Handoff Management Schemes in Wireless Mesh Networks

Zhenxia Zhang

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the Ph.D. degree in Computer Science

School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

Abstract

Recent advances in Wireless Mesh Networks (WMNs) have overcome the drawbacks of traditional wired networks and wireless ad hoc networks. WMNs will play a leading role in the next generation of networks, and the question of how to provide smooth mobility for WMNs is the driving force behind the research. The inherent characteristics of WMNs, such as relatively static backbones and highly mobile clients, require new handoff management solutions to be designed and implemented.

This thesis first presents our research work on handoff management schemes in traditional WMNs. In general, a handoff process includes two parts, the MAC layer handoff and the network layer handoff. For the MAC layer handoff, a self-configured handoff scheme with dynamic adaptation is presented. Before the mobile node starts the probe process, it configures parameters for each channel to optimize the scan process. Moreover, a fast authentication scheme to reduce authentication latency for WiFi-based mesh networks is introduced. A tunnel is introduced to forward data packets between the new access router and the original reliable access router to recover data communication before the complete authentication process is finished. To minimize the network layer handoff latency, a hybrid routing protocol for forwarding packets is proposed: this involves both the link layer routing and the network layer routing. Based on the hybrid routing protocol, both intra-domain and inter-domain handoff management have been designed to support smooth roaming in WMNs. In addition, we extend our work to Vehicular Mesh Networks (VMNs). Considering the characteristics of VMNs, a fast handoff scheme is introduced to reduce handoff latency by using a multi-hop clustering algorithm. Using this scheme, vehicle nodes are divided into different multi-hop clusters according to the relative mobility. Some vehicle nodes are selected as assistant nodes; and these assistant nodes will help the cluster head node to determine the next access router for minimizing handoff latency. Extensive simulation results demonstrate that the proposed scheme can reduce handoff latency significantly.

List of Publications Related to Thesis

Refereed Journal Papers

- Z. Zhang, A. Boukerche, and H. M. Ramadan: Design of a lightweight authentication scheme for IEEE 802.11p vehicular networks. *Elsevier Ad Hoc Networks* 10(2): 243-252 (2012)

- Z. Zhang, A. Boukerche, and H. M. Ramadan: TEASE: A novel Tunnel-based sEcure Authentication SchemE to support smooth handoff in IEEE 802.11 wireless networks. *Elsevier J. Parallel Distrib. Comput.* 71(7): 897-905 (2011)

- Z. Zhang, R. W. Pazzi, and A. Boukerche: A mobility management scheme for wireless mesh networks based on a hybrid routing protocol. *Elsevier Computer Networks* 54(4): 558-572 (2010)

- R. W. Pazzi, Z. Zhang, and A. Boukerche: Design and evaluation of a novel MAC layer handoff protocol for IEEE 802.11 wireless networks. *Elsevier Journal of Systems and Software* 83(8): 1364-1372 (2010)

- Z. Zhang, A. Boukerche and H. M. Ramadan: Design and Evaluation of A Fast MAC Layer Handoff Management Scheme for WiFi-based Multichannel Vehicular Mesh Networks. Accepted by *Elsevier Journal of Network and Computer Applications*

Refereed Conference Papers

- A. Boukerche, Z. Zhang, and X. Fei: Reducing handoff latency for NEMO-based vehicular ad hoc networks. *IEEE GLOBECOM* 2011: 1-5

- Z. Zhang and A. Boukerche, R. W. Pazzi: A novel multi-hop clustering scheme for vehicular ad-hoc networks. *ACM MobiWac* 2011: 1-8

- Z. Zhang, A. Boukerche, and R. W. Pazzi: A novel network mobility management scheme for vehicular networks. *IEEE GLOBECOM* 2010: 1-5
- Z. Zhang, R. W. Pazzi, and A. Boukerche: A fast MAC layer handoff protocol for WiFi-based wireless networks. *IEEE LCN* 2010: 684-690
- Z. Zhang, R. W. Pazzi, and A. Boukerche, B. Landfeldt: Reducing handoff latency for WiMAX networks using mobility patterns. *IEEE WCNC* 2010: 1-6
- Z. Zhang, R. W. Pazzi, and A. Boukerche: Design and evaluation of a fast authentication scheme for WiFi-based wireless networks. *IEEE WOWMOM* 2010: 1-6
- Z. Zhang and A. Boukerche: An efficient MAC layer handoff scheme for WiFi-based multichannel wireless mesh networks. *IEEE ICC* 2009: 1-5
- Z. Zhang and A. Boukerche: Design and implementation of a novel MAC layer handoff protocol for IEEE 802.11 wireless networks. *IEEE IPDPS* 2009: 1-5
- A. Boukerche, Z. Zhang, and S. Samarah: A WiFi-based wireless mesh network with inter-domain mobility management. *IEEE ISCC* 2009: 857-862
- A. Boukerche, Z. Zhang, and R. W. Pazzi: A self-configured handoff scheme for IEEE 802.11-based wireless networks. *IEEE LCN* 2009: 124-129
- R. W. Pazzi, Z. Zhang, and A. Boukerche: Performance evaluation of a fast MAC handoff scheme using dynamic adjustment of scanning parameters. *ACM MSWiM* 2009: 346-352
- Z. Zhang and A. Boukerche: A novel mobility management scheme for IEEE 802.11-based wireless mesh networks. *ICPP Workshops* 2008: 73-78
- A. Boukerche and Z. Zhang: A hybrid-routing based intra-domain mobility management scheme for wireless mesh networks. *ACM MSWiM* 2008: 268-275

