architecture

802.1X is a port based authentication which extends the Extensible Authentication Protocol (EAP) over a Local Area Network (LAN) through a process called Extensible Authentication Protocol Over LANs (EAPoL) [53]. It is an extension of IEEE 802 protocol so that it operates on the Layer 2 of the Open System Interface (OSI) model. This is the Data Link Layer and it present in both wired and wireless LAN. Port-based Authentication leverages three components to accomplish authentication process. 802.1X is probably the largest piece of the three, but could not effectively function without the other two. Those other two components are Extensible Authentication Protocol-Methods (EAP-Methods) and Remote Authentication Dial In User Service (RADIUS). As part of the network infrastructure, wireless AP controls the network service port but open an authentication point to all the device attempting to link to this network. AP encapsulates all the credential information in EAPoL package and RADIUS server will do decapsulation and validation. If successful, RADIUS server tells the access point the device wishing to access is allowed. If not, it will not open the network service port to the device.

802.1X consists of three components: Supplicant System, Authenticator System and Authentication Server System, shown in Figure 12.

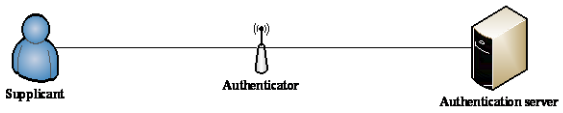

Figure 12 Standard three components in IEEE 802.1X

Supplicant

In physical perspective, supplicant usually is deployed on a mobile device, such as a laptop or smartphone, which attempts to connect to network. It needs to install client site software to initiate the 802.1X authentication process. For the sake of support port-based access control, client system need support EAPoL protocol. For example, Windows system comes with 802.1X built in with variety of EAP-Methods, such as EAP-TLS. The Supplicant communicates with the authentication server using EAP as the transport and specific EAP-Method that provides the actual authentication mechanism. EAP-Method actually is a group of EAP authentication protocols which utilize EAP authentication framework to make extension and variation.

Authenticator

Authenticator is deployed on a network device such as Ethernet switch or an AP. All these devices need to implement the 802.1X in order to support this 802.1X port-based authentication. This device corresponds to different port of user (can be physic port, MAC address, VLAN, IP and so on). Authenticator acts as a secure gate between the supplicant and the protected network. It opens a port to supplicant providing the layer 2 linkage. Until the authentication system verifies the credentials of the supplicant it opens the port and supplicant can access.

Authenticator also is a translator between the supplicant and the authentication server. It receives the credential information which supplicant sends to the authentication server for authentication in EAP frame, and removes the EAP-method data from EAPoL frame. After that, re-encapsulates it in RADIUS frame then forward it to authentication server. In reverse, it de-encapsulates the frame from authentication server and encapsulates it in EAP frame then sends to the supplicant.

Authentication Server

Port-based authentication standards and specification don’t make any particular type of authentication server. But in reality, Authenticator usually uses RADIUS server which provides the service of Authentication, Authorization and Accounting (AAA). It is in a separate part of network which may define many authentication systems at the same time. Authentication system is responsible for building access control list, user access policy and account property, and auditing the later connection of the user.

In some cases, authentication server can also be embedded to the authenticator. This distributed server model significantly reduces the authentication traffic and communication delay. This increases the network performance when the network is not in large coverage. However, when the user keeping adding to the network, manage the authentication database on these distributed authentication server is not that easy.

2.3.1.2 802.11i Authentication Protocol

IEEE 802.11i is widely used method on layer 2. In the 802.1i authentication mobile client and Authentication Server (AS) apply the 802.1X authentication model carrying out some negotiation to agree a pre-shared secret key pair or the Pairwise Master Key (PMK). All of above used by MC and AS to do the further authentication between them. AP can get a copy of the key from AS also, afterwards, a four-way handshake starts between the AP and the MC to generate encryption keys from the generated PMK. Encryption keys can assure confidential transfer between the MC and the AP. If the MC roams to a new AP, this MC will perform another full 802.1X authentication with the AS to derive a new PMK. For performance reasons, the PMK of the MC can be cached by the MC and the AP to be used for later re-association. The process depicted by the below Figure 13:

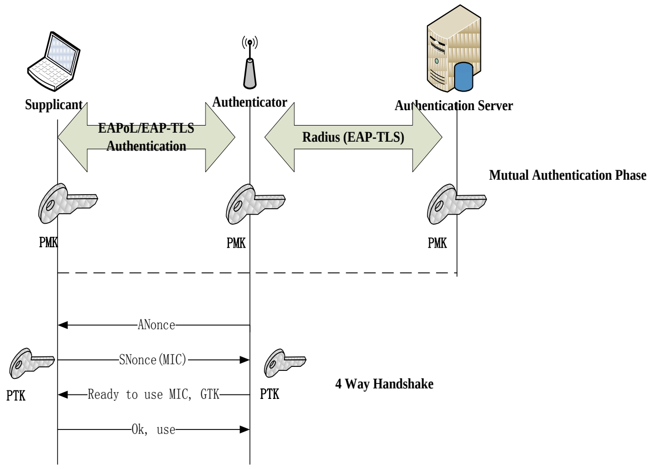

Figure 13 IEEE 802.1i Authentication Protocol

The features of 802.11i exhibit a potential vulnerability because a compromised AP can still authenticate itself to a MC and gain control over the connection. Furthermore, IEEE 802.11i authentication does not provide a solution for multi-hop communication.

2.3.1.3 Authentication Authorization and Accounting (AAA) Protocols

AAA refers to the authentication, authorization, and accounting to provide stand and unified service across heterogeneous access networks. A lightweight AAA infrastructure is proposed in providing continuous, on-demand, end-to-end security in many kinds of heterogeneous networks as in Figure 14.

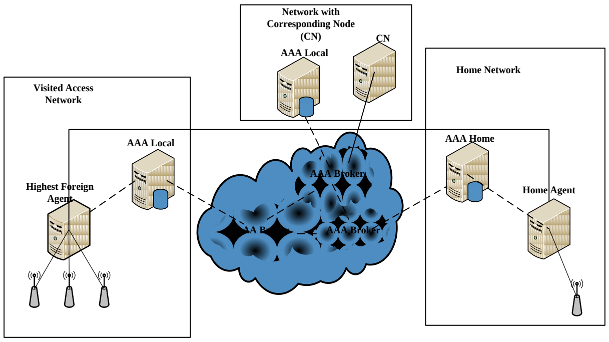

Figure 14 Light-Weight AAA Infrastructure for Mobility in Different Domains

This proposed method uses DIAMETER protocol in its architecture. Security association is reduced to only one so each MC is just require to association with its home AAA server. And the home DIAMETER server can communicate with foreign DIAMETER server in the other domains. Using AAA Broker architecture as distributed control architecture to automate SLA negotiation, load balancing, service composition and billing settlements between the different service providers (stakeholders). This architecture forms a virtual layer on top of the underlying mesh of network domains and intelligent enough to support user mobility as well as service mobility across multiple access networks i.e., dynamically provides AAA.

The pros are: this AAA infrastructure improves performance of the basic mobility protocol: authentication for signaling messages, accounting of network usage, minimal use of cryptographic keys, and the non-use of digital signatures. However the cons are: the main limitation of this architecture lies in the non-support of multi-hop communication between MCs. One way to overcome this limitation is by allowing extended mesh topology among the MCs. Furthermore, employing the IAPP limits the application of this architecture to a specific type of mobile devices.

2.3.1.4 Extensible Authentication Protocol (EAP)

EAP Method

EAP-methods are a group of authentication methods which based on the EAP framework. All the EAP-methods are end-to-end, logical communication mechanism. In port-based authentication system a specific EAP-method defines how the authentication takes places between supplicant and authentication server which including the data package structure and authentication process. And other protocols, such as EAPoL and RADIUS, merely transport the EAP-Methods data. EAP-Method defined EAP data, EAP packages (EAP-Request and EAP-Response packets) carry the EAP-Method protocol headers and data. EAPoL packets transport the EAP packet, and 802.3 (or 802.11) data frames carry the EAPoL packets. In addition to carrying the higher-layer protocol as data, each one of the protocols has its own function and defines packets that might not carry any of the higher-layer protocols.

The implementation and result of an EAP-Method is the goal of the port-based authentication system. The process in 802.1X makes use of different types of credentials, such as username/passwords, encryption keys, and digital certificates.

The standards require implementation of the following EAP-Methods

- MD5 challenge
- One-Time Password (OTP)
- Generic token card

In addition, there are many proprietary and RFC-based EAPoL-methods, such as EAP-TLS, EAP-TTLS, EAP-FAST, and EAP-LEAP. However, whatever it changes, the communication between the supplicant and authentication server are all the same. Figure14 and Figure 15 show the communication between these three entities and the data grams defined in the EAP and RADIUS protocol which used to support this kind of communication.

All EAP-Method packets have the same basic structure, which consist of Type and EAP-Method Data fields (Figure 15). The EAP-Method packet is carried in the Data field of an EAP packet. Hereafter explains the fields of the EAP-Method packet.

Octets:	1	Variable
	Type	EAP-Method Data

Figure 15 Basic EAP-Method packet structure

The EAP-Method Type field, an eight-octet value, identifies a specific EAP-Method. There are many EAP-Methods, some of which were originally defined by the EAP specification (RFC 3748) and many others that are optional and propriety. Table 1 shows the EAP-Method type registered in the RFC 3748. The value indicates the EAP packet type. Every number which fits in the Type field needs to be converted into one byte binary format.

Table 1 EAP-Method Types Registry

Range	Registration Proceed

1-191	Designated Expert with Specification Required
192-253	Standards Action
256-4294967295	Designated Expert with Specification Required

EAPoL Protocol

EAPoL is defined in the 802.1X standard to adapt EAP communications for operation over LANs. So, EAPoL provides additional header fields to EAP packages and creates some specialized EAP datagram. The layering process is similar to other computer network architectures. The collection of these protocols comprises an 802.1X port-based authentication system. EAPoL is the overall port-based authentication entity that requires transportation by a link protocol.

EAPoL Packet Types

EAP-Packet is defined by type “0” which means EAPoL packets merely pass through EAP packets. All the data in the payload is EAP-Method data. After link initiation, the most common EAPoL packets are EAP-Packet entities. EAPoL-Start generates at the beginning of the authentication when link state goes from down to up. EAPoL-Logoff when user indicates the log off from a particular system. EAPoL-Key can be sent wither by supplicant or the authenticator. If 802.1X implementation require keys, this is used to carry Key Descriptor with format in the packet body.

EAP Derivative

As we know, re-using the information of this initial authentication can speed up the following operation for authentication. Secure fast handoff mechanism has come out under this situation, which allows mutual authentication and provides access control protection through limiting the possibility of insider attackers during the re-authentication process. So, old shared authentication keys are no longer fit for this process, but the host only needs to keep the keys which they really needed.

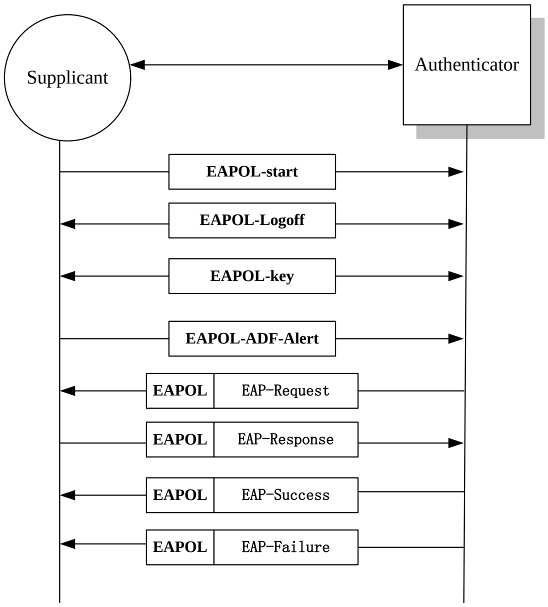

Figure 16 EAPoL package types

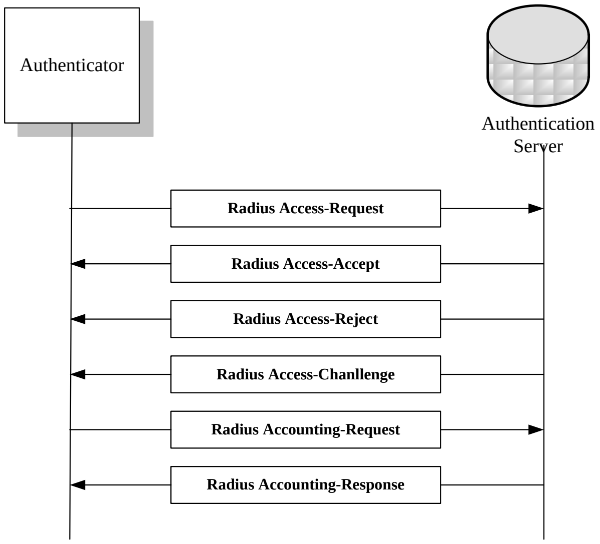

Figure 17 Radius package types

The EAP derivative method with token-based re-authentication means that when the 802.1X authentication ends both authenticating parties share a PMK key, and they can use it to perform re-authentications. The token is a signed with the PMK key and forwarded from the AP to the AS, it includes some identifier of the AP (ESSID). The path from the AP to the AS is a RADIUS tunnel, so that the token arrives to the AS with proof of authentication of the AP. The identifier included in the token must correspond with the RADIUS key used, and known to the AS. With last message, the PMK is moved to the AP. With modifications, the authenticating key used could be the AMSK key, from TLS protocol. Whenever the MC performs a handoff to another authenticator, the new authenticator should receive the PMK key to avoid a full re-authentication.

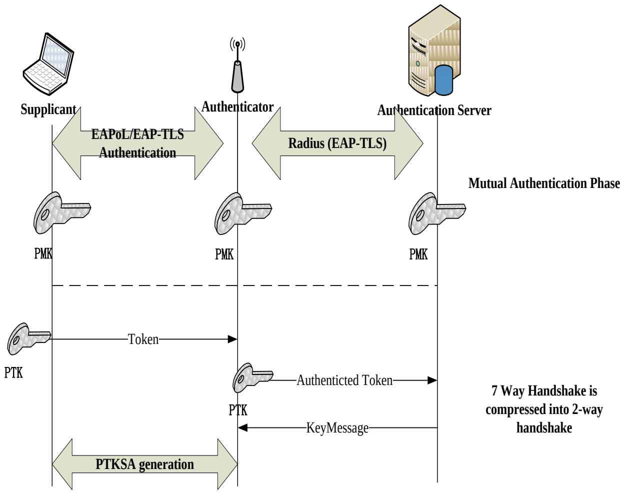

Figure 18 EAP with Token-Based Re-Authentication

2.3.1.5 Wireless Dual Authentication Protocol

Wireless Dual Authentication Protocol (WDAP) [61] is another protocol proposed for 802.11 WLAN. WDAP provides authentication for both MCs and APs and overcomes the shortcomings of other proposed mutual authentication protocols. WDAP provides authentication during the initial connection state and while roaming including three sub-protocols: an authentication protocol, a de-authentication protocol, and a roaming authentication protocol. Figure 19 illustrates the WDAP authentication process.

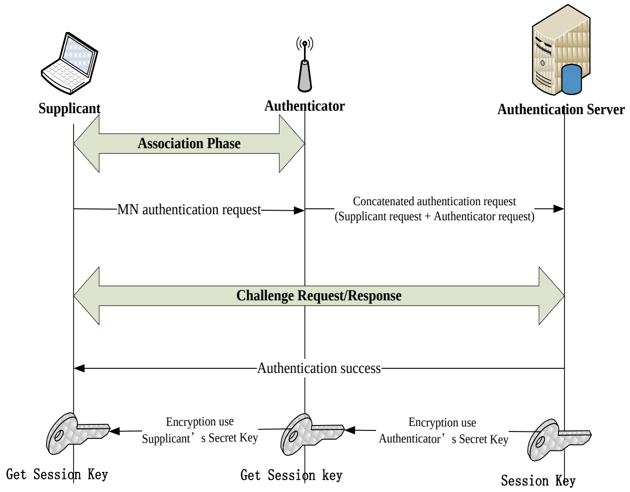

Figure 19 Wireless Dual Authentication Protocol-Authentication Phase

WDAP allows the mutual authentication between MC and authenticator. Also, WDAP can be used to assure the authentication between the authenticators themselves through authentication requests concatenation. In case of multi-hop communication in a multi-hop network, each pair of nodes can mutually authenticate through the session key generated by the AS.

2.3.1.6 Lightweight Hop-by-hop Access Protocol

There is another method to preventing unauthorized modes connect to a pervasive computing environment which is using authenticating transmitted data packets. Authenticating MC in wireless dynamic environments has been proposed in a Lightweight Hop-by-hop Access Protocol (LHAP) [62]. LHAP implements lightweight hop-by-hop authentication, where intermediate nodes authenticate all the packets they receive before forwarding them. In this protocol, MCs to first perform some inexpensive authentication operations to bootstrap a trust relationship with its neighbors, then they follow a lightweight protocol for subsequent traffic authentication. There are pros and cons about this protocol.

The pros are: this protocol is quite adaptable to WMN environments, especially for open scenarios of mesh network when the AS is not in place, preventing unauthorized MC participation in the communication and allowing hop-by-hop authentication. For secure roaming, LHAP can be useful in distributing session keys among MC employing a special type of packet designated for this issue. While the cons are: the focus of this protocol on resource consumption attacks' prevention restricts its application to a number of scenarios. Also, the fact that LHAP does not prevent insider attackers from carrying out malicious actions necessitates complementary solutions with such protocol.

2.3.2 Overview on Distributed Authentication

2.3.2.1 Basic Cryptology Techniques

In general, there are three cryptographic techniques that can be used to devise security mechanisms authentication: one-way hash functions, symmetric cryptosystems, and asymmetric (or public key) cryptosystems. An asymmetric cryptosystem is more efficient in key utilization in that the public key of a node can be used by all the other nodes; a symmetric cryptosystem requires the existence of a shared key between two communicating nodes. Hashing function is the most quickly and easy use method which can work together with the symmetric and asymmetric algorithm, such as a digital certificate or a keyed hash value (i.e., a keyed message authentication code).

Portable devices in a pervasive computing environment usually have limited battery life and must share a relatively limited transmission bandwidth. Therefore, symmetric cryptosystems are preferable in ad hoc scenarios due to their computational efficiency (conducting an asymmetric algorithm usually is three or four orders of magnitude slower than the symmetric counterpart). For a symmetric cryptosystem to work, a shared key must be established between each pair of communicating entities. The key establishment problem between two network principals is well understood for conventional communication networks, and generally can be resolved by key distribution or key agreement.

A symmetric cryptosystem uses shared key which is widely applied in the ad-hoc network due to the communication efficiency. Every shared key used by two entities needs to be identical which means over all there requires second shared keys. This method works well in a small group of users which may bring complicated problem in key distribution and management in big group. In asymmetric key algorithms group such as AES which need great amount of computation and most of time public-key cryptography is used to generate public certificate. The key pairs and certificate should be distributed to the end device and store on it which brings slow response in communication.

The classic key-distribution scheme Kerberos, requires an online centralized authority (CA) to generate and distribute the keys.

Key agreement protocols, such as the Diffie–Hellman key exchange protocol and many variations derived from it, do not need an online CA and compute the shared keys between nodes on-demand. These protocols are interactive schemes in that nodes need to exchange messages between them to establish the desired keys, for which active routes must pre-exist for such approaches to work. The assumption of pre-existing routes between two communicating parties, which may be multiple hops away from each other, contradicts the need to secure the routing discovery process between such nodes in the first place. Even if such an assumption is satisfied, network dynamics can tear routes down in the middle of the key negotiation, and as such no key can be agreed upon. Moreover, interactive key agreement protocols are not scalable in terms of communication overhead, because messages exchanged for key establishment can consume significant CPU cycles and wireless bandwidth in such a highly dynamic environment, which can become even worse if the shared keys between nodes need to be updated frequently.

2.3.2.2 Elliptic Curve Cryptology

Elliptic curve has been seen as one of the most cryptology algorithm as so far. In mathematics, an elliptic curve is a smooth, projective algebraic curve of genus one. An elliptic curve is in fact an abelian variety — that is, it has a multiplication defined algebraically with respect to which it is an abelian group. A typical elliptic curve can be written like: $y^2=x^3+ax+b$.

Based on the Elliptic curve algorithm, there comes Elliptic Curve Cryptology (ECC). ECC is a public-key cryptology utilized the intractability of certain mathematical problems. The set of points on such a curve — all solutions of the above equation together with a point at infinity — form an Abelian group, with the point at infinity as identity element. If the coordinates *x* and *y* are chosen from a finite field, the solutions form a finite abelian group. If the finite field is large, the discrete logarithm problem on such elliptic curve groups is believed to be more difficult than the corresponding problem in the underlying finite field's multiplicative group. Thus keys in elliptic curve cryptography can be chosen to be much shorter for a comparable level of security compared to integer-based methods.

By adding a "point at infinity", we obtain the projective version of this curve. If P and Q are two points on the curve, then we can uniquely describe a third point which is the intersection of the curve with the line through P and Q. If the line is tangent to the curve at a point, then that point is counted twice; and if the line is parallel to the y-axis, we define the third point as the point "at infinity". Exactly one of these conditions then holds for any pair of points on an elliptic curve.

2.3.2.3 Pairwise Key Pre-distribution Authentication

Public key cryptography is not feasible in mesh networks and therefore only symmetric schemes are applicable. The approach that all APs share the same secret key for authentication and encryption is not suited in mesh networks because APs provide only weak physical protection. In this case, once an adversary gains physical access to an AP in the network, she/he could read out the secret key, and thus, the entire network could be compromised. For this reason, sharing keys pair-wise seems to be a more reasonable approach. In addition, this approach enables entity authentication. Since APs have very constrained memory, they cannot store symmetric keys of every other AP in the network. To overcome this constraint, key pre-distribution protocols, which assign each AP a subset of the total set of symmetric keys, are proposed recently. Note that the APs of a network always belong to one domain. For most AP networks applications, it can be assumed that a trusted authority can set-up all APs before they are deployed. This process is called key pre-distribution.

Eschenauer and Gligor proposed a probabilistic key pre-distribution protocol in [58]. In their scheme, each AP is initialized with a random subset of keys out of the entire key pool. When two APs wish to securely communicate, they check if they directly share a secret key. If they do not, they have to try to find a common neighbor with whom they both share a key with and use this intermediate node(s) to establish a secure key.

In the pairwise key pre-distribution protocol proposed by Liu and Ning [59], the authors make use of the facts that most AP networks are static, i.e. AP do not move once deployed, and that the location of AP can be predicted. They argue that each AP has an expected location, thus, an AP can be initialized with a set of keys from its expected neighbors. The authors argue that APs can only talk to nodes in their direct neighborhood, because of their limited transmission range. By implementing a location-based approach, the probability that two neighboring nodes share a key is higher than in a probabilistic pre-distribution scheme. This approach is suited in static networks, in which the location of single nodes can be predicted.

2.3.2.4 Threshold Secret Sharing Authentication

The threshold secret sharing [54] is a recently developed concept to support such distributed authentication. In literature, there are two proposed schemes using threshold cryptography to distribute the services of certificate authority. Zhou and Hass [55] use a partially distributed certificate authority scheme, in which a group of special nodes is capable of generating partial certificates using their shares of the certificate signing key. A valid certificate can be obtained by combining k such partial certificates. This is the first threshold secret sharing scheme introduced for ad hoc security protocols and could serve as an excellent guide to the following work. The weakness of the solution is that it requires an administrative infrastructure available to distribute the shares to the special nodes. The scheme is further complicated by the normal nodes need to locate the server nodes.

2.3.2.5 Basic Self-certified Public Key

In an asymmetric cryptosystem, each user possesses secret and public keys. There might be an organization that takes care of these key pairs call third party authority. Third party Authority is used to ensure the authentication of the published keys and the certified the keys.

In CCITT X509, the certificate-based schemes, the guarantee G takes the form of digital signature of the pair (I, P). It is usually called certificate and is widely used in contemporary. In this way, (I, s, P, G) are distinct and need to be stored in the directory.

User i gets the (I, P, G) pairs, check the G using the authority's public key which is known to everybody. So that can ensure the reality of the P.

Shamir has proposed Identity-based schemes in 1984. He used the identity *I* to generate the public key (i.e. *P=I*). And the guarantee is the secret key s itself (i.e. G=s), so that it only needs to distribute the (I and s). This approach deduced the certification check and store but has great drawbacks. In particular, the authority can impersonate any user at any moment since secret keys are calculated by it.

Self-certified public key is a subset of certificate-based schemes algorithm. In ID-based cryptography (or Identity-Based Encryption (IBE) algorithm, the user identity is used to join the computation of its public key, for example the user ID, user's email address and so on. Because the public key is open to all the entities in the network and the information in the public key is also no need the credential. However due to the "publicity of" the public key, it faces higher risk to be as the attack target. Fake public keys substitute the true ones in the directory. There brings the problems that use must keep a guarantee G which can ensure that the identity I and the key pair (s, P) is really the true ones belong to the identity who asserted.

Self-Certified Key (SCK) follows the track of implicit verification in which the authenticity of a public key is verified when it is used for encryption or decryption, signature verification, key exchanging, or other cryptographic operations. Marc Girault had proposed a method of self-certified public key 1998. He used the RSA/Robin digital robin method to generate this kind of guarantee G which is equal to the public key (i.e. G=P). User using the self-certified public key, so it no needs to have separate certificate, and the public keys is not restrict to the identity. In addition, the secret keys are chosen by the user himself and remain unknown to the authority.

Chapter 3 HAWK Testbed with L2&L3 Cross-layer Fast Handoff and Fast Authentication Re-using Intermediate Credentials

3.1 Introduction

Although there is much work addressing L2 handoff in WLAN, fast handoff in 802.11-based WMN has not been well explored as it involves both L2 and L3 handoff. Although Mobile IP (MIP) has been frequently adopted as L3 handoff solution, it does not fit in the scenario of small / medium coverage of a WMN. So we aim to investigate and develop a practical solution including both L2 and L3 schemes to especially suit 802.11-based WMN. In this work, a large-scale 802.11-based WMN testbed is built on campus with dozens of custom-made nodes, which function as 802.11b AP and 802 11a Mesh Router (MR) to establish a mesh backbone for MC to access. We implement proprietary L2 and L3 fast handoff schemes: namely background selective channel scanning and location management-based routing update. Series of field tests have been conducted to investigate and recursively fine-tune key parameters to achieve fast handoff through analysis, which contain scanning and networking parameters. Eventually our WMN testbed achieves that L2 handoff delay is reduced to 8ms-19ms, L3 handoff delay 14ms-60ms, and the end-to-end delay varies around 30ms-80ms.

3.2 HAWK: an 802.11-based WMN Testbed

3.2.1 Architecture and Components

Our HAWK (Heterogeneous Advanced Wireless networKs) testbed is an infrastructure mesh network based on underlying IEEE 802.11 technology. As in Figure 20, it consists of two types of physical entities: MR (mobile router) and MC (mobile client). The MR T901 adopts ARM9 + Linux 2.4.20 and its successor T902 adopts IntelX86 + Linux 2.6.10.The MCs are Fedora Core 6 notebook PC equipped with 802.11a/b/g NIC.

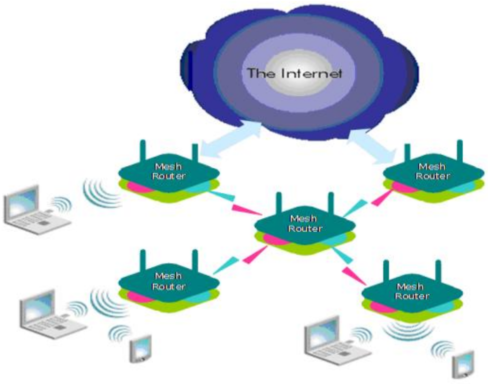

Figure 20 The architecture of HAWK testbed

The MR is enabled with an OLSR (Optimized Link State Routing) which is a flat routing protocol to enable mesh routing. Meanwhile, it is customized and utilized to support location management. The MC is equipped with corresponding processes to support vital functionalities including packet forwarding and handoff management.

3.2.2 Mobility Management with L2 Handoff and L3 Handoff

One of key characteristics of HAWK testbed is the mobility management which includes location management which introduces L3 handoff delays and L2 handoff scheme.

Generally it is based on OLSR routing update by continuously keeping track of the association between the MC and MR. Each MR periodically broadcasts the *Hello* messages to send updated topology information throughout the entire network. When handoff occurs, the new client-router association will be updated on the MC, the associating MR and other MRs. In this way, the communication packets will be forwarding to the new AP. Since routing is not the major concern of this paper, we mainly introduce the mobility management modules. As in Figure 21, the L3 handoff entities and the interactions are illustrated. On MC side, the L2 Handoff Monitor monitors the movement of the MC. When handoff occurs, it triggers the L3 Handoff Requestor and updates local client-router association by notifying the Handoff Request Handler to change associated AP. On the network side, the Route Update Handler sends HNA message to OLSR Daemon and updates the association information stored on each MR to change the routing information.

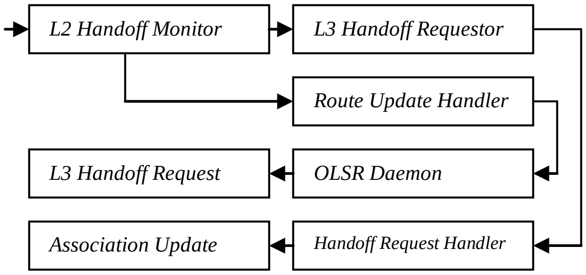

Figure 21 The functional entities of L2 and L3 handoff

For L2 handoff, the scanning is optimized by background selective scanning which decouples scanning from the conventional handoff process. The scanning process is initiated when scanning threshold is reached. In the course of scanning, only orthogonal channels e.g. CH1, CH6 and CH11 are scanned first. The scan list and results are stored adaptively maintained. Although two scanning modes are specified in the baseline standard: passive and active mode. We choose the active scanning mode for the scheme. The parameters *MaxChannelTime* and *MinChannelTime* constrain the delay of the active scanning mode. Based on the literature and our own test, we set up with *MaxChannelTime*=10ms and *MinChannelTime*=1ms respectively to balance the active scanning delay and discovery of APs.

3.2.3L2 and L3 Cross-layer Optimizations

As in Figure 22, we further look into the detailed procedure of L2 and L3 handoff. The original formula to calculate the handoff delay is:

$$Delay_{Total} = Delay_{L2} + Delay_{L3} + Delay_{L3\_process} * n(hop) \qquad (3.1)$$

After implementing background selective scanning, the Delay$_{L2}$ becomes mainly comprised of processing of scanning result, processing of association request and response, and processing of the authentication, which are at the order of a few milliseconds. Since the broadcast of Route Update from a MR to another is determined by hardware and wireless propagation, *the Delay$_{L3\_process}$ is* at the order of a few milliseconds too. Hence, we discover that the bottleneck of delay further lies in L3 Handoff triggering and the Local Route Update which occur on MC and MR.

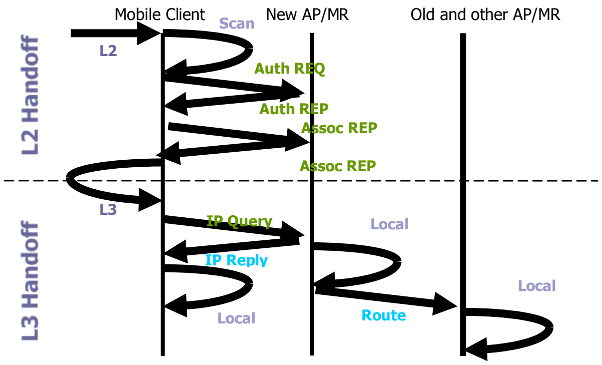

Figure 22 Detailed procedure of L2 and L3 handoff

The conventional way to trigger L3 handoff is the detection of movement on network layer, e.g. change of IP address. Obviously, it induces long delay. Therefore we utilize RSSI to trigger L3 handoff. When the RSSI indicator, which also functions as the L2 handoff trigger, exceeds certain threshold, the L2 and L3 handoff will both be triggered. In this way of cross-layer design, the abovementioned L3 handoff triggering delay is highly eliminated.

As for the delay induced by Local Route Update on MC and MR. We find that fine-tuning OS parameters can further significantly reduce processing delay. On MR side, we configure the OS routing parameters as the follows:

/proc/sys/net/ipv4/route/max_delay=0

/proc/sys/net/ipv4/route/min_delay=0

/proc/sys/net/ipv4/conf/all/send_redirects=0

/proc/sys/net/ipv4/conf/ath0/send_redirects = 0

On the MC side, we set up the routing parameters as follows:

/proc/sys/net/ipv4/conf/all/accept_redirects=0

/proc/sys/net/ipv4/conf/ath0/accept_redirects = 0

In addition to that, we configure ARP caching by bypassing the caching query. Based on the above L2/L3 handoff schemes and tuning measures, the HAWK testbed is evaluated with a series of experiments.

3.2.4 Experiments and Evaluations

3.2.4.1 Test Scenarios

We conduct experiments both in outdoor and indoor environments. As in the left photo in Figure 23, the outdoor environment is an open space in the campus. The HAWK testbed is deployed in a manner of 3 by 2 MRs in a square which has flowerbeds centered. As in the right photo of the indoor scenario, it is a linear deployment of a 4 MRs. The distance between two MRs is around 25 meters. We adjust the transmission power of NIC to establish mesh backbone with appropriate distance. The scanning trigger RSSI is set to 16 db.

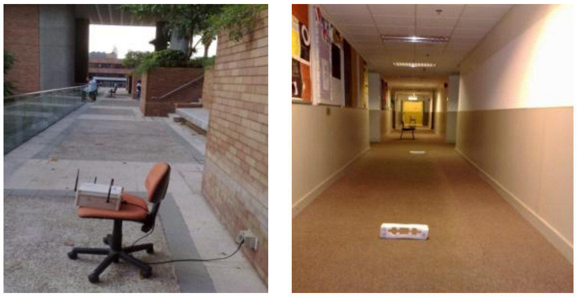

Figure 23 The outdoor and indoor test fields

As in Figure 24, we move the MC from one end of the test field to the other end. The MC does handoff in turn from one MR to another. All MRs have self-organized into a mesh backbone. We maintain a log of traces about the starting time and ending time of each packet in the procedures which is illustrated in Figure 24. After conducting series of experiments, the traces are the basis for our evaluation.

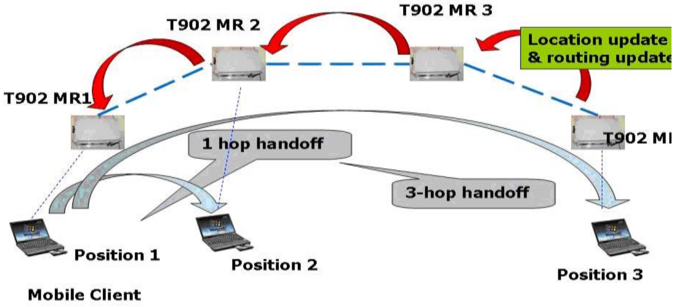

Figure 24 Illustration of test scenarios

3.2.4.2 Experimental Results before Optimizations

The primitive HAWK testbed without fast handoff scheme has also been tested and showed poor performance which is caused by both MAC layer and network layer. When the MC switches between two MRs as in Figure 25, it takes 220ms-610ms for L2 handoff and 155ms-280ms for L3 delay. The total delay is obviously unacceptable for VoIP service.

After the basic fast handoff scheme of background selective scanning is enabled, the results show that it takes 10ms-20ms for L2 handoff and 125ms-180ms for L3 delay. Such a total delay still does not meet requirement of VoIP.

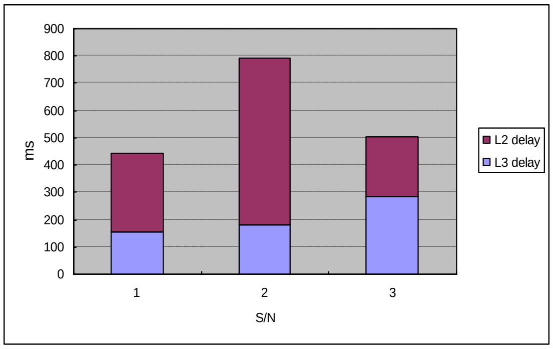

Figure 25 The handoff delay before implementing fast handoff scheme

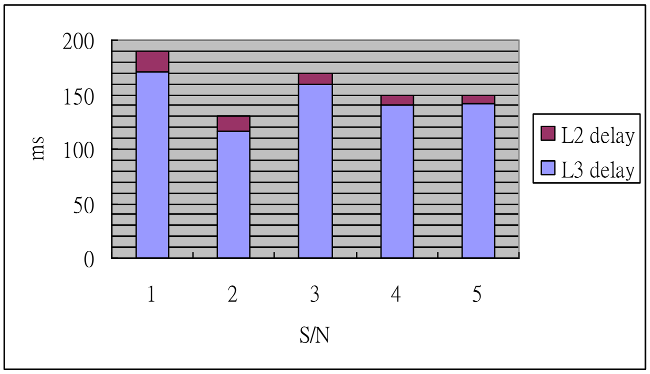

Figure 26 The handoff delay with selective scanning

3.2.4.3 Experimental Results after Optimizations

After we enable cross-layer optimizations on L3 handoff triggering and OS parameters, the handoff performance is improved as in Figure 27 and Figure 28: in an outdoor environment, it takes 4ms-22ms for L2 handoff and 14ms-38ms for L3 delay; while in an indoor environment, it takes 4ms-14ms for L2 handoff and 16ms-69.8ms for L3 delay. The average total delay is close to support VoIP in both environments.

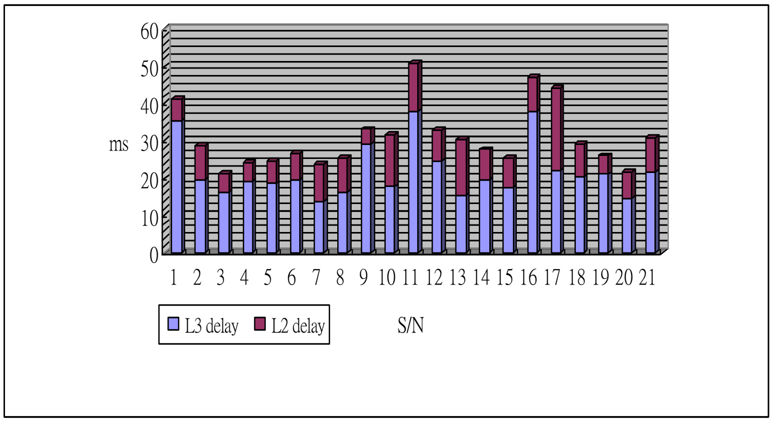

Figure 27 The outdoor handoff delay after L2&L3 cross-layer optimizations

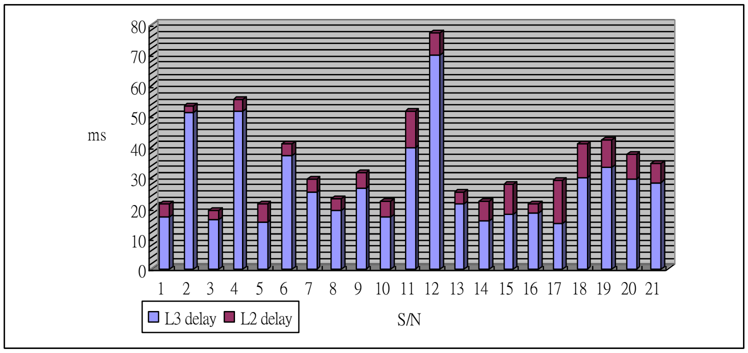

Figure 28 The indoor handoff delay after L2&L3 cross-layer optimizations

As introduced previously, authentication is usually ignored in early research, although it does not reflect the real-world fact. Therefore we enable WEP to test the performance. The delay is proportional to the number of messages exchanged between the AP and mobile client. But the actually processing time of authentication frames on MAC layer is at the order of milliseconds; L2 WEP authentication does not impose significant influence on handoff delay.

3.3 Fast Re-authentication by Re-using Intermediate Credentials in 802.11i

IEEE 802.11i is the most widely used secure handoff and authentication method in 802.11-based WMN and it is implemented in our HAWK testbed as the security mechanism for secure fast handoff by default.

As in Figure 29, in 802.11i authentication, MC and authentication server (AS) apply the 802.1X authentication model carrying out some negotiation to agree a pre-shared secret key pair or the Pairwise Master Key (PMK). All of above used by MC and AS to do the further authentication between them. AP can get a copy of the key from AS also, afterwards, a four-way handshake starts between the AP and the MC to generate encryption keys from the generated PMK. Encryption keys can assure confidential transfer between the MC and the AP. If the MC roams to a new AP, this MC will perform another full 802.1X authentication with the AS to derive a new PMK as in Figure 30.

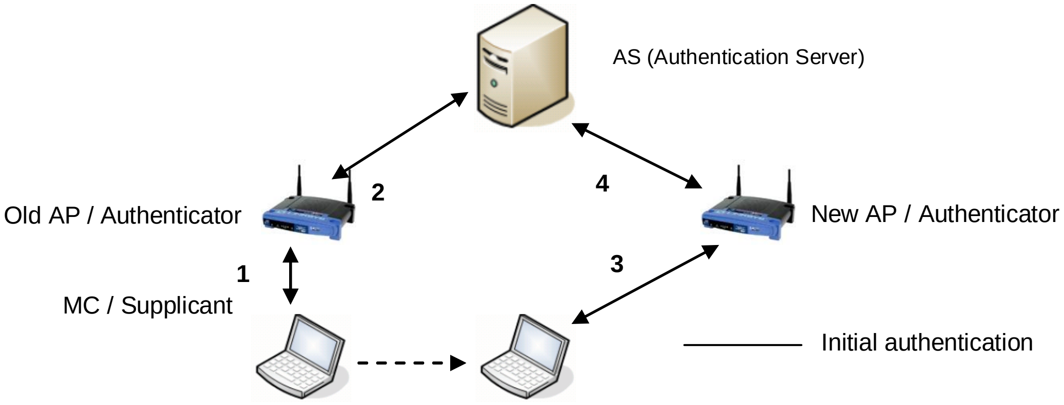

Figure 29 Conventional 802.11i authentication architecture

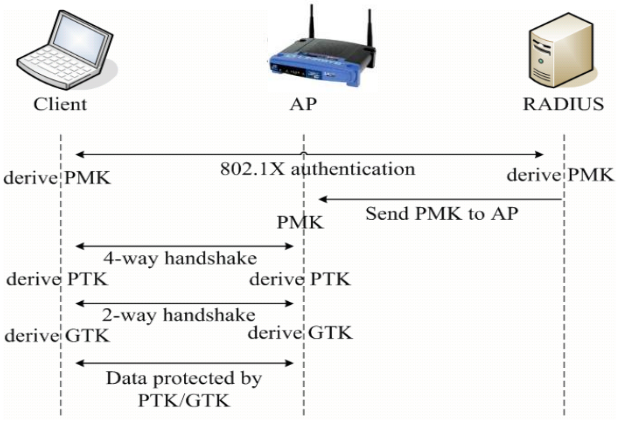

Figure 30 Conventional usage of PMK for new AP

3.3.1 Enhanced Protocol to Re-use Intermediate Credentials

In our enhanced protocol, the intermediate credentials are managed and re-used cooperatively by old AP and new APs in a distributed manner, instead of being re-generated by AS, assuming the communication between AS and the new AP takes longer time than that between the old AP and the new AP. For performance improvement, the PMK of the MC can be cached by the MC and the AP to be used for later re-association.