Acknowledgements

First of all, I thank my supervisor Professor Azzedine Boukerche for his guidance, advice and encouragement throughout my Ph.D. study. I have benefited tremendously from his vision, technical insights and profound thinking. This research work would not have been possible without his professionalism, encouragement, moral and financial support throughout the entire time of my Ph.D study. He is my great teacher.

I wish to thank the members of the Paradise Research Lab for their suggestions and helps on my work. I owe lots of thanks to Dr. Richard Pazzi, Xin Fei and Yunfeng Gu for their discussion and comments for my research work. I also wish to thank Robson Eduardo De Grande for providing consistent technical support to my working environment.

Last but not least, I thank my parents and my wife, I would never have made it without them.

Contents

List of Tables

List of Figures

Glossary

AAA: Authentication Authorization Accounting

APs: Access Points

ARP: Address Resolution Protocol

ARs: Access Routers

BMMM: Brownian Motion Mobility Model

BRMM: Bounded Random Mobility Model

CBR: Constant Bit Rate

CN: Correspondent Node

DHCP: Dynamic Host Configuration Protocol

EAP: Extensible Authentication Protocol

EAP-TLS: Extensible Authentication Protocol-Transport Layer Security

FHAP: Fast Handoff by Avoiding Probe wait

FHR: Frequent Handoff Region

GARP: Gratuitous Address Resolution Protocol

GMK: Group Master Key

GTK: Group Transient Key

HWMP: Hybrid Wireless Mesh Protocol

IGMG: Internet Gateway Multicast Group

MANET: Mobile Ad Hoc NETwork

NAT: Network Address Translation

NEMO: Network Mobility

PMK: Pairwise Master Key

PTK: Pairwise Transient Key

RADIUS: Remote Authentication Dial-In User Service

RDMM: Random Direction Mobility Model

RSS: Received Signal Strength

RWMM: Random Waypoint Mobility Model

SNR: Signal to Noise Ratio

VMNs: Vehicular Mesh Networks

VoIP: Voice over IP

WEP: Wired Equivalency Protocol

WLANs: Wireless Local Networks

WMNs: Wireless Mesh Networks

WSN: Wireless Sensor Network

Chapter 1

Introduction

Wireless Mesh Networks (WMNs) are a highly promising wireless network architecture that gained lots of advances in recent years. WMNs can use wireless connections to construct a backbone for relaying data packets. Compared to conventional wired networks, WMNs have many advantages such as easy deployment, low cost, dynamic self-configuration, and etc. Due to the inherent characteristics of wireless networks, such as low bandwidth, short transmission ranges, and the high possibility of congestion, the question of how to support smooth roaming in a wireless mesh network, considering the clients' mobility, has become a driving force behind research.

1.1 Background

In a general wireless mesh network, according to different functionalities and roles, mesh nodes can be divided into two groups: mesh clients and mesh routers [5]. Mesh clients are wireless terminals which can join and leave the wireless mesh network at any time. In addition, it is in these terminals' nature to connect to the Internet or local networks through the connection provided by the wireless mesh network. Mesh clients can be any type of communication devices with wireless antennas, such as laptops, PDAs, smart phones, wireless sensors and etc. Conversely, mesh routers are a type of special wire-

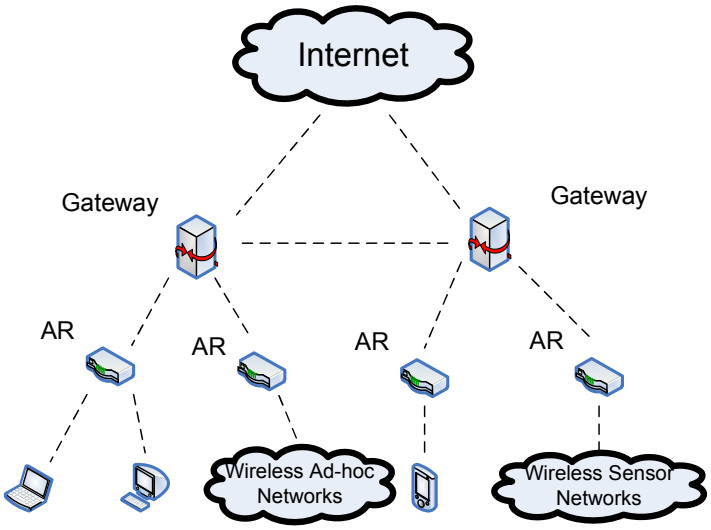


Figure 1.1: Wireless mesh networks

less nodes that provide the connection service for mesh clients. In general, there are three categories of mesh routers: gateways, relay routers and access routers. Gateways are used to provide communication among other nodes which are in a different domain through Internet or some local networks. In a wireless mesh network, the gateway is also connected with other mesh routers through wireless links. Relay routers are used to forward data packets between the gateway and access routers, and access routers (ARs) provide wireless connections for the mesh clients. Access routers are responsible for the last hop to mesh clients. Gateways, relay routers and access routers construct the relay backbone in wireless mesh networks. Usually, mesh routers are fixed nodes without mobility, and mesh clients can have greater mobility compared with mesh routers. Figure 1.1 shows a sample scenario of a wireless mesh network. In this example, there are two domains in the wireless mesh network. For each domain, a gateway is used to provide inter-domain communications. Access routers work as bridges to support data communication between gateways and mesh clients. Moreover, mesh clients can be any type of wireless nodes. Mesh clients can also be a subnetwork, such as a Mobile Ad hoc NETwork (MANET) [13] or a Wireless Sensor Network (WSN) [15].

In general, the architecture of WMNs is categorized into three types: client WMNs,

infrastructure/backbone WMNs, and hybrid WMNs [5]. In client WMNs, there are no mesh routers in the mesh networks, and all mesh nodes in the network are mesh clients. Therefore, in such mesh networks, all of the mesh nodes would have similar computing ability and communication resources. The routing, bridging and gateway functions are also implemented by the mesh clients. This type of mesh network is similar to the traditional wireless mobile ad-hoc networks. The main advantage of client WMNs is flexibility. In infrastructure/backbone WMNs, there are fixed relay backbones which are composed of relay routers, access routers and gateways. The relay backbones provide connectivity to the mesh clients, and they also provide connections with other networks in different domains. The existence of relay backbones makes the infrastructure WMNs more reliable and it also increases the throughput of the mesh networks. Therefore, infrastructure mesh networks are the most common architecture used to deploy wireless mesh networks. Finally, hybrid WMNs combine the characteristics of client mesh networks and infrastructure mesh networks. There would be also a relay backbone in hybrid WMNs. The backbone is composed of mesh clients and mesh routers, since both mesh clients and mesh routers can provide routing and gateway function. However, the backbone is not fixed in hybrid WMNs because of the high mobility of mesh clients.

A Vehicular Mesh Network (VMN) is a special wireless mesh network which is developed for high speed scenarios. Using VMNs, a lot of new applications can be introduced, such as intelligent transportation systems and the mobile entertainment environment. In a vehicular mesh network, mesh routers are usually deployed on the roadside; therefore mesh routers can be called roadside units. Mesh routers also have to provide gateway function, relay function and routing function. Conversely, mesh clients are vehicles equipped with antennas for wireless communication. Therefore, vehicles can communicate with each other in ad-hoc mode or access the Internet through roadside nodes. The communication among vehicles is called vehicle to vehicle (V2V) communication; and the communication between vehicles and roadside nodes is called vehicle to infrastructure (V2I) communication [16]. To construct VMNs, various wireless technologies