As in the Figure 31, the conventional 802.11i authentication method is re-engineered from a centralized way to a distributed way by passing the intermediate credentials not only from Authentication Server to APs but also from AP to AP. The intermediate credentials which are generated of the initial authentication are temporarily stored on APs for re-authentication.

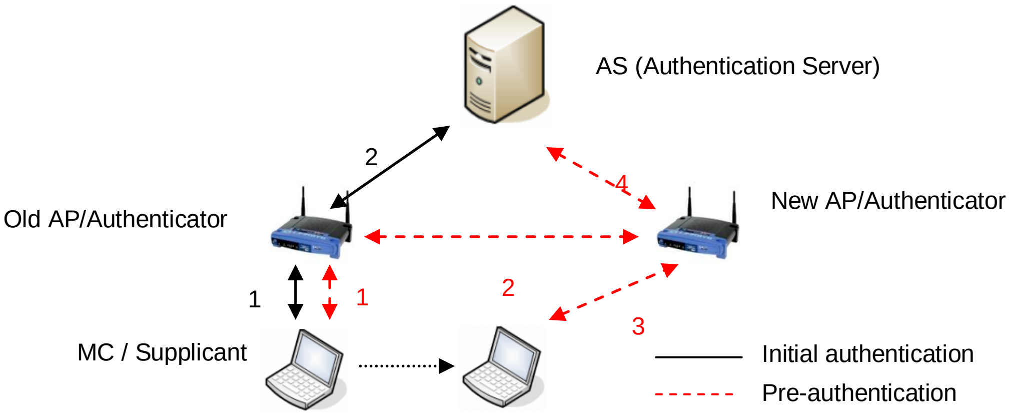

Figure 31 Re-engineer 802.11i to re-use inter mediate credentials for re-authentication

The related procedures of 802.11i have been modified and optimized as well. In Figure 32 major intermediates including PMK and PTK are generated in 4-way handshake and pass to new AP from the old AP rather than Authentication Server. And the derive PTK and GTK are passed to new AP from the old AP as well.

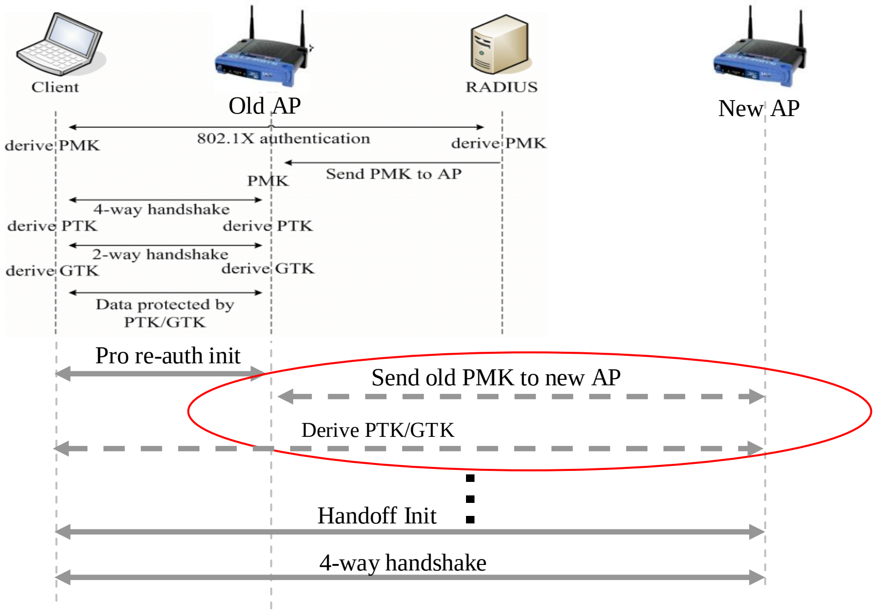

Figure 32 Reusing PMK and PTK/GTK for new AP

Refreshment of Keys

In order to reduce the security risk introduced by passing the intermediate credentials among APs, e.g. from old AP to new AP, rather than passing the credentials from AS to APs, the intermediate credentials are periodically triggered to refresh the key, e.g. PMK/PTK, to satisfy the security requirement of fresh key derivation at APs. Mobility of MC and density of APs are taken into account for key refreshment, which is represented by mobility pattern.

$$P_{MN} = \frac{N_{AP}}{\Delta t} \tag{3.2}$$

where N_{AP} is the numbers of AP associated previously. The neighbor graph G is constructed by all APs which could be associated by MC next.

P_{MN} is used to determine the set of next APs and the time to refresh the PMT/PTK.

In this way, a proactive authentication mechanism to control MC to proactively do AP authentication server has been incorporated to utilize the 802.11i pre-authentication and extends it to IP routing-based multi-domain network to reduce the handoff delay in 802.11-based WMN. As the Figure 33 shows, the L2 delay increases insignificantly and remains at the same level as in Figure 27 and Figure 28. The end-to-end delay is around 30ms to 80ms which does not comprise the support for time-sensitive applications.

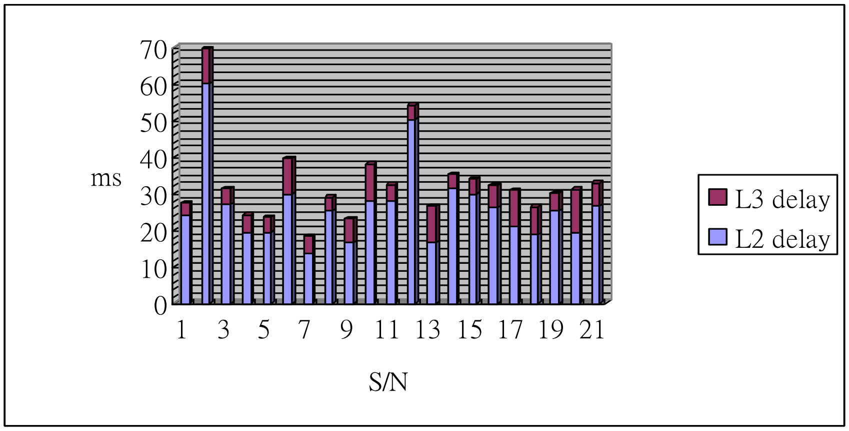

Figure 33 The handoff delay with fast re-authentication in 802.11i

Chapter 4 Proactive Approach with Network-assisted Radio Signatures and Dual Re-Authentication for Secure Fast Handoff

Authentication issue has been mostly ignored to ensure fast handoff in 802.11 Wireless Mesh Network (WMN). With the proliferation of WMNs in recent years for practical deployment, Secure Fast Handoff has drawn much attention to enforce authenticated access while reduce the extra delay caused by enabling authentication operations. In this section, we present an overview on the state-of-the-art advance in this field and tackle the problem from a practical perspective based on experiments and analysis on our real-world testbed HAWK. We propose a novel fast handoff scheme Network-assisted Radio Signature (NRS) to eliminate probing delay by taking advantage of the characteristic of the actual dynamic topology about mesh routers in WMN. Moreover, we apply an optimistic authentication mechanism Dual Re-authentication to counteract the authentication delay while providing the secured wireless access. In this manner, we have reduced the end-to-end handoff delay of WMN back again to a level below 50ms to achieve secured handoff and support time-sensitive applications. We describe detailed mechanisms, simulation, implementation and experimental results. To our best knowledge, we are the first to achieve such an optimal performance of Secure Fast Handoff.

4.1 Introduction

Provision of pervasive wireless Internet access to citizens anywhere and anytime has been a hectic drive for many governments and operators across the globe. To serve this purpose, IEEE 802.11-based Wireless Mesh Network (WMN) has just been matured to a point of wide acceptance as a cost-efficient technology. Hence, authentication issue naturally becomes a significant concern and ongoing effort when it comes to the stage of commercial deployment, though it has been mostly ignored in early studies on Layer 2 handoff. Extra delay is introduced into handoff process by enabling authentication operations like challenge-response, multiple way handshakes and key distribution etc. so that it severely undermines prior efforts towards fast handoff in 802.11-based networks.

We have been carrying out empirical studies on mobility management regarding multiple radio technologies under HAWK (wmn.comp.polyu.edu.hk) project. In this work, we address the challenges of existing IEEE 802.11 handoff mechanisms. By identifying the bottleneck and the direction for improvement, we propose specific and applicable solutions for fast and authenticated handoff to support Internet access. However, the primitive fast handoff and authentication mechanisms do not meet higher requirements of application scenarios in reality.

IEEE 802.11-based WMN consists of a self-organized and self-configurable mesh backbone and conventional IEEE 802.11 MCs. The mesh routers of the backbone communicate with each other via multi-hop wireless links while serving as Access Point (AP) for MC. Although a WMN shares similarities as ad hoc network, it distinguishes itself with a static topology or trivial mobility of mesh routers. In early deployments of WMNs, open access was a common practice which incurs public concerns on security later. So subsequently authentication measures are enforced which include basic WEP, WPA and WPA2. However, although the security objective is fulfilled, the performance of WMNs is influenced, particularly the handoff.

Basically, the fundamental performance bottleneck of handoff in IEEE 802.11-based WMN is caused by interruptions of communication occur in the event of MC changing its point of attachment, e.g. AP. Under the circumstance, the time-sensitive applications such as streaming media demand a delay of less than 300ms while VoIP demands one less than 50ms, specified by the long accredited International Telecommunication Union (ITU) standard [6]. Such requirements impose difficulties upon the primitive IEEE 802.11 network mechanisms.

Handoff involves multiple dimensions of operations from the bottom physical layer to the upmost application layer. Conventionally, research focuses on Layer 2 and Layer 3 due to that they contribute the most of the delay. Concerning Layer 3 handoff, it covers IP acquisition, route update, etc. and has been long investigated. Mobile IP (MIP) and its derivatives [42] [43] have been standardized as major solutions. The situation about Layer 2 handoff is somewhat different and far from being mature. It involves operations including probing for candidate APs, authentication of MC and AP, and association with AP. The delay caused by association is determined by hardware. It is relatively fixed and simple similar to channel switch on Layer 1. However, on the contrary, there are diversified probing and authentication mechanisms existed in vendors' implementation.

Early research targets probing delay with assumption of open authentication which means no authentication at all; currently Secure Fast Handoff tackles with probing delay with authentication enabled. It is obvious that varied authentication methods, e.g. WEP/802.11X/, WPA/802.11i/WPA2, involve various ways of handshaking and exchanges of authentication packets to interrupt communication. The delay accumulates. The longer authentication procedure is, the performance of secured handoff deteriorates more. In order to achieve an optimal overall performance in real world, we have proposed an integrated solution, namely a novel Layer 2 fast handoff scheme Network-assisted Radio Signature (NRS) and an optimistic authentication mechanism Dual Re-authentication. The NRS scheme eliminates the probing delay by taking advantage of the characteristic of the actual topology of a WMN, that is the actually topology of mesh routers is much less dynamic. We assume that probing for candidate APs can be conducted proactively and the outcome of probing can be valid as long as the topology remains unchanged. In this sense, the delay caused by reactive probing is totally eliminated. Moreover, the Dual Re-authentication mechanism counteracts the authentication delay when providing secured wireless Internet access by utilizing optimistic Access-before-Authenticate approach. Based on the total solution, we have eventually achieved an end-to-end secured handoff delay which is around 50ms.

We claim our contributions as follows:

- We propose and implement a total solution of Secure Fast Handoff which includes NRS and Dual Re-authentication mechanisms in an 802.11-based

WMN. To our best knowledge, we are the first to achieve such a performance of Secure Fast Handoff.

- NRS is the first Layer 2 fast handoff scheme to realize proactive probing by making the most of the static topology of a WMN in real world.

- The Dual Re-authentication mechanism improves the performance of secured handoff in a manner of optimistic access, meanwhile, it does not comprise the security by providing its unique design of one time ticket (OTT).

4.2 Related Works

The Secure Fast Handoff of 802.11WMNinvolves two main aspects of operations: namely probing and authentication. In this section, we present an overview on the existing probing schemes and fast authentication mechanisms for Layer 2 fast handoff in Table 2. A simple analysis on pros and cons of the related works is offered subsequently. Note that the probing and scanning are interchangeable in the following sections. Since the probing delay is one of the two main handoff schemes aim to reduce this lengthy process. Some schemes rely on reducing the number of scanning channels, the time taken on each channel, scanning-related timers [25]. The tuning technique targets to find an optimal value for the time taken on each channel, i.e. *MinChannelTime* and *MaxChannelTime* to reduce active scanning delays. Intelligent channel scanning aims to minimize the probing and waiting time on each channel [39]. SyncScan synchronizes MCs and APs and instructs MCs to scan the channel by switching channels at the exact moment when a beacon is about to arrive with periodic beacon broadcasting from APs

[31]. Instead of probing all available channels individually, selective scanning reduces the number of channels required to discover APs. A number of selective scanning approaches have been proposed, such as selective scanning plus AP caching methods [40] designed to reduce L2 handoff delay to a level where VoIP communication becomes seamless [41]. Another typical scheme is Neighbor Graphs approaches, which aim to reduce the total number of probed channels and the probe-waiting time on each channel via selective scanning as well as using caching techniques to solve the problem of packet loss [37]. MultiScan uses multiple radio interfaces equipped on MC to search proactively for alternate APs while being associated with an AP and interleaving data communication [38].

Any authentication mechanism will introduce additional delay and network overhead into the network. Three major approaches to optimize and realize fast authentication in secure fast handoff have been proposed. The first category of approaches is to integrate and merge the overlapping procedures in authentication to reduce the total time delay. In [44], Optimized Integrated Registration Procedure for MIP and SIP with AAA operations is proposed and attempts to reduce the delay by combining the round trips between the MC and the home AAA server when MIP and SIP are both deployed. Another similar case is MPA [45]. Both of them employ cross-layer approach to reduce redundant changes of messages in authentication procedures as one of basic technique to optimize authentication delay, nevertheless they cannot eliminate the delay. Proactive authentication approaches constitute another direction for fast authentication, e.g. Shadow Registration [46], which is proposed to optimize secure handoffs by proactively conducting the authentication before the handoff really occurs. A security association is established between the MC and every neighboring APs before the former associates with one of them. Similar mechanisms are employed in AAA context transfer [47] and P2P context transfer [48]. Such proactive authentication mechanisms can further reduce the authentication delay but may cause the problem of heavy traffic and waste of resources induced by redundant information transfer. The third approach is called optimistic access which postpones the real authentication to a later moment as addressed in [49]. During the handoff process, the network allows an optimistic access by the MC to the new mesh router using a light-weight authentication instead of executing a normal authentication. The full authentication is executed in a later stage. If the MC cannot pass the full authentication, the current access will be terminated. This technique makes a trade-off between security and performance. The possibility of suffering from vulnerability is reduced by the small time window before the full authentication completes.

Table 2 Comparison of fast probing for fast handoff

	Features	**Pros**	**Cons**	**Delays**
Timer tuning	Optimize values of scanning timers	Simple and effective	Limited improvement	*>250ms*
Intelligent channel scanning	Minimize the probing and waiting time on channels	Simple and effective	Limited improvement	*100ms-250ms*
SyncScan	Periodically probe channels in case of communication	Reduce the delay significantly	Large overhead	*<50ms*
MultiScan	Use one radio to probe channels while another radio for communication	Minimize the delay to close to 0ms	Extra hardware is required	*<50ms*

Neighbor graph	Reduce the probed channels and probe-waiting time by using NG	Systematic enhancement of timer tuning and channel selection	Large overhead, limited improvement	*<50ms*
Procedure optimization	Merge overlapping procedures in different authentication mechanisms	Simple and applicable	Limited improvement on current mechanisms	*<250ms*

Based on the survey on the related work, we observe that the early research on Fast Handoff schemes mainly compare with each other in terms of handoff delay, which has been significantly reduced from over 1000ms to a minimal level, e.g. a few milliseconds. Therefore the above performance comparisons become self-explanatory. However we also observe that no significant outcome has been addressed in literature to further achieve optimal performance of Secure Fast Handoff in term of end-to-end delay which varies from a few seconds to 500ms due to the complicity of security mechanisms on system level.

Table 3 Comparison of authentications for secure fast handoff

Authentication				
Offline Authentication	Online Authentication			
CA(Certification Authority) < *n seconds*	Distributed Approaches		Centralized Approaches with AS (Authentication Server)	
	Distributed without AS <*n seconds*	Distributed with AS < *n seconds*	Reactive	Proactive
			802.11i *<1s*	802.11i Pre-authentication

				500ms

Following the empirical approach of HAWK project, we aim to provide a solution which combines mechanisms of fast handoff with authentication for secured Internet access, namely Network-assisted Radio Signature and Dual Re-authentication.

4.3 Secure Fast Handoff Solution

4.3.1 Network-assisted Radio Signature (NRS) Scheme for Layer 2 Fast Handoff

4.3.1.1 Motivation

If we look into the handoff schemes, it is unobvious that there are certain assumptions behind their common practice: first the design philosophy beneath all probing schemes is that the network topology keeps changing, so that the topology information needs to be collected in an on-the-fly manner and keeps updating; secondly, the schemes are mainly client-initiated, i.e. the handoff intelligence is employed at MC side instead of network side, which means the network intelligence has not been utilized. We notice that such assumptions do not always hold when it comes to a WMN in the real world.

Hence, we argue that if the network topology is relatively static, the handoff decision can be made not only in a predetermined manner but also stand long. The delay caused by probing can be totally avoided and the collected topology information can be re-used as long as the mesh routers remain unchanged in the life cycle of the deployment. In order to proactively collect the predetermined network topology information, i.e. candidate APs for handoff, we decouple a training phase from the operational phase in the life cycle of the WMN. In order to comply with the fact of single radio for most MCs (otherwise the schemes which utilize extra radios like MultiScan [38] could be applied), multiple APs have to be coordinated to monitor and manage the handoff. Based on the above arguments and analysis, we present the details of NRS scheme in next section.

4.3.1.2 Overview

The fundamental design philosophy of the proposed NRS scheme is that since the effort of continuous and proactive probing deals with highly dynamic network scenario which is not necessarily right in all occasions, we tradeoff such an advantage to gain the improvement of handoff performance.

As illustrated in Figure 34, when a WMN equipped with NRS is put into operation, an MC moves to a certain location in the network. The most preferred AP to associate for the client is predefined in the training phase. The radio characteristics of the client are monitored by multiple APs: e.g. AP_1, AP_2, AP_3, in the network. For an instance, the on-the-fly measurement of the Received Signal Strength Indicator (RSSI) is transmitted by the MC. It constitutes the radio signature of the MC on the scene. The mapping relationship between the preferred AP and the corresponding radio signature of an MC is stored in a knowledge base in advance. Thus the knowledge base is looked up to find the most matched radio signature stored against the radio signature monitored by the APs.

According to the match, a corresponding AP, which is also the preferred AP to associate with, is determined. After the lookup, the handoff is initiated and managed by the network side. In this manner, handoff delay is totally eliminated since the MC does not probe for candidate APs anymore when handoff occurs.

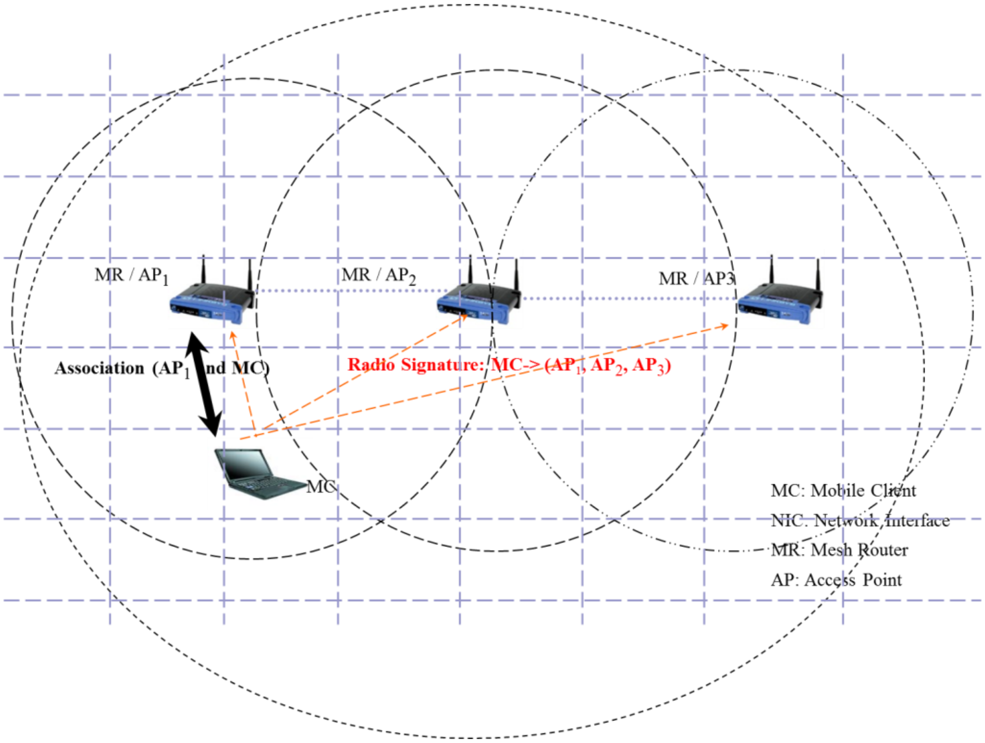

Figure 34 Illustration of NRS Scheme

How to build such a radio signature knowledge base? In line of the operational phase, the MC needs to traverse the coverage area of the network to collect topology information in another training phase. The coverage area is separated into subarea with certain granularity. The MC is used for measurement in each subarea one by one. In a subarea, the radio signal transmitted by the MC is recorded, processed and stored in the knowledge base by AP_1, AP_2, and AP_3. The most preferred AP is arbitrary and determined by the designer of NRS system according to specific application requirements, for an instance, the AP which provides highest data rate by the actual measurement. This implies that the most preferred AP can be either of AP_1, AP_2, or AP_3, but not necessarily be one of them. After the preferred AP is determined, the mapping between this AP and the radio signature is recorded into the knowledge base.

4.3.1.3 NRS Operations and Algorithms

A. Training Phase

The objective of the training phase is to build the radio signature knowledge base. In practice, we use a notebook PC which is equipped with an 802.11 STA to traverse the subarea of the WMN. The radio characteristics of the STA, which is the so-called radio signature, is measured, processed and stored by the functional entities resided on APs. As in Figure 36, the process is illustrated step by step as follows:

Step 1: The most preferred AP is determined by a predefined metric, e.g. data rate, RSSI, or Signal Interference Ratio (SIR). To simplify the implementation, we adopt RSSI as the indicator. It is noticeable that the implication of this RSSI is different from the RSSI used as radio signature. We aim to establish the relationship:

$$max\{RSSI_1; RSSI_2; \ldots; RSSI_i\} \rightarrow AP_{preferred}$$

```
{ //MC traverses the network area
 for each subarea $A_i \in Area_{network}$ A do
 {
  for each $AP_{i,} \in AP_{NRS}$ do
  { //Monitor and calculate the real time radio signature.
     //Process the radio signature with application-specific
fuzzy functions.
   $\mu_{APi\_RSSI\ high} = f_{fuzzy}(AP_i\_RSSI);$
   RS = ($AP_1$_RSSI, high,) $\mu_{AP1\_RSSI\ high}$), ($AP_2$_RSSI,
   high, $\mu_{AP2\_RSSI\ high}$),…, ($AP_i$_RSSI, high, $\mu_{APi\_RSSI\ high}$)
   // Associate RS with the most preferred AP according to
   data rate or RSSI etc., write in database KBD.
  }
```

Figure 35 The algorithm of NRS training stage

Step 2: The area of 802.11 WMN is divided into m*n subareas: A_1 ,A_2 ,…, A_i.

Step 3: In the subarea A_i, manually associate MC with AP_1, …, AP_i, measure the RSSI, to determine the most preferred $AP_{preferred}$. e.g. when max{$RSSI_1$, $RSSI_2$,…,$RSSI_i$} is fulfilled, MC is associated with AP_i, then we determine AP_i is the preferred $AP_{preferred}$.

Step 4: In the meantime, measure the radio signature of the MC, e.g. RSSI of MC at AP_1, …, AP_i, e.g. AP_1_RSSI=-60dbm, AP_2_RSSI=-70dbm, …, a multiple tuple (AP_1_RSSI, AP_2_RSSI, … AP_i_RSSI) is constructed to represent the raw numeric value of a radio signature.

Step 5: We further process the radio signatures by employing functions according to specific application requirements. We apply fuzzy functions in our implementation since the fuzzification deals with the vagueness of measurement introduced by granularity of subareas, e.g. Radio Signature Process Function$_1$:

$$F_{fuzzy}(\text{AP}_{\text{i}}\text{\_RSSI}) = \begin{cases} 0 & \text{AP}_{\text{i}}\text{\_RSSI} < 10 \text{ dB} \\ \mathbf{1} + \log \frac{\mathbf{APi\_RSSI}}{\mathbf{10}} & \text{AP}_{\text{i}}\text{\_RSSI} \geq 10 \text{ dB} \end{cases} \tag{4.1}$$

Without loss of generosity, we use $\mu_{\boldsymbol{AP1\_RSSI\,high}}$ to denote the numerical result of the function.

$f_{fuzzy}(\text{AP}_1\text{\_RSSI}) =: \mu_{\boldsymbol{APi\_RSSI\,high}}$

$f_{fuzzy}(\text{AP}_2\text{\_RSSI}) = \mu_{\boldsymbol{AP2\_RSSI\,high}}$

……

$f_{fuzzy}(\text{AP}_{\text{i}}\text{\_RSSI}) = \mu_{\boldsymbol{APi\_RSSI\,high}}$

Step 6: We combine the outcomes and the original radio signatures to build a knowledge base.

(AP$_1$_RSSI, high, $\mu_{\boldsymbol{AP1\_RSSI\,high}}$),

(AP$_2$_RSSI, high $\mu_{\boldsymbol{AP2\_RSSI\,high}}$),

……

(APi_RSSI, high, $\mu_{\boldsymbol{APi\_RSSI\,high}}$).

e.g. (AP$_1$_RSSI, high, $\mu_{\boldsymbol{AP1\_RSSI\,high}}$) means the signature of MC heard by AP$_1$ is = 0.8.

Step 7: We have an entry in the knowledge base for the subarea A$_{\text{i}}$, which is

f: KB(A$_1$) -> AP$_{\text{preferred}}$, e.g. {(AP$_1$_RSSI, high, 0.8), (AP$_2$_RSSI, high, 0.7), (AP$_2$_RSSI, high, 0.5)**)}** -> AP$_{\text{preferred}}$

Step 8: We move MC to the next position A_2, repeat the above procedures. After MC traverses the whole area, the knowledge base is built. e.g. KBD = {KB(A_1), KB(A_2),…, KB(A_i)}, ($1< i<n$).

B. Operational Phase

Basically, the operational phase is the reversal process of the training phase. In the operational phase, the real-time radio characteristics of the STA is monitored and processed by APs, which is considered as the radio signature on the "crime scene", the captured radio signature will be compared with the radio signature in the knowledge base. The similarity between the captured one and the records will be measured and the matched radio signature will be identified. In case of no identical match, the most similar one of all records will be identified. Corresponding to the identified radio signature, the most preferred AP to be associated is determined. By this means, the handoff is achieved proactively and without any on-the-fly operation of STA. As illustrated in Figure 36, the algorithm of the operational stage is listed below:

```
{ //MC moves to a sub-area of the network coverage
 for each sub-area subarea $A_i$ $\in$ Area do
 {
   for each $AP_{i,}$ do
   { //Calculate the real time radio signature.
   RS(current) = ($AP_1$_RSSI, high, $\mu_{AP1\_RSSI\ high}$),
   ($AP_1$_RSSI, high, $\mu_{AP2\_RSSI\ high}$), …, ($AP_i$_RSSI,
   high, $\mu_{APi\_RSSI\ high}$)
    }
 }
  //Look up and compare with each entries in the
knowledge base and select the one match.
   for each $KB_i$ $\in$ KBD do
   { //Deal with user preference and compare the real
time measurement with the stored radio signature.
      if (exist new preferred APj)
      then   select new APj.
      else
        if RS (current) ($A_i$) = KB($A_k$)
        {     $A_k$ is the $AP_{preferred}$
       } } //end of operation
```

Figure 36 The algorithm for operational stage

Step 1: When the 802.11-based WMN is put into operation, MC moves in the network. AP_1, … , AP_i, monitor the movement of MC.

Step 2: At given time I, AP_1, … , AP_i measure the radio signature of MC, e.g. when MC is in subarea AP_1, the measurement is (AP_1_RSSI, AP_2_RSSI, ・ ・ ・ , AP_i_RSSI).

Step 3: We apply the same fuzzy functions as those employed in the training stage for further processing, e.g. RS (current) = {(AP_1_RSSI, high, 0.5), (AP_2_RSSI, high, 0.5), (AP_i_RSSI, high, 0.7)}.

Step 4: The radio signature is looked up in the knowledge base, calculate the match degree between the radio signature on the scene and the entries in the knowledge base. If we discover: e.g RS (current) (A_1) = KB(A_1), which is, {(AP_1_RSSI, high, 0.8), (AP_2_ RSSI, high, 0.7); (AP_2_RSSI, high, 0.5)} → $AP_{preferred}$.

Step 5: We conclude that AP_1 is $AP_{preferred}$, then instruct the MC to handoff to the AP_1. The probing delay is eliminated until this phase. However, after the preferred AP is determined, the authentication still takes time to grant the Internet access for the MC. We further present the solution in the next section.

4.3.2 Dual Re-authentication Mechanism for Secure Handoff

4.3.2.1 Overview

As analyzed in the related work, most recent efforts of fast authentication focus on proactive approaches by transferring security context to potential target mesh routers before handoff occurs. However, the efficiency and effectiveness of such approaches are significantly affected by the density of potential targets and the accuracy of handoff prediction which is in fact a very difficult task. So we propose a reactive scheme to reduce re-authentication delay which consists of two steps to exploit the tradeoff between system performance and security. The fundamental design philosophy of our proposed mechanism is Access-before-Authenticate, which means to grant the user an “immediate” access based on a lightweight authentication while a strong authentication is executed afterwards. The original authentication process is separated to two phases, namely Immediate Authentication (IA) and Full Authentication (FA). With IA an MC has a particular evidence which can prove it have associated with a trusted AP within a granted period. After that, the new AP allows the MC to access the network temporarily for a certain time of duration. The design of evidence makes it to be validated quickly without any assistance from remote Authentication Server (AS). Full Authentication (FA) is executed after that. If the strong authentication is failed, the access will be terminated.

4.3.2.2 IA and FA Procedures

To simplify the implementation, our scheme is based on IEEE 802.11i standard which adopts IEEE 802.1X framework. The FA phase adopts standard IEEE 802.1X authentication. The IA phase uses One-Time Ticket (OTT) as the authentication evidence.

A. One-Time Ticket

OTT provides a method to validate whether an MC has associated with another AP dominated by the same AS before.

In the IEEE 802.11i, when an MC entries the network initially, the MC should conduct IEEE 802.1X authentication to decide whether data access is allowed or not. In our scheme, an OTT is delivered from AS to the MC accompanying the initial IEEE 802.1X authentication. When the MC leaves the current AP’s domain, the MC sends its OTT to new AP through OTT-request message. The new AP can compare the OTT with the entries of OTT cache table which it maintains. If the OTT is valid, the new AP trusts the MC for a certain while and initiates a four-way handshake to establish the session keys used for traffic encryption. Otherwise, AP notifies the network access request is denied through OTT-response. The MC has to conduct IEEE 802.1X authentication and another four-way handshake within a given duration contained in the received OTT-response. The OTT is established by hash chain [43].

Let $H^n_{s,k}$ denote a set of hashed keys established by a seed, s. k is a secret key, and n is the number of hashed keys. $H^n_{s,k}$ is formulated by

$$H^n_{s,k} = \{h_i | h_i = HMAC_k(h_{i-1}), h_0 = s, 0 < i \le n\} \tag{4.2}$$

where $HMAC_k$ is a keyed-Hash Message Authentication Code. h_i is an OTT value, and bound to the time when the MC is guaranteed. For example, if one OTT is guaranteed for l sec, $H^n_{s,k}$ is valid for $n \times l$ sec. If an MC tries to associate with an AP at time t_i, the corresponding OTT is h_j, where $j = n - \lfloor (t_i - t_s) \times l^{-1}$ and t_s denotes the time when the seed was distributed. Since $H^n_{s,k}$ is bound to time, AP$_s$ and AS should be time synchronized.

B. Immediate Authentication Procedure

A seed *s*, randomly chosen by AS, has to be shared between authorized APs and AS in advance, *s* can be distributed through two modes: *passive* and *active*. In the passive mode, AS broadcasts the seed to all authorized APs regularly through *Seed-advertisement* message (P1 in Figure 37). *k* is used for securing hash function. Particularly, t_c is the current time stamp used for time synchronization between AS and the authorized AP. In the active mode, when an AP boots up, the AP requests *s* from the AS through *Seed-solicitation* message.

The AS replies the AP through *Seed-advertisement*. IEEE 802.11i assumes that AP and AS can established a secure channel such as IEEE802.1X, we can assume *s* is delivered securely. The *Seed-advertisement* contains seed number *s*, the number of keys *n*, unit time *l*, usage starting time of the seed number *ts*, and OTT valid duration time t_v.

Since *s* and *k* are shared at both APs and AS sides, they can generate common outcome of $H^n_{s,k}$ individually. When an MC attempts to access the network initially, IEEE802.1X authentication and four-way handshake are conducted consequently. While the EAP procedure at the AAA server such as RADIUS server, AS sends the MC the current time t_r and an appropriate OTT (P2 in Figure 38). It is the same as IEEE 802.1X authentication and four-way handshake procedure except an OTT attached inside.

The MC should keep the received OTT and t_r. When the MC leaves from the current serving AP, the MC starts to probe available wireless channels and associates with a new AP, AP_n. In order to reduce the re-authentication delay, the MC sends its OTT to AP_n

through *OTT-request* (P3 in Figure 38). Since the wireless channel between MC and AP is insecure, { $f_h(t_r||ID||\mathrm{OTT})$, t_r, *ID*} are transmitted from MC to AP_n as an *OTT-request*, where f_h is an one-way hash function such as MD5 and SHA-1. *ID*, and the MC's identity such as MAC address, are used for preventing attacks such as replay and man-in-the-middle.

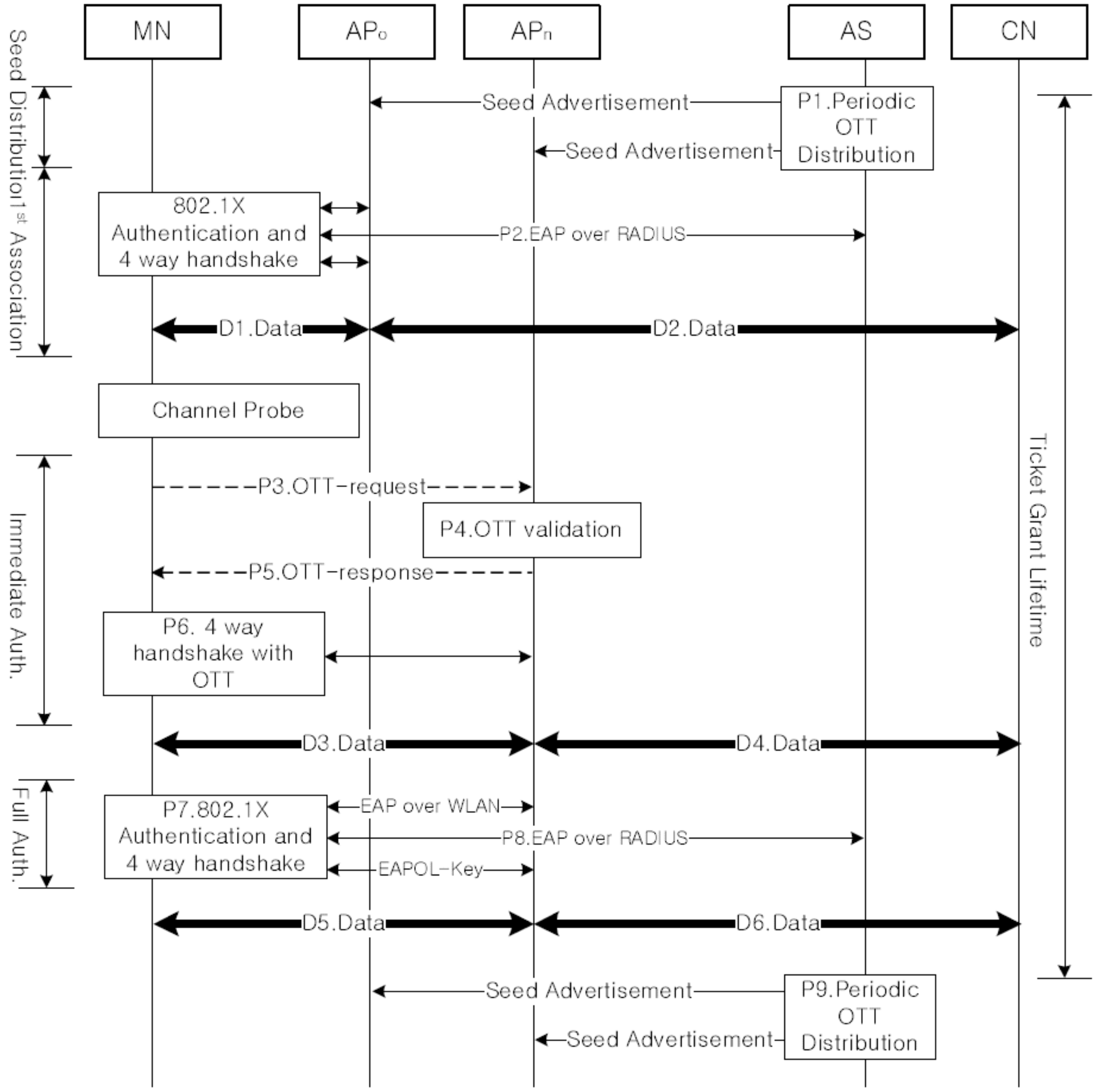

Figure 37 Proposed Dual Re-authentications scheme

When receiving the *OTT request*, *AP_n* is able to check immediately whether the MC has ever associated with other AP_s also dominated by the same AS through comparing the validation of received OTT to the OTT cache table maintained itself (P4 in Figure 38). If the OTT from the MC is valid, *AP_n* notifies MC's admission through *OTT-response* (P5 in Figure 37), and initiates a four-way handshake to establish session key used for data traffic encryption (P6 in Figure 37).

The session keys are derived from the pairwise master key (PMK), Authenticator address (AA), Supplicant address (SPA), Authenticator nonce (AN_{once}), and Supplicant N_{once}(SN_{once}). IEEE 802.11i also specifies a 256 bit Pre-Shared Key (PSK) shared beforehand between AP and MC, which derives PMK through Pseudo Random Function (PRF) without AS. IA uses the OTT as a PSK. Consequently, IA requires no modifications to the four-way handshake procedures in IEEE 802.11i. After P6 is completed, the MC is able to access the network on secure channel (D3) established by the procedure P6. By that way, the re-authentication delay of our report IA is only from P3 to P6. Within these procedures, there are no information exchanging required therefore packets exchanged with *APo* or AS so far. Therefore, no packets transmission is needed. Conventional IEEE 802.1X authentication and four-way handshake will be accomplished within a given expired time for restraining IA's optimism (P7 in Figure 37). After that, OTT of MC is updated (P8 in Figure 37). AS should broadcast *Seed-advertisement* within the Ticket Grant Lifetime (P9 in Figure 37).

C. OTT Validation

AP should maintain OTT-cache table which contains valid OTTs. OTT entries of the cache table is formulated by

$$\widehat{H_{t_r,ID}} = \{\xi_i | \xi_i = f_h(t_r \| ID \| h_i), \\ h_i \in H^n_{s,k}, \zeta_l \le n - i < \zeta_h\}, \tag{4.3}$$

where $\max(0, \lfloor(\widetilde{t_c} - t_v - t_s) \times l^{-1}\rfloor), \zeta_h = \lfloor(\widetilde{t_c} - t_s) \times l^{-1}\rfloor$ and $\widetilde{t_c}$ is the current time. Let h_o be $f_h(t_r \| ID \| \mathrm{OTT})$ contained in *OTT-request* message. If$h_o \in \widehat{H_{t_r,ID}}$, this *OTT-request* is valid, thus AP notifies MC is allowed to access the network through *OTT-response* (P5 in Figure 37). *OTT-response* contains the time (allowed FA guaranteed duration) within which indicates the MC must accomplish conventional IEEE 802.1X authentication and four-way handshake. Otherwise, it is probably that the MC is not a normal user such as an abused reusing cached old OTTs (replay or man-in-the-middle attacks). The AP can adjust an allowed FA time bound to correspondent OTT in order to mitigate the effect of OTT abuse. Even if OTT is revealed to attackers, the access of the network is only within the allowed FA guaranteed duration.

4.4 Experiments and Evaluation

In order to investigate the feasibility and effectiveness of the proposed mechanisms, we implement them in our HAWK integrated experiment environment which is consisted with prototype testbed and simulation tools to test and evaluate the system performance:

4.4.1 Layer 2 Fast Handoff Delay

We compare the end-to-end delay of a VOIP application considering two cases without and with our proposed NRS scheme in the same simulation scenario. A MC roaming through three neighbouring APs connecting though wireless mesh backbone. Three APs collaboratively monitor the radio signatures of MC, the signature is compared with the entries recorded in the knowledge base which has been prepared. The preferred AP is calculated and stored in the list for MC to associate in case handoff will be initiated. We run a VOIP application on the local MC and the remote correspondent. We arbitrarily make MC to handoff by setting handoff threshold and observe the time delay. We keep measuring the end-to-end delay of packets between MC and the remote correspondent. When we manipulate Layer 2 handoff to occur by moving MC, the end-to-end delay will vary and indicate the start and the end of the handoff. Referring to Figure 38, we can see that our proposed NRS scheme can reduce the time-consuming scanning delay during the handoff procedure greatly. Under the normal circumstance without NRS scheme, the delay is around 70ms. By enabling NRS scheme, the total handoff delay is shortened to 10ms, which are accumulated by open authentication and association excluding probing. It is obvious that the proposed mechanism works. However, the accuracy of the handoff becomes our concern. The accuracy issue of NRS schemes majorly comes from the dividing the area into subareas. The granularity of the subarea determines whether a handoff occurs according to the training phase. If a handoff occurs in the operational phase while it is not expected in the training phase, we consider it inaccurate. In general, smaller the granularity is, more unnecessary handoff occurs. Nevertheless, if the granularity is too large, it loses the meaning of handoff in the first place and only an ideal

AP with the unlimited coverage can support that. We use false negative (FN) and the false positive (FP) rate to represent the accuracy of the proposed NRS scheme. We apply the two-ray ground reflection model as the wireless channel propagation representation in our experiments. We set the size of the whole area is $100m \times 100m$. As the size of the subarea which is a square as the basic union of territory increases, the FN and FP go to their theoretical values fluctuating around 21% and 36% respectively. We can notice that when the size of subarea goes over 10m, the rate of FN and FP are relatively stable which can be set as a recommended system parameter of our proposed NRS scheme in implementation.