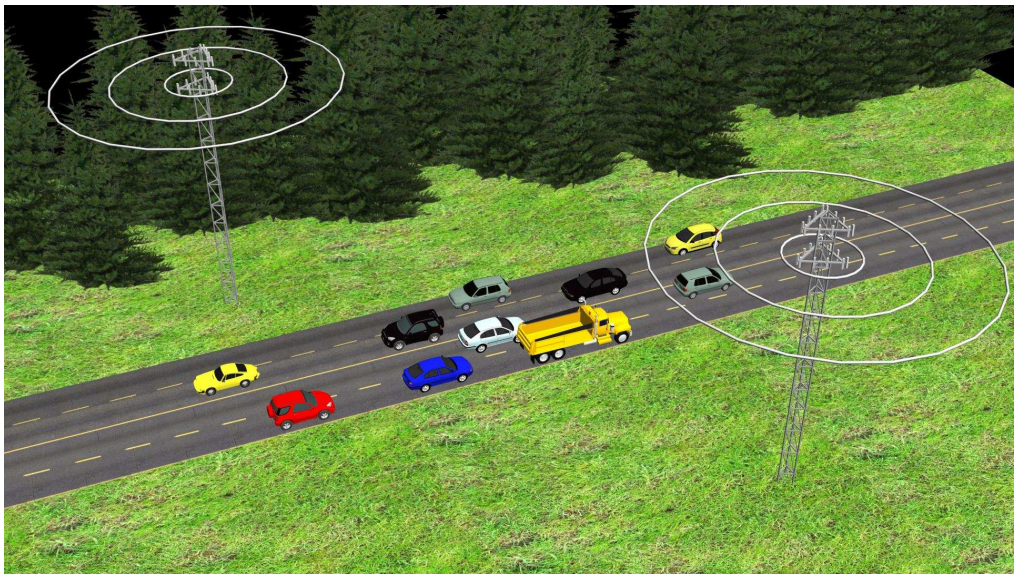


Figure 1.2: A sample scenario of VMNs in the freeway environment

can be used to provide wireless links in the physical layer, such as IEEE 802.11 [31], WiMax [35], UMTS, and etc. In a vehicular mesh network, roadside units are always static nodes and vehicle nodes usually have a higher speed when compared with mesh clients in a traditional wireless mesh network. Therefore, the schemes used to support smooth roaming in a vehicular mesh network should be designed separately. A sample scenario of vehicular mesh networks in the freeway environment is shown in Figure 1.2.

1.2 Problem Statement

The many advantages of Wireless Local Networks (WLANs) such as low cost and ease of maintenance have enabled the rapid growth and wide deployment of wireless applications. Wireless mesh networks are an important alternative to Wireless Local Area Networks. In these WMNs, wireless terminals can transmit data frames among each other through wireless connections using any standards defined in the physical layer. However, the transmission range of wireless antennas used in WMNs is still limited. For example, the regular transmission range of IEEE 802.11b/g antennas will be less than 150 meters, and

the transmission range of IEEE 802.16e antennas will be around 1000 meters. In a mobile environment, the mesh clients will have free mobility; therefore, when a mesh client moves out of the access router's coverage, the mesh client will lose wireless connection with the original access router. Consequently, the network application launched in the mesh clients, such as VoIP [37], online games, email services, and etc., will also be interrupted. The question of how to maintain network applications is very important in the mobile environment. The answer is the handoff. Handoff, which can also be called handover, is the process of switching access routers, access points or base stations during the movement of mobile nodes. During the movement, when the mesh client finds that the current wireless connection is lost, it should establish a new wireless connection with another access router. After the handoff process is completed, the mesh client can continue the interrupted network applications [14, 13, 15].

Although the handoff solves the problem of how to support free mobility in the mobile environment, it introduces another question for WMNs. This question is how to support smooth roaming in WMNs. Smooth roaming means that during a mesh client's movement, the user would not notice that the mesh client does handoffs to maintain wireless connections. During the handoff process, the mobile mesh client will disconnect with the original access router. The network communication is interrupted in this phase. The time spent to complete the handoff process is called handoff latency. If the handoff latency is too long, large jitter will appear and the quality of service will not be acceptable for users. Therefore, the handoff latency should be minimized to support smooth roaming in WMNs. In WMNs, frequent handoffs are inevitable for mobile clients to keep network connections during their movement; as a result, smooth roaming is very important in order to support real-time applications in wireless mesh networks.

Real-time applications have strict requirements on time synchronization. Long handoff latency introduces long interruption time, and long interruption time causes large jitter. Large jitter is not tolerable for real-time applications because it breaks the time synchronization. As a result, the user will have a bad user experience on real-time ap-