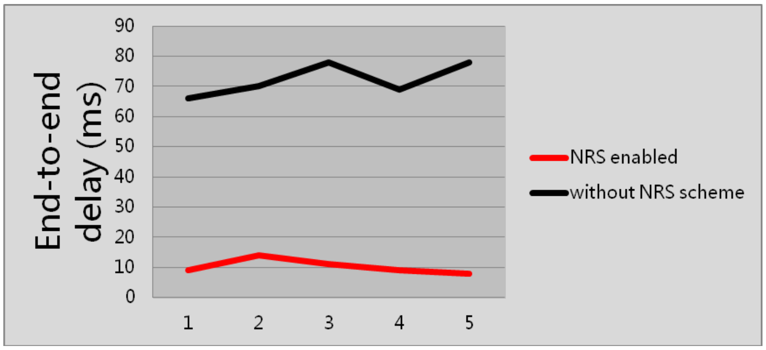

Figure 38 The latency due to handoff delay without vs. with employing NRS Scheme

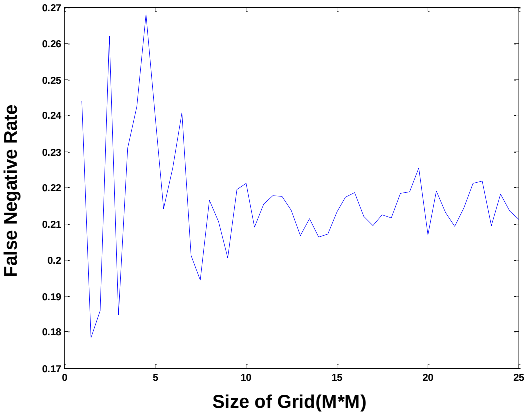

Figure 39 The False Negative Rate of Handoff Decision

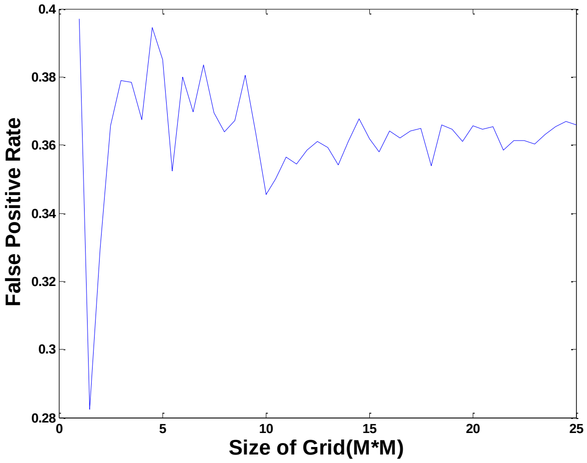

Figure 40 The False Positive Rate of Handoff Decision

4.4.2 Dual Re-authentication

This section provides performance analysis of the proposed scheme in terms of authentication delay. We also make a quantification model to measure the degree of security for our report.

Handoff Delay

Let T_r denote the response time from the point when re-association is initiated to the point when four-way handshake is completed. T_r is formulated by

$$T_r = T_{reassoc} + T_{auth} + T_{fwh} \tag{4.4}$$

where $T_{reassoc}$ is a delay for re-association, and T_{fwh} is a delay for four-way handshake. In IEEE 802.11i, T_{auth} is a delay for IEEE 802.1X authentication. In the proposed scheme, T_{auth} denotes operation time for OTT validation in Section II. C. T_{auth} is determined by tv

and *l*. The total possible number of keys assigned to each MC per AP is formulated by _ $\eta = \frac{t_v}{l}$. Since T_{auth} is the time to conduct a linear searching based on hashing and comparing, T_{auth} is stochastically proportional to η. The OTT-Validation takes $O(\eta)$ times. The expectation of T_{auth} is defined as:

$$E[T_{auth}] = (\eta \times T_h)/2 \tag{4.5}$$

where T_h is the time for f_h computation process time. Our analysis adopts SHA-1 as f_h. T_h is measured by Benchmarks of Crypto++ 5.5 Library [51]. Conclusively, T_r, re-authentication delay in our scheme can be represented as the function: $f(\eta) = T_r$,where $f(\eta)$ is the re-authentication delay function. Given that delays for re-association, IEEE 802.1X, and four-way handshake average 2, 250, and 60ms, respectively, re-authentication delay in the conventional IEEE 802.11i is totally 312ms. 250ms for IEEE 802.1X authentication is measured in fast resume mode [50]. Although full IEEE 802.1X authentication takes 1000ms, we compares with the fixed version for unbiased comparison.

Quantification for the Degree of Security

As *l* is increased, the probability that the same OTT is assigned to a number of MCs are increased. Let λ_a denote the total arrival rate, .i.e., the total association rate. λ_a is defined as $\lambda_1 + \lambda_2$. Let μ_d denote the total departure rate. μ_d is defined as $\mu_1 + \mu_2$. Let $P_l(N = n)$ denote the probability that the number of associated MCs are *n* within a period *l*. The probability that the same OTT is assigned to a number of MCs before handoff is equal to the probability that at least two MCs arrive within a period *l*, i.e., $P_l(N \geq 2)$.

$$P_l(N = 0) = 1 - \lambda_a e^{-\lambda_a l} - \int_0^l \lambda_a e^{-\lambda_a t} \times \mu_d e^{-\mu_d (l-t)} dt \tag{4.6}$$

where the integral part indicates the case that an MC leaves within a period $l - t$, even though arrives at $t(\leq l)$.

$$P_l(N = 1) = \lambda_a e^{-\lambda_a l} \tag{4.7}$$

$$P_l(N \geq 2) = 1 - P_l(N = 0) - P_l(N = 1)$$

$$= \int_0^l \lambda_a e^{-\lambda_a t} \times \mu_d e^{-\mu_d (l-t)} dt \tag{4.8}$$

The duplicated OTT problem during initial association does not threaten the security of network system critically, because OTT is used after handoff. The case of after-handoff duplication should be considered. If MC_0 moves into AP_n, the probability that MCs assigned the same OTT to MC_0 exist is:

$$p_d = 1 - \left(1 - \frac{1}{\eta}\right)^{m_n} \tag{4.9}$$

This implies a trade-off between OTT acquisition speed and security. The reduced re-authentication delay decreases the degree of security measured by *pd*, comparing with original IEEE 802.11i. Additionally, even though OTT is revealed to a later-coming MC, since the hash function for OTT chain generation varies depending on a secret key *k*, the hash chain is difficult to be derived by an adversary. It shows the relationship between p_d and MC. The black dot indicates that η is 180. Likewise, the white dot indicates η is 360. In Figure 41, we find out that the probability of OTT duplication is decreased faster, as η

_ increasing. p_d is also defined as a function as follows: $g(m_n, \eta) = p_d$, where $g(m_n, \eta)$ is a security measure function.

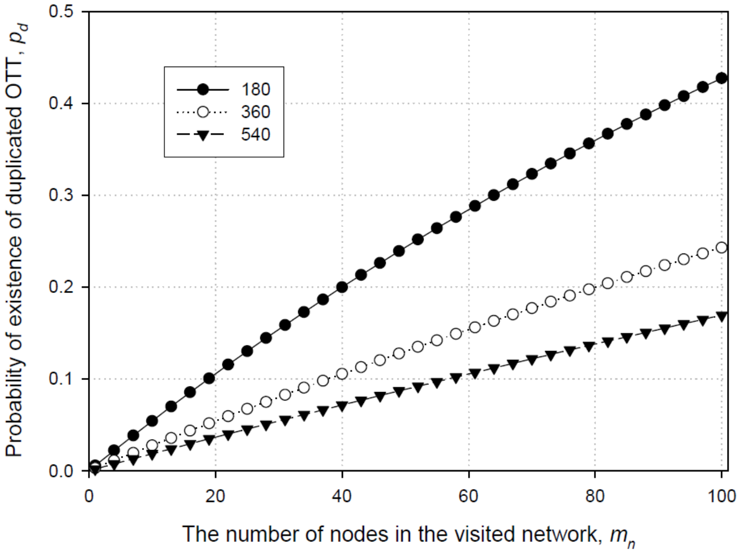

Figure 41 Probability of existence of duplicated OTTs after handoff

4.5 Discussion on Limitations

In our previous work of Network-assisted Radio Signature (NRS) probing, we employ multiple APs to monitor a MC and produce a vector of Received Signal Strength Indications (RSSIs) as radio signature of a MC corresponding to the most appropriate AP. It is an arbitrary handoff mechanism which completely eliminates the on-the-fly scanning procedure together with the handoff delay whereas also introduces the complexity of a training stage prior to network operation and the non-scalability of fixed deployment [10].

Chapter 5 Enhanced Approach of Network-assisted Radio Signature for Fast Handoff in AP-dense Environment

5.1 Introduction

Recently the deployment of IEEE 802.11 Wireless Mesh Networks (WMNs) has been growing exponentially because of emerging of sharing economy of Wi-Fi. As a result, the handoff delay caused by mobile client (MC) probing for the next preferred AP becomes more challenging in a dense WMN. In this paper, we propose a fast handoff scheme T-NRS (Temporal-NRS) which leverages historical knowledge of APs associated in time series to assist in handoff decision in addition to the Network-assisted Radio Signature(NRS) technique based on spatial knowledge. The enhancement improves the performance whilst eliminates the inflexibility. We implement and evaluate the mechanism in about 200 APs in shopping malls. The results prove the performance of the proposed scheme.

5.1.1 Motivations

Provision of pervasive wireless Internet access to mobile users anywhere has been a hectic driving force for industry and academic community. Recently the deployment of IEEE 802.11-based Wireless Mesh Networks (WMNs) has been growing exponentially in many places like chain stores and shopping malls due to the ermerging of sharing economy of Wi-Fi, so as to deliver value-added mobile applications such as advertising or e-commerce etc. to nearby smartphone users.

Consequently, time-sensitive mobile APPs on smartphones face more frequent handoffs with dense access points (APs) and mobile clients (MCs) in the network, which imposes severe challenges upon fast handoff mechanisms in such scenarios [66].

As previously illustrated in Figure 1, the handoff issue involves interruptions of multiple dimensional operations of the network from physical layer on the bottom to application layer on the top. We address on Layer 2 handoff issue in this paper due to which contributes the majority of the delay. The existing handoff mechanisms can be classified into three categories: the legacy handoff, the terminal-based handoff and the network-assisted handoff. The legacy 802.11 handoff operations are comprised of three procedures as shown in Figure 2: namely scanning, authentication and re-association. The scanning is that a MC collects information about neighboring AP and chooses the most appropriate AP to handoff. The authentication procedure is to authenticate and authorize the access to the AP. And the association procedure is to establish the association with AP. The scanning is the most time-consuming and left for manufacturers and researchers to explore.

Terminal-based approaches of fast handoff have been thoroughly explored at early stage since the terminal is convenient to tackle with rather than network infrastructure. They manage how a MC switches channels, listens for beacons sent by APs passively, or actively sends probe requests and receives probe responses from APs, etc. Nevertheless, it is constrained by limited capacity of terminals and has the problem of forward-compatibility of existing terminals.

Network-assisted approaches orchestrate multiple APs to work collaboratively to probe and produce knowledge for handoff. Multiple APs listen for beacons from a MC rather than the original opposite design. Specialized centralized or distributed mechanisms are implemented on APs. In the meantime, network-assisted approaches are welcome and embraced by Wi-Fi operators because it solves the practical problem of compatibility of legacy terminals on the market.

In our previous work, we proposed a fast handoff scheme Network-assisted Radio Signature (NRS). The technique employs multiple APs to collaboratively monitor a MC and produce a tuple of Received Signal Strength Indications (RSSIs) as the radio signature. The correlation between radio signatures and the most preferred AP is established in a knowledge base by using a MC to traverse the network and keeping a record of all the measurements in a training phase prior to network operation. Thus in the following operational phase, the handoff can be decided reversely by comparing the real-time measurement of radio signatures and the knowledge base. The NRS scheme completely eliminates probing and the relevant delay whereas introduces an extra training phase which causes the problem of non-scalability of the fixed deployment of APs [50].

5.1.2 T-NRS Scheme: Enhancement based on Temporal Knowledge of Next APs

In this section, we consider the abovementioned disadvantage and further propose an enhancement of the scheme based on temporal knowledge of next APs which is named as T-NRS (Temporal-NRS) scheme. It is featured by recording the handoff decision, i.e. the next AP to associate along with radio signatures, of the legacy handoff mechanism when the network is in operation. The correlation between and the most preferred AP and radio signatures is established in a spontaneous manner when a MC moves in the network. Therefore the knowledge base of mapping relationship between radio signatures and the next AP is built adaptively.

At the beginning of operation of network, the legacy handoff is employed because T-NRS is invalid since it works under the condition that the proper historical knowledge is previously recorded. But T-NRS becomes fully operational in the end after MCs have traversed the network with a learning-while-working process rather than the standalone training phase in the original NRS approach.

Moreover, the T-NRS scheme improves the approach in an AP-dense network by recording the APs associated after each handoff. The probability of the next AP to associate in time-series is taken into account based on the temporal information in order to improve the handoff accuracy of the radio signature approach.

We observe that the scale and density of 802.11 WMNs has been increasing exponentially. In an urban area a MC received radio signals from more than 10 APs most of the time as in [26, 34, 31] but nowadays there are over 30 SSIDs are accessible for a MC and an AP serves over 50-70 smartphones in rush hours as in our controlled scenario of shopping malls. We implement the T-NRS scheme in our testbed and it turns out that at least three RSSIs are always available so that the scheme works most of the time.

5.1.3 Contributions

The contributions of this chapter are summarized as follows:

- We propose an enhanced fast handoff mechanism T-NRS based on our previous work. The new scheme eliminates the standalone training phase and achieves the same purpose adaptively and gradually in the course of network operation, by utilizing the mobility of all MCs spontaneously. Temporal information about next APs associated is taken into account to calculate the probability of the next AP. Both spatial and temporal information are utilized to improve the handoff accuracy.
- To the best of our knowledge, we are the first to combine and utilize the characteristics of dense APs and MCs and the temporal information of handoff to improve the performance.
- We evaluate the solution with experiments in a testbed. The results prove that the performance of the mechanism satisfied the requirements of current popular mobile applications.

5.2 Temporal Network-Assisted Radio Signature Scheme

5.2.1 Motivation

We observe that there is an strong assumption behind the conventional fast handoff mechanisms which is the network infrastructure topology would keep changing, so that probing for next APs is carried out in an on-the-fly manner. The assumption is not necessarily true therefore we proposed NRS scheme to decouple a training phrase to establish the relationship between radio signatures and the preferred AP, prior to the operational phase of the network.

In the operational phase, a MC moves with random-walk in the network and the handoff decision is made reversely by measuring the real-time radio signatures i.e. RSSIs, looking up in the knowledge base of the relationship, and calculating the probability of the next AP.

Nevertheless, it is not cost-efficient to afford this training phase before the network operates in all scenarios. Moreover, once APs change their position in the life cycle of the deployment, the NRS scheme has to go through training phase again. So we propose an enhancement of the scheme Temporal-NRS (T-NRS), which is to collect the information of relationship between radio signatures and the next APs while the network is in operation. The dedicated training phase is removed and the learning process is carried out regularly in order to build the knowledge base in a gradual and adaptive way. With multiple MCs assisted in the learning process in parallel, the time to complete the knowledge base is expected to be reduced in a MC-dense environment. Also it becomes flexible to the changes of the network topology.

We observe that since the fundamental of measurements of RSSIs of MC is to utilize the spatial information to assist in handoff and due to the highly dynamic nature of the radio signals, the handoff accuracy, i.e. the probability of the next AP could be further improved by introducing temporal information into handoff decision. In T-NRS scheme, we take the historic records of APs ever associated in time series into account. The probability of the next AP is independently calculated based on temporal information. With the combination of both, the accuracy of handoff is greatly improved while the cost of measurements is reduced.

T-NRS scheme is required to employ the legacy handoff mechanism initially and then becomes fully operational after a period of time. It is obviously a tradeoff between the instant effectiveness of the original NRS scheme and the flexibility of network deployment for the enhanced T-NRS scheme.

5.2.2 Network Scenario

As illustrated in Fig 42, a WMN with APs installed with T-NRS mechanism is in operation, APs are connected with each other via IAPP communications and connected with Internet. MCs including notebook PCs and smartphones move randomly in the network and are associated with APs to have access to Internet.

The T-NRS scheme cooperates with the legacy handoff scheme. At the beginning, the legacy scheme is initiated to build the knowledge base first. In the meanwhile, the T-NRS scheme continuously calculates the probability of the next AP and decides whether or not to switch between the legacy scheme and itself.

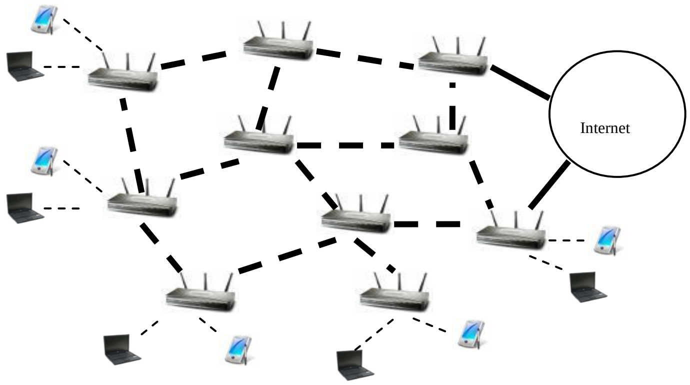

Figure 42 The illustration of the network scenario

5.2.3 Handoff Decision based on Measurement of Radio Signatures

The radio signature of a MC is measured by multiple APs: AP_1, AP_2, AP_3, … AP_j, i.e., the Received Signal Strength Indicators (RSSIs) transmitted from the MC. When the legacy handoff is completed, the mapping relationship between the preferred AP and the corresponding radio signature of a MC is recorded and stored into a knowledge base.

Definition 1: The mapping relationship is denoted as:

$\{RSSI_{(\mathrm{AP1})}, RSSI_{(\mathrm{AP2})}, \ldots , RSSI_{(\mathrm{APj})}\} \rightarrow \mathrm{AP}_{\mathrm{preferred}}$;

A learning process is carried out regularly and interweaved with the handoff process, so the knowledge base is built by multiple MCs simultaneously.

Definition 2: The radio signatures and the knowledge base are denoted as:

RS(MC$_1$)= （$RSSI_1^1, RSSI_2^1, \ldots, RSSI_n^1$) → AP$_1$;

RS(MC$_2$)= （$RSSI_1^2, RSSI_2^2, \ldots, RSSI_n^2$) → AP$_2$;

……

KBD={*RS*(MC$_1$), *RS*(MC$_2$), …, *RS*(MC$_i$)}, (1< i<n)

When a handoff occurs and seek for the next preferred AP, T-NRS scheme is initiated. First it calculates the probability of the next AP based on the comparison of real-time measurement of the radio signature and the historical records in knowledge base. If the probability is higher lower than the threshold which means the knowledge base does not meet the requirement, the legacy handoff is employed to decide the handoff and the result is recorded into the knowledge base. If the probability is higher than the threshold, the AP with the highest probability will be chosen and the MC will be instructed to connect with it. In this way, the dedicated training phase is eliminated and the learning process is merged with handoff itself.

Procedures: The detailed procedure will be illustrated step by step as follows:

Step 1: Though the most preferred AP can be determined by a predefined metric, we adopt RSSI as the indicator to simplify the implementation. We aim to establish the mapping relationship between the measurement of RSSI and the next AP. The largest RSSI indicates the next AP:

$$\text{Max}\ \{RSSI_1, RSSI_2, \ldots, RSSI_i\} \rightarrow \text{AP}_{\text{preferred}}$$

Step 2: Initially there is no record in the knowledge base. Regarding MC$_1$, when an 802.11 legacy handoff is initiated and the reactive probing is carried out to find the most preferred AP, the T-NRS scheme keeps the record of the multiple-tuple as the radio signature and add it into knowledge base:

$$RS_i = RS(\mathrm{MC}_1) = (RSSI_1^1, RSSI_2^1, \ldots, RSSI_n^1) \rightarrow \mathrm{AP}_1$$

Step 3: Regarding MC_1, MC_2,…, MC_n, the legacy handoff scheme is always initiated while T-NRS does not satisfy the requirement of handoff accuracy. We combine the records to build the knowledge base.

$$RS_1 = (RSSI_1^1, RSSI_2^1, \ldots, RSSI_n^1) \rightarrow \mathrm{AP}_1;$$

$$RS_2 = (RSSI_1^2, RSSI_2^2, \ldots, RSSI_n^2) \rightarrow \mathrm{AP}_2;$$

……

$$\mathrm{KBD} = \{RS_1, RS_2, \ldots, RS_i\}, (i = 1, \ldots n)$$

Step 4: When a handoff is initiated, the T-NRS scheme first looks up in the knowledge base and compare the current measurement, i.e. $RS_{current} = (RSSI_1^{cur} \ldots RSSI_N^{cur})$, and the historical records of radio signatures, i.e. $RS_i = (RSSI_1^1, RSSI_2^1, \ldots, RSSI_N^1)$, where $RSSI_i^{cur} (i = 1, \ldots N)$ are the RSSI value measured between the current MC and AP_i. All the APs that contain one or more RSSIs with the same set of APs of the $RSSI_{current}$. Assume AP_i is one of them, and one RSSI is RS_i.

The relative distance between $RS_{current}$ and RS_i is calculated:

$$D_i = \sum_{k=1}^{n} \frac{\left|RSSI_k^{cur} - RSSI_k^i\right|}{\left|RSSI_k^i\right|} \tag{5.1}$$

The $AP_i$with the minimum D_i will be selected as the next AP, where $i = (1, \ldots n)$.

Step 5: The above step will be repeated and applied to all the APs that contain the same set of APs with the $RS_{current}$. Without loss of generosity, we assume that we have three APs: $\mathrm{AP_i}$, $\mathrm{AP_j}$ and $\mathrm{AP_k}$, and the corresponding distances for these APs are D_i , D_j and D_k. Then we can calculate the probability of $\mathrm{AP_i}$ being the next AP as the below:

$$P_{measurement}(\mathrm{AP_i}) = P(i) = \frac{\frac{1}{D_i}}{\frac{1}{D_i}+\frac{1}{D_j}+\frac{1}{D_k}} \tag{5.2}$$

Step 6: As long as the probability is above the predefined threshold, T-NRS will choose the $\mathrm{AP_i}$ as the potential next AP to associate.

5.2.4 Handoff with Historical Information of Associations

The fundamental of handoff decision by measuring radio signatures is to calculate the probability of the next AP based on spatial information. Due to the highly dynamic nature of radio signals, we introduce the historical records of APs associated in time series to be taken into account to improve the handoff accuracy of T-NRS scheme.

Generally speaking, we utilize the statistics of the historical association information to obtain the probability of the next AP. For instance, in Figure 43, we illustrate the basic idea. After we have collected the historical association information, we know that if the current serving AP is AP$_1$, then its next AP can be AP$_2$, AP$_3$ and AP$_4$. Furthermore, we calculate the statistics for each of these three potential next APs. For example, if the current serving AP is AP$_1$, there are 100 traces in which AP$_2$ are the next AP, 50 traces in which AP$_3$ are the next AP, and 10 traces in which AP$_4$ are the next AP, then the probability that AP$_2$, AP$_3$, AP$_4$ being the next AP is 100/160, 50/160, 10/160, respectively.

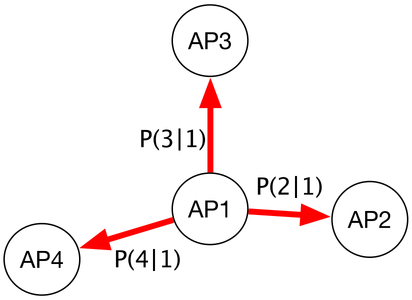

Figure 43 The probability of the next AP based on historical records

Step 8: Given the historical records of handoff collected in Step 3: regarding AP$_{1,}$ AP$_{2,}$ …… AP$_i$, the probability of the next AP based on the temporal information is calculated based on statistics of traces T$_j$:

$$P_{temporal}(AP_i \rightarrow AP_j) = \frac{N(AP_i \rightarrow AP_j)}{N(AP_i)} \qquad (5.3)$$

where $P_{temporal}(AP_i \rightarrow AP_j)$ is the provability that AP_j being the next AP of AP_i, $N(AP_i \rightarrow AP_j)$ is the number of traces that AP_j is the next AP of AP_i, and $N(AP_i)$ is the total number of traces containing AP_i.

Step 9: Combine the results of spatial calculation and the temporal prediction, ttherefore we have the final probability and choose the AP with the maximum value as the most preferred AP to associate.

$$P\ (\mathrm{AP_i}) = P_{measurement}(\mathrm{AP_i}) \times \mathrm{P_{temporal}}\ (\mathrm{AP_1}) \quad (5.4)$$

$$Max\{P(\mathrm{AP_i})\ \} \rightarrow \mathrm{AP_{preferred}},\ \text{where } i = (1, \dots n) \quad (5.5)$$

5.3 Evaluation

5.3.1 Testing Scenario

We deployed a total of 198 APs in two shopping malls in CBD of Shenzhen. The APs are all IEEE 802.11b/n 2.4GHz band. The hardware chipsets of APs include Atheros 9344 and MTK 9531. The topology of the whole 802.11 WMN is mesh networks connected by wired and wireless links distributed on different floors of the two shopping malls.

We implement the proposed scheme in the firmware of the APs. During these experiments, the deployed APs first collect the packets from the MCs nearby and update the database, and then predict the next AP that a MC should connect to when necessary. Note in these experiments, the system only records and does not actually control to which AP a MC should connect. In addition, the mobile phone of any person in the shopping malls can be the MC that contributes the data in the experiments.

5.3.2 Performance Criteria

We utilize the *handoff accuracy* as the measure to evaluate the performance. We define that if the next-hop AP predicted by the proposed scheme is the same with the one that a MC actually selected, then this handoff is regarded as correct. The handoff accuracy is then defined as the ratio of the number of correct handoffs to the total number of handoffs occurred during the experiment.

5.3.3 Performance Evaluation

We will show the performance of the proposed scheme during the two tests in a shopping mall in which 56 APs are deployed. Both of the tests lasted for 10 hours. Test 1 is carried out on a weekend from 11:00am to 10:00pm, and Test 2 is implemented on a weekday from 10:00pm to 9:00pm. In both tests, the system initially collects the information from the mobile phones in the shopping mall to establish the database, including the RSSI information, and the temporal association information. In the meantime, for any MC which initiates the handoff, the system will calculate, accordingly to the proposed scheme, the next potential AP to which it will connect.

Figure 45 shows how the handoff accuracy changes with time in the two tests. We can see that in Test 1, the handoff accuracy is as low as 50% in the first hour, and then starts to gradually increase to 91% at the end of the experiments. The initially low performance of the system can be attributed to the fact that the database that helps to predict has not been fully established. The results of Test 2 show the similar pattern as in Test 1. However, we can see that the initial accuracy of Test 2 is lower than Test 1. This is intuitively correct since during the weekend, the number of MCs is greater than that during the weekday of Test 2. As illustrated in Figure 45, it compares the total number of MCs that all 56 APs served during each hour of the two tests.

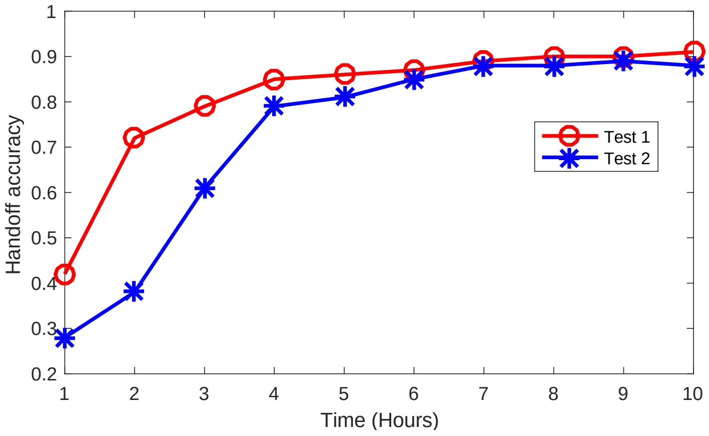

Figure 44 The handoff accuracy of experiments carried out in the shopping mall with 56 APs

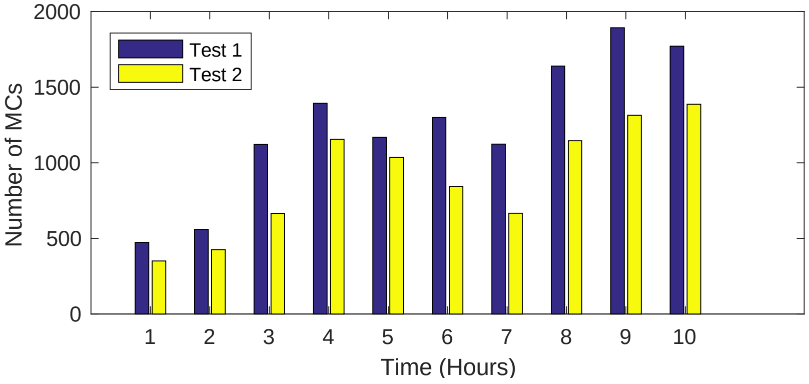

Figure 45 Total number of MCs that all 56 APs served during each hour of the two tests

To demonstrate the effect of using the RSSI information and the temporal association information, we also compare the performance of the system when these two types of information are utilized separately in Test 1. Figure 47 compares the handoff accuracy of the system when both the RSSI and the temporal association information are utilized, when only the RSSI information is utilized, and when only the temporal association information is utilized. We can see that when only the temporal association information is utilized, and the maximum handoff accuracy can only achieve about 85%, which is about 6% lower compared to when both information are utilized. In addition, when only the RSSI information is utilized, the handoff accuracy can only achieve about 76%. This indicates that the temporal association information is more informative or reliable than the RSSI information.

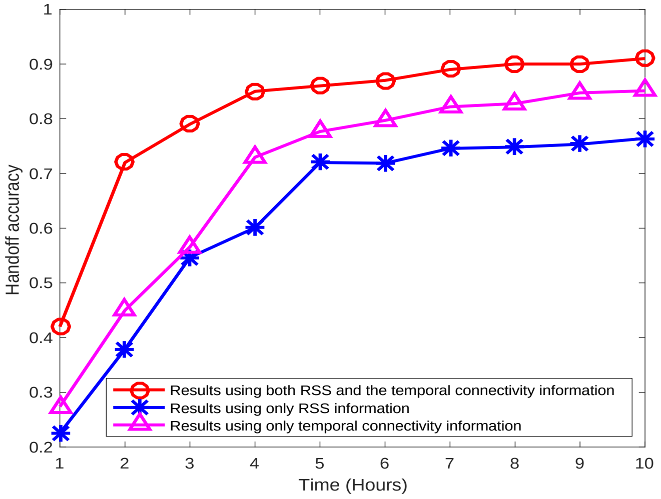

Figure 46 The handoff accuracy of the Test 1 using three types of information. (a) when both RSSI and the temporal connnectivity information are utilized (b) when only the RSSI information is utilized and (c) when only the temporal connnectivity information is utilized.

Figure 48 shows the handoff accuracy of the experiment we obtained in another shopping center where a total of 142 APs are deployed. This test is carried out during the weekend from 11:00am to 8:00pm. Compared to Figure 45, we can see that the handoff accuracy of the system deployed in a denser environment of APs is slightly lower. This can be explained as the fact that in a denser environment, the system needs to determine an AP from more number of candidate APs for a MC.

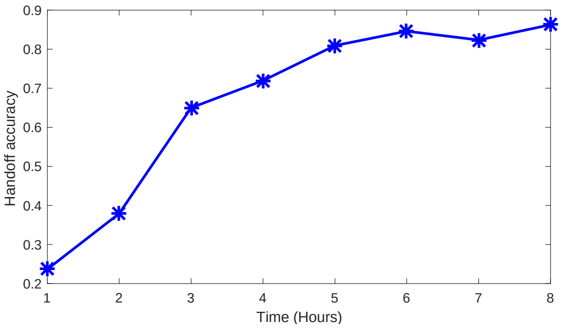

Figure 47 The handoff accuracy of the experiments carried out in the shopping mall with 142 APs

5.4 CONCLUSIONS

Fast handoff becomes a more severe issue to deal with in dense 802.11 WMNs because it is essential to fulfill the requirements of real-time applications and dense APs and MCs means more frequent handoff occurs.

In this paper, we propose an enhancement of network-assisted fast handoff scheme T-NRS based on previous radio signature technique. With spatial information based on measurement of RSSIs and the temporal information of historical records of associated APs, the handoff is decided based on probability of the next AP. The dedicated training phase of previous NRS scheme is eliminated and the learning procedure is interwoven with network operation. We have done experiments to evaluate the performance and it proves that handoff accuracy is acceptable while the original cost is avoided. Since the frequency of handoff would increase greatly in the coming era of dense networks, the combination of fast handoff approach based on spatial information and temporal information is expected to be explored deeper in our future work to forward the application in this area.

Chapter 6 Opportunistic Approach with Virtual Radio for Fast Handoff in AP-Dense Environment

6.1 Introduction

6.1.1 Motivations

With the prosperity of so-called “Business Sharing Wi-Fi”, IEEE 802.11-based Wireless Mesh Networks (WMNs) gains a new momentum to widespread exponentially in chain stores, shopping malls in China in past three years. Correspondently, when smartphone users are accessing the WMN, time-sensitive mobile APPs such as video chat, WeChat, and game King of Glory, face more frequent handoffs with denser APs and MCs in the network. Thus it imposes more severe challenge upon fast handoff mechanisms in such scenario, that is the time delay caused by interruptions of communication between Access Points (APs) and Mobile Clients (MCs) [66].

The legacy 802.11 handoff operation is comprised of three procedures: namely scanning, authentication and re-association. The scanning is that a MC collects information about neighboring APs and chooses the most appropriate AP for itself to handoff. The authentication procedure is to authenticate and authorize the access to the AP for a MC. And the association procedure is to re-establish the association with AP. The scanning is the most time-consuming, but the enhancements of fast handoff are not standardized and left for manufacturers and researchers to develop.

The essential work of fast handoff is: i) to probe-before-break, i.e. proactively probe information and decide which AP to associate before handoff occurs; and ii) to reduce the time of probing that becomes the key techniques which include terminal-based approaches and network-assisted approaches.

Terminal-based approaches of fast handoff have been thoroughly explored at earlier stage since it is convenient for academic community to tackle with terminals instead of network infrastructure. It basically controls how a MC switches channels, passively listens for beacons sent by APs ， or actively sends probe requests and receives probe responses from APs, etc. However, it is constrained by limited capacity of terminals and cannot solve the problem of compatibility of legacy terminals.

Thereafter network-assisted approaches gradually prevail in 802.11-based WMNs inspired by dedicated pilot channel and specialized function entities for handover on infrastructure-side in mobile cellular systems. It orchestrates multiple APs to work collaboratively to probe and produce information of handoff. APs listen for beacons from a MC rather than the opposite of the original scanning and specialized centralized mechanisms are implemented on APs both are because the network-side is powerful in terms of capacity of computing and memory to support sophisticated solution.

In the meantime, network-assisted approaches are welcome by Wi-Fi operators because it solves the practical problem of compatibility of existing legacy terminals. This is how we have built a testbed of over 200 APs in two shopping malls on the basis of a commercial platform comprised of thousands of APs in cooperation with a Business Wi-Fi operator in

China [67]. It is convenient and feasible for us to fiddle with AP, Access Controller (AC) and other network entities to study the problem space of network-assisted approaches.

In our previous work of network-assisted Radio Signature (NRS) probing, we have employed multiple APs to collaboratively monitor MCs and produce a vector of Received Signal Strength Indications (RSSIs) as radio signature of a MC. The most appropriate AP to associate is pre-defined in a training stage prior to the operational stage of the network, so that in the operational stage, handoff is determined by the measurement of radio signatures It is an arbitrary handoff scheme which completely eliminates the on-the-fly probing and the handoff delay, whereas the complexity of a training stage and the non-scalability for fixed installation of APs is introduced [71].

To avoid the abovementioned problem, in this paper we further propose a new network-assisted fast handoff scheme based on collaborative probing as well. It is called as opportunistic scanning (OppoScan) featured by a method of so-called virtual radio. In short, an AP serving a MC instructs its neighboring APs to assist in probing information for handoff by assuming the radios of other APs as its virtual radios. OppoScan utilizes neighboring AP's radios to probe other MCs for its own purpose instead of equipped itself with more physical radios.

Traditionally, it is expected that the AP serving a MC on one channel controls and exchanges information with all neighboring APs on all working channels to determine the most preferable AP via inter-AP communication of wired or wireless mesh links. However, the mechanism works with light-loaded APs and sparse MCs but would interfere with normal traffic when it comes to large-scale 802.11-based WMN with heavy-loaded and dense APs and MCs.

To address this issue, we improve the approach by selectively utilizing some of the neighboring APs in an opportunistic manner based on the geographic proximity of dense APs and MCs.

6.1.2 Contributions

We observe that the scale and density of 802.11-based WMNs has been increasing exponentially. In an urban area a MC received radio signals from more than 10 APs most of the time as in [26, 31, 34] but nowadays there are over 30 SSIDs are accessible for a MC and an AP serves over 50-70 smartphones in rush hours as in our controlled scenario of shopping malls.

Based on extensive field studies, we discover that when the density of neighboring APs involved in collaborative probing is greater than 50%, the OppoScan scheme can guarantee the accuracy of handoff to be over 90% without assistance of virtual radio of APs on different channels. Hence we simplify the mechanism by filtering some of neighboring APs when they are in the proximity with serving AP. Furthermore, when the density of MCs is greater than 10/m2, instead of instructing the virtual radio of neighboring APs to switch to other channels and probe, a MC can simply re-use the probing knowledge of another neighboring MC in the proximity in an opportunistic manner without switching to different channel at all. Basically this method reduces the overhead to probe if unnecessary.

The contributions of this chapter are summarized as follows:

- We propose a network-assisted fast handoff scheme OppoScan which collaboratively probes for APs by utilizing the virtual radios of neighboring APs in addition to its own radio.
- To the best of our knowledge, we are the first to analyze and utilize the correlations between the density of network nodes i.e. APs and MCs and the accuracy of handoff.
- We improve our approach of OppoScan to achieve optimized performance in an opportunistic manner by taking advantage of high density of neighboring APs and neighboring MCs which is a novel technical advance in the area of fast handoff
- We evaluate the solution with simulations and experiments in a real environment testbed use a proprietary version of cloud AC system WiSense running on top of openwrt. The results prove that the performance of the mechanism satisfied the requirements of current popular mobile applications

This remainder of this Section is organized as follows: in Section 6.2, we discuss the previous works with an overview of terminal-based handoff schemes and particularly the newly updates on network-assisted approaches. In Section 6.3, we elaborate our novel network-assisted mechanism OppoScan, a proactive, collaborative handoff scheme with a novel technique of opportunistic probing by utilizing virtual radios of neighboring APs. In Section 6.4, we evaluate our solution based on empirical and analytic models based on our discoveries on the correlations between density of APs and MCs and handoff accuracy in 802.11-based WMNs. In Section 6.5, we point out the limitation and our future work direction.

6.2 Background and Related works

6.2.1 Legacy Handoff Initiation and Parameter Optimization

We refer the primitive handoff procedures in 802.11 standards as legacy handoff hereafter. We have briefly addressed on the handoff issue which is originated from disruptions of the communication between a MC and its serving AP [66]. Due to the improvident nature of 802.11 standards on fast handoff schemes, the issue has been left open for hardware manufacturers and application developers to realize their proprietary solutions. Thus it has drawn amount of attractions from academic communities.

Normally the legacy handoff is initiated when the current connection between a MC and its serving AP is degraded too much to sustain. Nevertheless an improved handoff initiation is to measure the difference between the RSSI of its serving AP and the RSSI of its preferable AP. It works as a trigger to determine whether a handoff is initiated after it increases over a threshold. Mhatre et al. have explored different handoff initiation algorithms and triggers including beacon detection, threshold-based, hysteresis-based as well as trend-based algorithms [67]. Except signal strength, the quality of connection between MC and AP can be measured and evaluate in multiple dimensions including SNR, SIR, throughput, delay, jitter, AP load, and even QoS of mobile applications on upper layers. So far the initiation of handoff still remains a problem space to study while our current work mainly utilizes signal strength RSSI for simplicity.

Parameters related to legacy handoff have been studied to optimize for fast handoff. The normal legacy handoff consists of scanning procedure, authentication procedure and re-association procedure. Regarding the major source of delay, i.e. the probing procedure, MC searches for potential APs to associate with. It sends a Probe Request frame on every channel and listens on that channel for a MinChannelTime interval. If there is no answer, the channel is declared empty. On the other hand, if the MC receives at least one Probe Response from an AP, it waits until a MaxChannelTime interval expires to collect more information. Therefore early work on fast handoff adjusts the parameters in order to stay less time on each channel and probe fewer channels.

Mishra et al. made an empirical study on 802.11 MAC layer handoff process and proposed optimization in a heuristics manner which is to minimize the number of scanning channels, the response time to stay on channel [25]. They further introduced Neighbor Graphs of APs to reduce the total number of channels to probe and Non-overlap Graph to reduce the waiting time on each channel as well as caching techniques to solve the problem of packet loss [26]. H. Velayos et al have investigated techniques to start probing on the most possible channels, probe channels selectively, and derived values 1ms and 10.24ms for MinChannelTime and MaxChannelTime [34]. Y. Liao et al. propose a selective scanning which further reduces the channels to scan by utilizing adaptively method [68].

6.2.2 Terminal-based Approaches

Proactive probing schemes have been comprehensively studied. In the original process of legacy handoff, the scanning procedure is initiated after disconnection with the serving AP, that is referred as probe-after-break. Thereby in most commonly-adopted fast handoff schemes, the probing procedure is expected to be carried out prior to real handoff occurs, which is referred as probe-before-break. It requires the probing procedure to be conducted as a constant operation on the background.

Terminal-based approaches of proactive probing have been explored by many researchers. I. Ramani et al. designs SyncScan to synchronize MCs and APs first and then instructs MCs to probe the beacon periodically broadcasted by APs on each channel by switching to the channel at the exact time slot when a beacon is broadcast [31]. Essentially, it transfers one-time overhead of probing for beacons into evenly-distributed overhead by time-division with fixed time slots. Similarly, Haitao Wu et al. proposes ProactiveScan which instructs MCs to collect all information of neighboring APs in a one-time manner before switching back to current working channel [69], thus the overhead is distributed irregularly over time series.

It is MultiScan which is the first one to creatively propose a new way of fast handoff by equipping multiple radio interfaces on a MC to exploit hardware for proactive probing. One radio is used to search for alternate APs while the other radio is associated with AP for interleaving data communication [38].

SyncScan and ProactiveScan impose overhead of timing, synchronization and caching upon MCs, while MultiScan pays the price of redundant hardware on MCs. In [74] Singh et al. evaluate such impacts upon MC and AP.

The terminal-based approaches are convenient to implement for academic community since they require less modifications on network side. However, when it comes to 802.11-based WMNs in real world, the terminal-based approaches might introduce too much workload on terminals and more than that, they have the practical issue of incompatibility with existing legacy terminals.

6.2.3 Network-assisted Approaches

Network-assisted approaches are preferable in real-world large-scale networks for two justifications: i) One is that the network-side is more powerful than the terminal-side in terms of computation and communication capacity. By deploying intelligence on network-side, the overhead of probing, processing and caching is shifted from MCs to APs. Dedicated resources can be contributed to support fast handoff such as pilot channel and handoff management entities. ii) The other is that the network-assisted approaches enable compatibility with legacy terminals which have prevailed in markets.

Singh et al. [73] propose to set up a signaling channel to disseminate information of APs in the vicinity similar to the functionality of pilot channel in cellular networks. In addition to that they suggest eliminating the probing process when necessary. M. E. Berezin et al. [71] propose to use virtual APs to monitor MC on a dedicated channel and design a mechanism to enable a MC to permanently regard itself connected with the same logical AP so that no need to initiate handoff to incur delay. S. Jin et al. are inspired by MultiScan and propose a novel scheme by installing multiple radios on an AP [72], which basically contribute one radio to probing constantly so that the other radio can be spared from handoff operations.

In our previous work [50], we have proposed a network-assisted fast handoff scheme to utilize APs to monitor a MC and manage its handoff by building a correlation of radio signature of the MC and its serving AP. We eliminate the probing process totally at a cost of introducing a training stage prior to operational stage of the network.

We conclude that the network-assisted approaches have significantly reduced the overhead on terminal-side, enabled compatibility with terminals available on markets. Nevertheless, the cost of the approaches and the impact on the communication between MCs and its serving APs both are expected to be minimized, this is the major concern in our current research.

6.3 Opportunistic Probing with Virtual Radio

Before we introduce the proposed OppoScan, we first briefly reviewed the procedures of the traditional network-assisted fast handoff which are shown as follows：

1. MC_1 is associated with AP_1 on the channel CH_1. MC_1 starts moving and AP_1 detects the RSSI of MC_1 is less than a predefined threshold $T_{handoff}$.
2. AP_1 exchanges information with the single-hop neighboring APs on the same or different channels via inter-AP communication of wired or wireless mesh links.

3. Each of the single-hop neighboring APs starts to measure the RSSI between itself and the MC.
    a) If a neighboring AP is on the same channel of MC_1, it listens to beacons sent from MC_1 for a period of time and can obtain the RSSI values from the received packets. If no beacons received, then it will send a probe request message to MC_1 and extract the RSSI information from the probe response message sent from MC_1.
    b) If a neighboring AP is not on the same channel of MC_1, then it will temporarily disconnect all the MCs it is serving, switch to the same channel of MC_1, and obtain the RSSI information by listening or exchanging the messages with MC_1.
4. The AP with the largest RSSI will be chosen as the next serving AP of MC_1.

From the procedures above, we can see that the major problem lies in the condition when an AP wishes to measure the RSSI between itself and the MC in a different channel. An example is shown in Figure 48. In this example, AP_1 is the serving AP of MC_1, and they work in a different channel of AP_2. When AP_2 wishes to measure the RSSI between itself and MC_1, AP_2 needs to switch to the channel of MC_1, thus disrupt all the communications between AP_2 and the MCs it is serving..

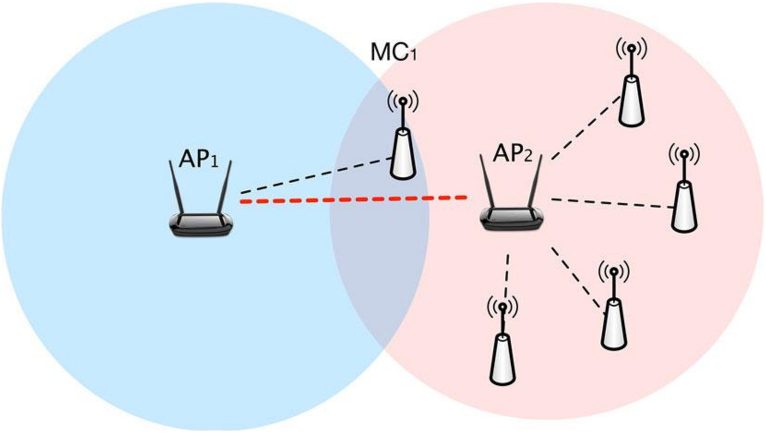

Figure 48 AP2 needs to switch to the channel of MC1 thus disconnecting all MCs currently serving

Therefore, we have the following question, for a MC and its potential next serving AP but is currently working in a different channel, is it possible to estimate the RSSI without requiring the AP to switch to the channel of the MC? To be more specific, we wish to find out, for a MC, the AP working on a different channel but have the good RSSI value.

Assume AP_1 is the serving AP of MC_1 working on a channel CH_1 and the RSSI value between them is $RSSI_1$. AP_2 is a single-hop neighboring AP working on a different channel CH_2. AP_2 is the serving AP of MC_2 and the RSSI value between them is $RSSI_2$. We have the following observations.

1. If AP_2 is in geographic proximity to AP_1, then after AP_2 switching to CH_1, the measured RSSI value between MC_1 and AP_2 will be similar to $RSSI_1$. This means that if AP_1 detects the RSSI of MC_1 is less than a predefined threshold and starts to looking for the next serving AP in its single-hop neighboring AP list, AP_2 should not be considered. This observation is shown in Figure 49.

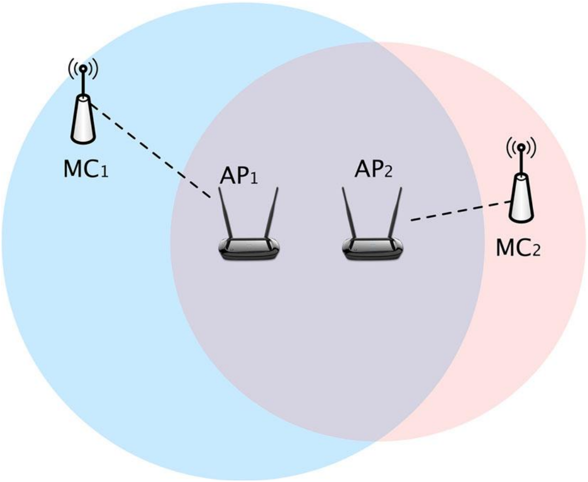

Figure 49 Observation 1: AP2 to be filtered as next serving AP when geographically near AP1

2. If MC_1 is in geographic proximity to MC_2, then after AP_2 switching to CH_2, the measured RSSI value between MC_1 and AP_2 will be similar to $RSSI_2$. This means that

in some conditions, the RSSI values between a MC and an AP can be re-used by a nearby MC as the estimate of the RSSI if it is connected to the AP. This observation is shown in Figure 50.

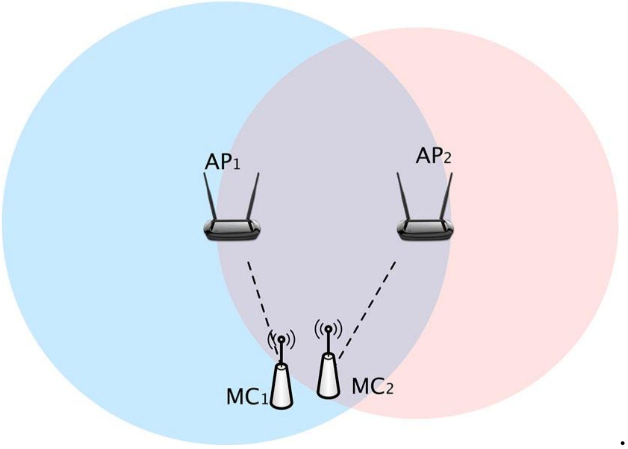

Figure 50 Observation 2: RSSI to be re-used when MC1 and MC2 are close to each other

We utilize Figure 51 as an example to show how to apply the two observations above in real conditions. In Fig.5, AP_1 is the serving AP of MC_1 and has four single-hop AP neighbors: AP_2 ~ AP_5, and we assume that all of these 4 neighboring APs are not working in the same channel of AP_1. When AP_1 finds that the RSSI of MC_1 is less than a predefined handoff threshold and starts to look for the next serving AP, it first eliminates AP_4 and AP_5 from the list as these two APs are very close to AP_1. In addition, MC_1 finds out that it is close to MC_2 which connects to AP_2, then $RSSI_2$ will be utilized as the RSSI value between MC_1 and AP_2. Similarly, as MC_3 is near to MC_1, then $RSSI_3$ will be utilized as the estimate of RSSI between MC_1 and AP_3. If, for example, both $RSSI_2$ and $RSSI_3$ are larger than the handoff threshold and $RSSI_2$ is greater than $RSSI_3$, then MC_1 will choose AP_2 as the next serving AP. In the scenario shown in Fig. 5, although there are 4 APs that are working on a different channel from AP_1, none of these APs need to switch to the channel of AP_1 to measure the RSSI with MC_1, thus realizing fast handoff without requiring disruption of communications between any APs and their serving MCs.

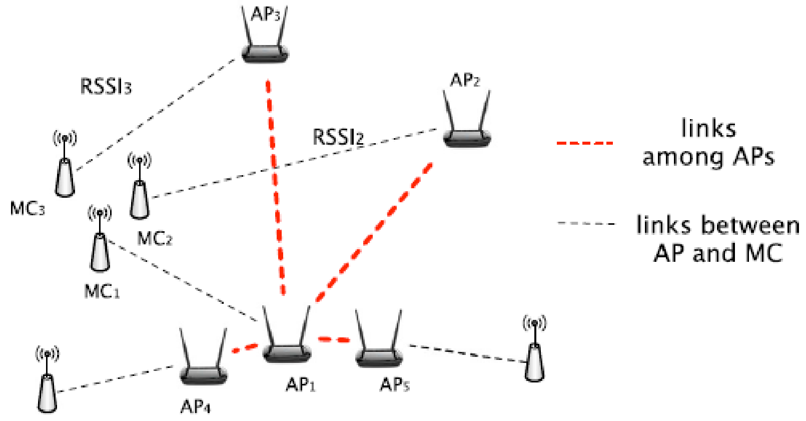

Figure 51 Example of using the two observations above.

It can be seen from the example above that, in order to estimate the RSSI value between a MC to an AP working on a different channel without requiring the switching of AP, we utilize the RSSI values measured by a neighboring MC to the AP, if such a neighboring MC does exist. From this perspective, we virtually utilize the radios of neighboring APs on different channels in an opportunistic way. This is why we call this method the opportunistic probing with virtual radios.

Presumption:

- APi is the serving AP for MCi and they work on the channel CHi.
- APi detects the RSSI of MCi is less than a predefined threshold $T_{handoff}$.

Input:

- The list of all the single-hop neighboring APs of APi : $\{AP_{\text{single-hop}}\}$
- The list of all the APs that are geographically near to APi with the distances less than D_{AP} meters: $\{AP_{\text{near}}\}$
- The list of MCs that are geographically near to MCi with the distances less than D_{MC} meters: $\{MC_{\text{near}}\}$

Output: A single AP which will be utilized as the serving AP for MCi

1. APi will delete all the APs in $\{AP_{\text{near}}\}$ from $\{AP_{\text{single-hop}}\}$
2. $\{AP_{\text{single-hop}}\}$ is divied into two sets: those APs in the same channel of APi and different channel of APi. The former is denoted as $\{AP_{\text{single-hop}}\}_A$, and the latter is denoted as $\{AP_{\text{single-hop}}\}_B$.
3. APi sends a Probe Request message to all its single-hop neighboring APs via the mesh links between APs.
4. APi will find out the serving APs for each MC in the $\{MC_{\text{near}}\}$ and generate a list of APs: $\{AP(MC_{\text{near}})\}$.
5. APi will find APs which are in different channel of APi but have a MC near the MCi: $\{AP_{\text{virtual}}\} = \{AP(MC_{\text{near}})\} \cup \{AP_{\text{single-hop}}\}_B$
6. Using the RSSI values measured between each of the APs in the set $\{AP_{\text{virtual}}\}$ and the corresponding MC as the ones measured between MCi and the APs in the $\{AP_{\text{virtual}}\}$.
7. For those APs in the same channel of APi :$\{AP_{\text{single-hop}}\}_A$, APi will receive the probe response to collects the RSSI measurement.
8. For those APs in the $\{AP_{\text{single-hop}}\}_B$ and are not in the set of $\{AP_{\text{virtual}}\}$, they will switch to the channel of MCi to measure the RSSI.
9. Using the RSSI values in $\{AP_{\text{virtual}}\}$, the $\{AP_{\text{single-hop}}\}_A$, and the ones measured in Step 8, the AP with the maximum RSSI will be chosen as the next serving AP (denoted as AP*next*.)

Figure 52 Algorithm of the proposed scheme OppoScan

Note that if the potential single-hop APs of AP_i are working in the same channel with the current MC, then AP_i will send a probe request to ask these APs to measure the RSSI between themselves and the MC. In addition, for the remaining APs that are not in the same channel of AP_i and are not in geographic proximity to AP_i, and at the same time, do not have any MCs near to the current MC_i, these APs will switch to the channel of MC_i to measure the RSSI. Algorithm 1 shows the procedures of the proposed method.

To determine the single-hop neighboring of an AP is relatively easy, as the deployment of APs is generally handily available. To find out the neighboring MCs for a certain MC, we can utilize the methods proposed in [75][76][77], which are generally encounter-based methods that can detect the nearby mobile phone users using different technologies like Wi-Fi and Bluetooth.

6.4 EVALUATION

6.4.1 Testing Scenario

We deployed a total of 198 APs in two shopping malls in CBD of Shenzhen. The APs are all IEEE 802.11b/n 2.4GHz band. The hardware chipsets of APs include Atheros 9344 and MTK 9531. The topology of the whole 802.11-based WMN is mesh networks connected by wired and wireless links distributed on different floors of the two shopping malls.

We implement the proposed scheme in the firmware of the APs. We configure the network so that the experiments are conducted in a controlled environment. *MinChannelTime* is set to 1ms and *MaxChannelTime is set to* 10ms. The channel switch time is 300us. Three non-overlapping channels CH_1, CH_6, CH_{13} are used. *RTS/CTS* handshaking is used in the virtual AP environment with different data rates. The handoff delay time is defined as the period that a MC cannot exchange date packets with any AP during the handoff process.

6.4.2 Performance Criteria

We utilize the *handoff accuracy* as the measure to evaluate the performance. We define that if using the proposed OppoScan scheme, an AP has found its single-hop AP with the greatest RSSI value, then this handoff is regarded as correct. The handoff accuracy is then defined as the ratio of the number of correct handoffs to the total number of handoffs occurred during the experiment.

Besides the handoff accuracy, we also evaluate the benefit of the proposed OppoScan using the *switching ratio*, which is defined as the percentage of the number of APs that have been switched to nother channels. Obviously, the smaller the value, the greater the benefit of OppoScan is.

6.4.3 Performance Evaluations

We will show the performance of the OppoScan in a test in a shopping mall in which 56 APs are deployed. This test last for 30 minutes and involved 25 mobile phone users who randomly walked around in the shopping mall. In this test, the threshold to determine whether to APs are geographically together is set to 15 meters.

Figure 53 shows the handoff accuracy of the OppoScan. We can see that on average, the handoff accuracy can reach about 65%~90% in most cases, which indicates that OppoScan can find out the best next-hop AP with the probability of 65%~90%.

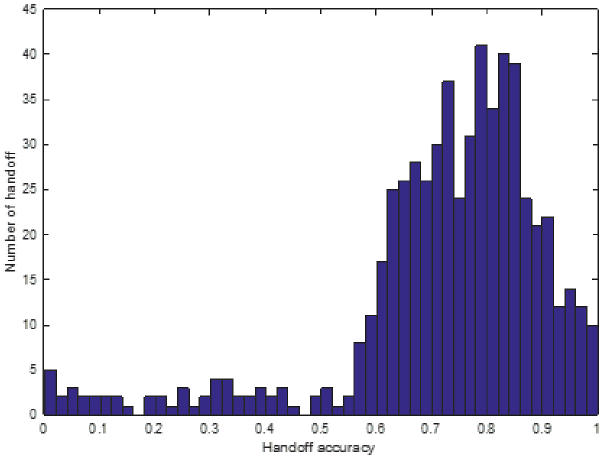

Figure 53 Handoff accuracy of the experiment carried out in the shopping mall with 56 APs

Figure 54 shows the switching ratio of the system during the test. We can see that the average switching ratio is about 0.3 which indicates that the OppoScan can save about 70% switches of APs on average.

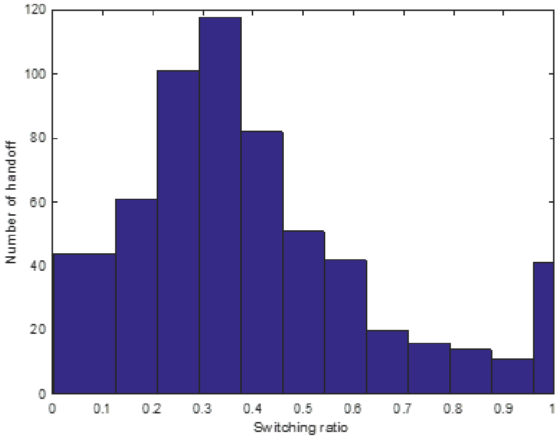

Figure 54 Switching ratio of the experiment carried out in the shopping mall with 56 APs.

Figure 55 shows the handoff accuracy of the experiment we obtained in another shopping center where a total of 142 APs are deployed. Compared to Figure 53, we can see that the handoff accuracy of the system deployed in a denser environment of APs is slightly lower.

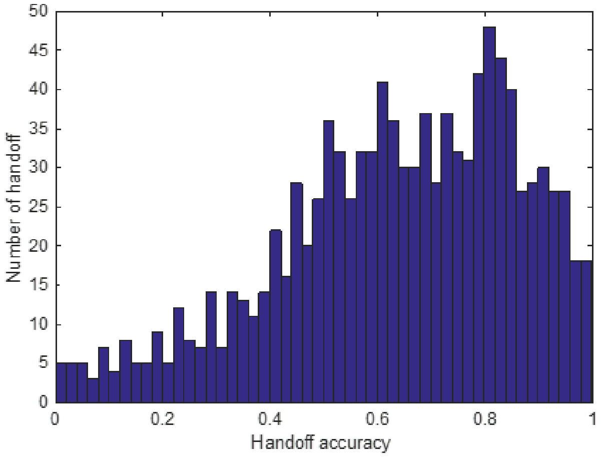

Figure 55 Handoff accuracy of the experiment carried out in the shopping mall with 142 APs.

Figure 56 shows the switching ratio of the system during the test above. Compared to Figure 54, we can see that the switching ratio is much lower. This indicates that the OppoScan can save more switches of APs in a denser environment.

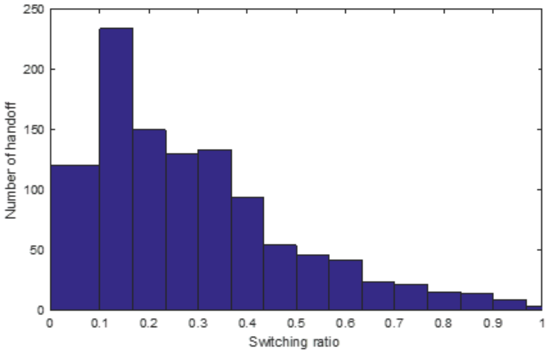

Figure 56 Switching ratio of the experiment in the shopping mall with 142 APs

6.4.4 The effect of important parameters

In this section, we evaluate the effect of important parameters for the OppoScan.

The first parameter we wish to evaluate is the threshold that determines whether two APs are geographically near to each other. We have carried out a number of tests in the shopping mall with 147 APs, with each test corresponding to a different threshold ranging from 10m, 15m, 18m, 21m, 24m, 27m and 30m. Figure 57 shows how the average accuracy and the switching ratio change with the change of this threshold. It can be seen that generally speaking, with the increase of the threshold, both the handoff accuracy and the switching ratio will be decreased. This can be easily explained as the fact that for a certain APi, a larger threshold will delete more neighboring APs, thus decrease the switching ratio. However, due to the same reason, the price to pay is that the handoff accuracy will be decreased.

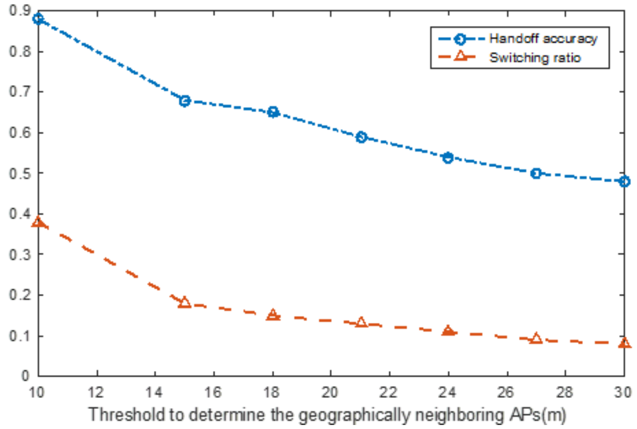

Figure 57 Effect of the threshold of APs proximity on handoff accuracy and switching ratio

The second parameter we wish to evaluate is the density of MCs. Similarly, we carried out 4 tests in the shopping mall with 147 APs, with each test corresponding to different density of MCs. To control the density, we ask the different number of mobile phone users walking randomly in the same floor of the shopping mall at approximately the same speed. The numbers of mobile phone users in these tests are 5, 10, 15, and 25. In these tests, the threshold to determine whether two APs are geographically near to each other is set to be 15m. Figure 58 shows how the average accuracy and the switching ratio change with the change of the number of mobile phones. It can be seen that generally speaking, with the increase of the mobile phone density, the handoff accuracy slightly decreases but generally remains in the same level. On the other hand, with the increase of the density, the switching ratio is significantly decreased. This can be explained as the fact that with more number of mobile phones, a certain AP$_i$, can have higher probability to find a nearby MC and utilize its RSSI values.

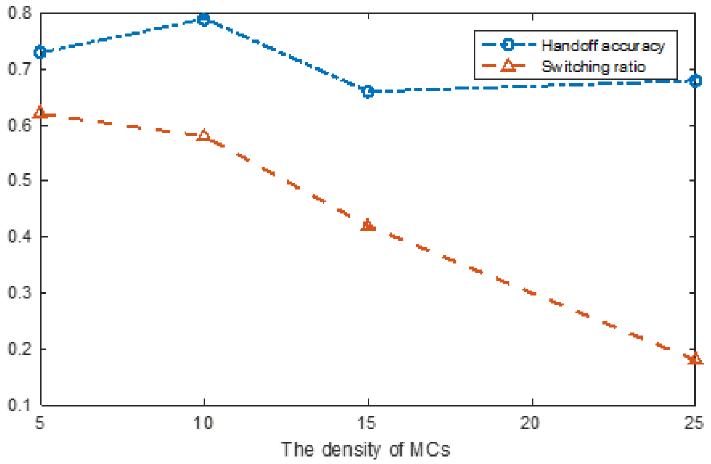

Figure 58   Effect of density of MCs on handoff accuracy and switching ratio

6.5 Conclusions

Fast handoff is a very important issue to deal with in 802.11-based WMN because it is essential to fulfill the real-time applications with stringent delay requirements. Currently the rapid business deployments of 802.11-based WMN require network-assisted approaches rather than terminal-based approaches because of practical purpose of being compatible with legacy smartphones on the market.

In this section, we propose a novel network-assisted fast handoff mechanism. With opportunistic and collaborative probing of partial APs instead of orchestrate all neighboring APs to probe, the connection interruption occurred by APs switching between channels is significantly alleviated, while at the same time, the accuracy of handoff is maintained. We have done extensive experiments for evaluation and we will be working to forward network-assisted approaches in our future research in large-scale dense 802.11-based networks.

.

Chapter 7 Conclusions and Future Research

7.1 Conclusions

In this thesis we present the detailed mechanisms, simulations, implementation and experimental results of Secure Fast Handoff solutions in 802.11 Wireless Mesh Networks. We have conducted systematic investigations and studies on key issues. With respect to each issue, we identify problems which are important but not yet well addressed and propose corresponding solutions. We conclude our works as follows.

To address secure fast handoff in 802.11-based WMN based on current off-the-shelf hardware, we have proposed Network-assisted Radio Signature scheme as a novel Layer 2 fast handoff mechanism to realize proactive probing by making the most of the static topology of an actual 802.11-based WMN, with a trade-off of training phase prior to operational phase. We also propose a proactive dual re-authentication mechanism which achieves fast handoff without great degradation in security criteria. It improves the security of optimistic access by providing one time ticket while effectively reduces the authentication delay. We have quantitatively analyzed both re-authentication delay and security degradation and also proposed numerical model for optimizing parameters for system performance. To our best knowledge, we are the first to achieve such an optimal performance of Secure Fast Handoff at no cost of extra hardware or overhead, comparing with other pilot projects.

Furthermore, to address the handoff problem in an AP-dense WMN environment and push forward the above technique Network-assisted Radio Signature, we further propose Temporal-NRS (T-NRS) scheme. It leverages historical knowledge of APs associated in time series to assist in handoff decision in addition to NRS technique based on spatial knowledge. The enhanced scheme improves the performance whilst it greatly eliminates the inflexibility of the original approach by no longer requires a training phase.

At last, to continue addressing the handoff problem in an AP-dense WMN environment, a collaborative scheme OppoScan (Opportunistic Scanning) supported by virtual radio is proposed. OppoScan opportunistically leverages nearby MCs and APs to produce the required information of neighboring AP for handoff, thus significantly decrease the number of switching channel of APs. Our evaluation based on experiments indicates that OppoScan can efficiently achieve low delay while maintaining handoff in more practical scenarios for 802.11-based WMN.

7.2 Future Research

The work presented in this thesis is still on its early stage for mass application and still requires much more improvements. We conclude the thesis by providing some suggestions for future research. Specifically, we believe the below aspects are worthy of further exploring.

First, given a highly-dense AP and MC scenario in the future, we believe distributed approaches will further improve both AP probing and re-authentication to achieve secured fast handoff. Since security is very important in real-world deployment, we will study how to reduce re-authentication delay after introducing more secure mechanisms while maintaining the overhead of security management among Aps and MCs to be at a low level. The existing work uses proactive re-authentication but relies on communication with AS. In our next developing scheme, the intermediate credentials will be managed and re-used cooperatively by Aps and MCs in a highly distributed manner instead of being re-generated by AS. With this work we aim to further reduce re-authentication delay re-use of PMK/PTK and decrease dependency on Authentication Server.

Second, to evaluate the tradeoff between secure fast handoff and the cost of the mechanisms, we will carry out a comprehensive study on the overhead introduced. We will compare time cost and message overhead for each scheme by experiments. We will also identify the application scenarios for each scheme in terms of network size and user mobility pattern.

Thirdly, we see a future of booming of 802.11-based WMN co-existing with cellular systems 5G to extend the coverage of pervasive Internet access. This will bring new opportunities and challenges to provide seamless communication and applications for mobile users to roam between 802.11-based WMN and cellular systems. The integration of handover mechanisms of cellular systems will be a new direction of secure fast handoff in 802.11 networks.

At last but not the end, new application scenarios always bring new challenges upon secure fast handoff mechanisms such as IoT and blockchain. We will further carry out research on the massive connections introduced by IoT and constant updates introduced by blockchain applications. We will aim to improve handoff performance by taking advantage of the highly-distributed mobile nodes for computing and storage.

References

[1] B. Aboba, D. Harkins, A. Alimian, M. Lefkowitz, Thinking About the Site Report. IEEE 802.11r submission, 2004.

[2] W. Arbaugh, Improving the latency of the Probe Phase during 802.11 Handoff. IEEE 802.11k submission, 2004.

[3] Atheros, Power Consumption and Energy Efficiency Comparisons of WLAN Products,Online: http://www.atheros.com/pt/whitepapers/atheros_power_whitepaper.pdf, 2003.

[4] Atheros, 802.11ag The Clear Choice, Online: http://www.atheros.com/pt/whitepapers/atheros_802.11ag_whitepaper.pdf, 2005.

[5] M.S. Bargh, R. J. Hulsebosch, H. E. Ertink, A. Prasad, H. Wang, P. Schoo. Fast Authentication Methods for Handovers between IEEE 802.11 Wireless LANs. WMASH'04, Philadelphia, Pennsylvania, USA, 2004.

[6] CCITT, ITU (International Telecommunication Union), General Characteristics of International Telephone Connections and International Telephone Circuits, 1988.

[7] Cisco 2004. IEEE 802.11b DSSS Channel Allocations. Online: http://www.cisco.com/application/pdf/en/us/guest/products/ps469/c1650/ccmigration_09186a008008883b.pdf

[8] Cisco, 2005, “Cisco Wireless IP Phone 7920”, Online: http://www.cisco.com/en/US/products/hw/phones/ps379/products_data_sheet09186a00801739bb.html

[9] Cisco, 2005, “Capacity Coverage and Deployment Considerations for IEEE 802.11G” Online:http://cco.cisco.com/application/pdf/en/us/guest/products/ps430/c1244/ccmigration_09186a00801d61a3.pdf

[10] The Economist, 2005, “Hearing Voices”, Online: http://www.economist.com/printedition/displayStory.cfm?Story_id=3307396

[11] The Economist, 2005, “The war of the wires”, Online: http://www.economist.com/business/PrinterFriendly.cfm?Story_ID=4232442

[12] Electronics, N. 2005. "802.11n or UWB?" Online: http://neasia.nikkeibp.com/neasia/002271

[13] Engadget, 2005, “LG’s new CL400 UMA Wi-Fi Phone”, Online: http://www.engadget.com/entry/1234000147059793/

[14]Geier, J. 2004. "802.11a vs. 11g in Homes" Online: http://www.Wi-Fiplanet.com/tutorials/article.php/3337861

[15] IEEE, 1999. "IEEE Standards for Information Technology -- Telecommunications and Information Exchange between Systems -- Local and Metropolitan Area

Network -- Specific Requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications."

[16] IEEE, 1999, "IEEE Standard for Information technology— Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications— Amendment 1: High-speed Physical Layer in the 5 GHz band."

[17] IEEE, 1999, "Supplement to IEEE Standard for Information technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements— Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band."

[18] IEEE, 2003, "802.11F IEEE Trial-Use Recommended Practice for Multi-Vendor Access Point Interoperability Via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation."

[19] IEEE, 2003, "IEEE Standards for Information Technology -- Telecommunications and Information Exchange between Systems -- Local and Metropolitan Area Network -- Specific Requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 5: Spectrum and Transmit Power Management Extensions in the 5 GHz band in Europe"

[20] IEEE, 2004, "IEEE Standards for Information Technology -- Telecommunications and Information Exchange between Systems -- Local and Metropolitan Area Network -- Specific Requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 6: Medium Access Control (MAC) Security Enhancements."

[21] InfoSyncWorld. Preview Nokia N91 Online: http://www.infosyncworld.com/news/n/5944.html

[22] M. R. Jeong, F. Watanabe, T. Kawahara. Fast Active Scan for Measurement and Handoff. IEEE submission <P802.11-03/623r0>, 2003.

[23] S. Matta et al, Proposed Text for Neighbor Report Enhancements. IEEE 802.11k proposal <11-04-0735-03-000k-site-report-enhancements.doc>, 2004.

[24] P. Molinero-Fernandez, N. McKeown, H. Zhang. Is IP going to take over the world (of communications)? ACM SIGCOMM Computer Communication Review, vol. 33, no. 1, pp. 113-118, 2003.

[25] A. Mishra., M. Shin, W. Arbaugh. An empirical analysis of the IEEE 802.11 MAC layer handoff process. ACM SIGCOMM Computer Communication Review, vol.33, no.2, pp. 93-102, 2003.

[26] A. Mishra., M. Shin, W. Arbaugh. Context Caching using Neighbor Graphs for Fast Handoffs in a Wireless Network. Joint Conference of the IEEE Computer and Communications Societies, 1:361, 2004.

[27] A. Mishra, M. Shin, W. Arbaugh. Proactive Key Distribution using Neighbor Graphs. IEEE Wireless Communications, vol. 11, no. 1, pp. 26-36, Feb 2004.

[28] Nortel, 2001, Voice over Packet: An Assessment of Voice Performance on Packet Networks, Online: http://www.nortel.com/products/library/collateral/74007.25-09-01.pdf

[29] R. Pries, K. Heck. Performance Comparison of Handover Mechanisms in Wireless LAN Networks, University of Wurzburg, Germany, Tech. Report. 339, 2004.

[30] R. Pries, K. Heck. Simulative Study of the WLAN Handover Performance. OPNETWORK 2005, Washington D.C.

[31] I. Ramani., S. Savage, 2005. SyncScan: Practical Fast Handoff for 802.11 Infrastructure Networks. IEEE INFOCOM, vol. 1, pp. 675-684, 2005.

[32] J.-O. Vatn. An Experimental Study of IEEE 802.11b Handover Performance and It's Effect on Voice Traffic. Royal Institute of Technology, Stockholm, Sweden, Tech. Rep. TRITA-IMIT-TXLAB R 03:01 KTH, 2003.

[33] S. J. Vaughan-Nichols. 802.11a: Wait until Next Year! Online: http://www.Wi-Fiplanet.com/columns/article.php/1465111

[34] H. Velayos, G. Karlsson. Techniques to Reduce IEEE 802.11b MAC Layer Handover Time. RTH Royal Institute of Technology, Tech. Rep. TRITA-IMIT-LCN R- 03/02, 2003

[35] The World Paper, 2004, Wireless Internet is Growing Virally, Online: http://www.worldpaper.com/archive/2003/september_04/september2.html

[36] Gartner Research: http://www.gartner.com.

[37] M. Shin, A. Mishra and W. Arbaugh. Improving the latency of 802.11 hand-offs using neighbor graphs. Proc. of International Conference on Mobile Systems, pp. 70-83, June 2004.

[38] V. Brik, A. Mishra and S. Banerjee. Eliminating handoff latencies in 802.11 WLANs using Multiple Radios: Applications, Experience, and Evaluation. ACM Sigcomm Conference on Internet Measurement, p. 27, Oct. 2005.

[39] K. Kwon and C. Lee. A fast handoff algorithm using intelligent channel scan for IEEE 802.11 WLANs. Proc. of the ICACT 2004, pp. 46-50, Phoenix Park, Republic of Korea, Feb. 9-11, 2004.

[40] S. Pack, J. Choi, T. Kwon and Y. Choi, Fast-handoff support in IEEE 802.11 wireless networks, IEEE Communications Surveys & Tutorials, vol. 9, no. 1, pp. 2-12, Jan. 2007.

[41] S. Shin, A.G. Forte, A.S. Rawat and H. Schelzrinne. Reducing MAC layer handoff latency in IEEE 802.11 wireless LANs. Proc. of ACM MobiWac'04, pp. 19- 26, Philadelphia, PA, USA, Oct. 1, 2004.

[42] D. Johnson, C. Perkins and J. Arkko. Mobility Support in IPv6. IETF RFC3775, June 2004.

[43] R. Ramjee, K. Varadhan, L. Salgarelli, S.R. Thuel, Shie-Yuan Wang, T. La Porta. HAWAII: A domain-based approach for supporting mobility in wide-area wireless networks. IEEE/ACM Transactions on Network,vol:10, no. 3, pp.396-410, 2002.

[44] P. Xu et al. Optimized Integrated Registration Procedure of Mobile IP and SIP with AAA Operations. Proc. 20th International Conference on Advanced Information Networking and Applications, vol. 1 (AINA'06), pp. 926-931, 2006.

[45] A. Dutta et al. MPA Assisted Optimized Proactive Handoff Scheme. Proc. 2nd Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services 2005 (MobiQuitous 2005), pp. 155-165, 2005,.

[46] T. Kwon, M. Gerla, and S. Das. Mobility Management for VoIP: Mobile IP vs. SIP, IEEE Wireless Communication Magazine., vol. 9, no. 5, pp. 66-75, Oct. 2002.

[47] M. Georgiades et al. AAA Context Transfer for Seamless and Secure Multimedia Services over All-IP Infrastructures. Proc. of 5th European Wireless Conference, Spain, Feb. 2004.

[48] T. Braun and K. Hahnsang. Efficient Authentication and Authorization of Mobile Users Based on Peer-to-Peer Network Mechanisms. Proc. 38th Annual Hawaii International Conference on System Sciences (HICSS '05), Jan. 2005.

[49] T. Aura and M. Roe. Reducing Re-authentication Delay in Wireless Networks, Proc. of the First International Conference on Security and Privacy for Emerging Areas in Communication Networks (SecureComm 2005), pp. 139-148, Sept. 2005.

[50] G. Yao, J.N. Cao, Y. Yan, Y.S. Ji. Secured Fast Handoff in 802.11-based Wireless Mesh Networks for Pervasive Internet Access. IEICE Transactions on Information and Systems. vol.E93-D, no.3, pp. 411-420, Mar. 2010.

[51] Crypto++ 5.5 Benchmarks. [Online]: http://www.cryptopp.com/benchmarks.html

[52] IEEE 802.11, Standard Speciation for Wireless Local Area Networks, http://standards.ieee.org/wireless/

[53] IEEE 802.11X, Standard Specification for Wireless Local Area Networks

[54] IEEE 802.11i, Standard Specifications for Wireless Local Area Networks

[55] A. Shamir. How to Share a Secret, Communications of the ACM, vol. 22, no. 11, pp. 612-613, November 1979.

[56] A. Weimerskirch and D. Westho. Zero Common-Knowledge Authentication for Pervasive Networks. Proc. of Tenth Annual International Workshop on Selected Areas in Cryptography (SAC 2003), pp. 73-78, 2003.

[57] M. Girault. Self-certied Public Keys, Advances in Cryptology- EUROCRYPT '91, D.W. Davies (Ed.), LNCS 547, Springer-Verlag, pp. 490-497, 1991.

[58] L. Eschenauer and V.D. Gligor. A key-management scheme for distributed AP networks, Proc. of 9th ACM conference on Computer and Communications Security, pp. 41-47, 2002.

[59] D. Liu and P. Ning. Location-Based Pairwise Key Establishments for Static Sensor Networks. Proc. of the First ACM Workshop Security of Ad Hoc and Sensor Networks (SASN) '03, pp. 72-82, 2003.

[60] A. Weimerskirch and D. Westho. Identity Certified Authentication for Ad-hoc Networks, Proc. of the First ACM workshop on Security of ad hoc and AP networks (SASN), pp. 33-40, 2003.

[61] X. Zheng, C. Chen, C. T. Huang, M. Matthews, and N. Santhapuri. A Dual Authentication Protocol for IEEE 802.11 Wireless LANs. Proc. of IEEE Second International Symposium on Wireless Communication Systems, pp. 565-569, 2005.

[62] S. Zhu, S. Xu, S. Setia, and S. Jajodia. LHAP: A Lightweight Hop-by-Hop Authentication Protocol for Ad-hoc Networks. Proc. of IEEE 23rd International Conference on Distributed Computing Systems Workshops (ICDCSW ’03), pp. 749-755, 2003.

[63] C.S. Zhang, J.N. Cao, G. Yao. CoDA: Connectivity-Oriented Data Dissemination Algorithm for Vehicular Internet Access Networks. MSN 2015: 186-193.

[64] Y.M. Deng, G.J. Wang, J.N. Cao, X. Xiao. Practical Secure and Fast Handoff Framework for Pervasive Wi-Fi Access. IET Information Security, vol. 7, no. 1, pp. 22-29, 2013.

[65] K.H. Chi, Y.C. Shih, H.H. Liu, J.T. Wang. Fast Handoff in Secure IEEE 802.11s Mesh Networks. IEEE Transactions on Vehicular Technology, vol. 60, no. 1, pp. 219-232. Jan. 2011.

[66] https://en.wikipedia.org/wiki/IEEE_802.11

[67] V. Mhatre and K. Papagiannaki, Using Smart Triggers for Improved User Performance in 802.11 Wireless Networks, in Proceedings of the 4th International Conference on Mobile Systems, Applications and Services, Jan. 2006, pp. 246–259.

[68] Y. Liao and L. Gao. Practical Schemes for Smooth MAC Layer Handoff in 802.11 Wireless Networks. WoWMoM, 2006.

[69] Haitao Wu, K. Tan, Y. Zhang, and Q. Zhang, Proactive Scan: Fast Handoff with Smart Triggers for 802.11 Wireless LAN. INFOCOM, 2007.

[70] J. Teng, C. Xu, W. Jia, and D. Xuan. D-Scan: Enabling Fast and Smooth Handoffs in AP-dense 802.11 Wireless Networks. INFOCOM, 2009.

[71] M. E. Berezin, F. Rousseau, and A. Duda, Multichannel Virtual Access Points for Seamless Handoffs in IEEE 802.11 Wireless Networks, in Proc. IEEE 73rd Vehicular Technology Conference (VTC Spring), May 2011.

[72] S. Jin and S. Choi. A Seamless Handoff With Multiple Radios in IEEE 802.11 WLANs, in IEEE Transactions on Vehicular Technology, vol. 63, no. 3, pp. 1408-1418, March 2014.

[73] G. Singh, A Atwal, B. Sohi, A Signaling Technique for Disseminating Neighbouring AP Channel Information to Mobile Stations, in: ICDCN2006, Springer, Berlin, 2006.

[74] G. Singh, A. Atwal, B. Sohi. Effect of Background Scan on Performance of Neighboring Channels in 802.11-based networks, International Journal of Communication Networks and Distributed Systems vol.1 (IJCNmesh links) 2008.

[75] Jun J, Gu Y, Cheng L, et al. Social-Loc: Improving Indoor Localization with Social Sensing. SenSys 2013.

[76] Chen H, Chen Y L, Wu C H, et al. EcoLoc: Toward Universal Location Sensing by Encounter-Based Collaborative Indoor Localization. Proceedings of the Second International Conference on Internet-of-Things Design and Implementation. ACM, 2017: 215-220.

[77] Vanderhulst G, Mashhadi A, Dashti M, et al. Detecting Human Encounters from Wi-Fi Radio Signals. Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia. ACM, 2015: 97-108.

[78] Gang Yao, Jian-nong Cao, Xuefeng Liu, et al. OppoScan: Enabling Fast Handoff in Dense 802.11 WMNs via Opportunistic Probing with Virtual Radio. IEEE 14th Mobile Ad Hoc and Sensor Systems (MASS17). Oct. 2017.

[79] Gang Yao, Jian-nong Cao, Xuefeng Liu, et al. Fast Handoff based on Enhancement of Network-assisted Radio Signature in 802.11 Dense WMNs. Accepted by ACM International Conference on Distributed Computing and Networking (ICDCN18). Jan. 2018.