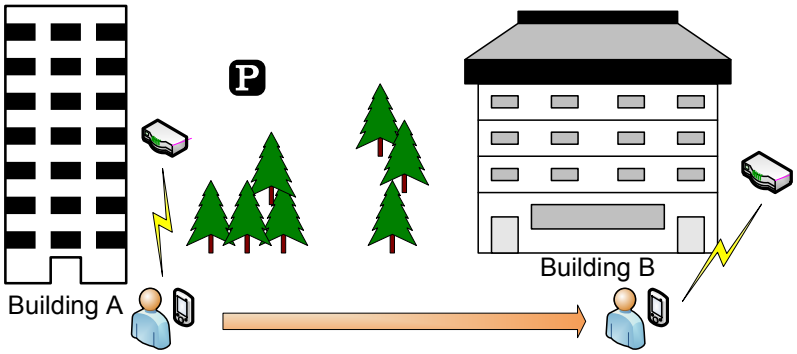


Figure 1.3: An application scenario featuring handoff

plication. For example, as illustrated in Figure 1.3, if we have a conference call using Skype [68] and we have to move from one building to another building in campus, large jitter caused by handoffs would harm the quality of video conference. It is possible that we cannot hear anything during the handoff process and the video would also be stuck during the movement. Smooth roaming is also very important for VMNs. In a VMN, the mesh clients which are vehicles have much higher speed compared with the clients in traditional WMNs. Consequently, the handoff occurs more frequently in vehicular mesh networks; and the quality of service is significantly affected by the handoff latency. For instance, a user takes the bus for a long distance trip, and the user would like to play an online-game during the trip, such as World of Warcraft [78]. If there is no smooth roaming supported, when the computer has to switch access routers, large jitter is introduced; and the control of his character will be lost. The character then would lose a battle or be killed by the enemy. Thus, smooth roaming is needed to be implemented in WMNs.

Handoff schemes are solutions to implement handoff processes. An efficient handoff scheme can reduce handoff latency and minimize packet loss ratio. Figure 1.4 illustrates components of a complete handoff scheme. In general, a handoff process can be divided into two phase: the MAC layer handoff and the network layer handoff. When a mesh client finds that the quality of the signal, which can be measured by the Received Signal Strength (RSS) or the Signal to Noise Ratio (SNR) below the predefined level in the

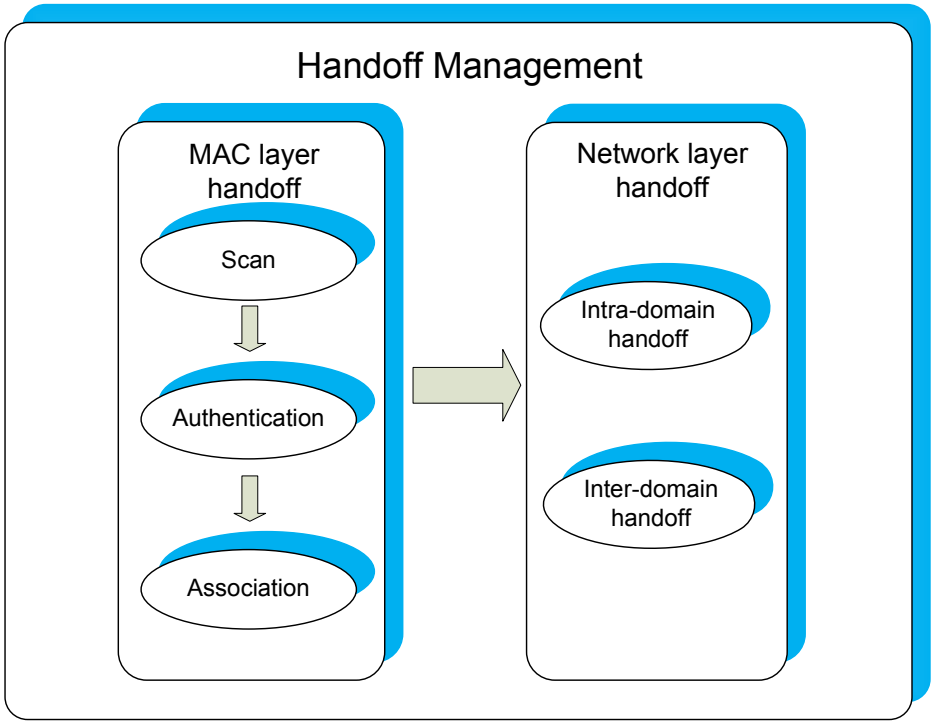


Figure 1.4: Components of a handoff management scheme

MAC layer, the mesh client initiates the MAC layer handoff process to establish wireless connection with the new access router in the physical layer. The MAC layer handoff process includes three steps: scan, authentication and association. In the first step, the mesh client has to find the new access router that has the best quality of signal in its neighborhood. Usually, there are two scan methods that can be used to probe available access routers: passive scan and active scan. In the passive scan mode, mesh clients capture beacon messages from access routers in each channel, and select the routers with the best signal quality as the next access routers. The waiting time in a passive scan should be long enough to guarantee that mesh clients can receive the beacon messages. On the other hand, active scan allows mesh clients to broadcast probe request messages in different channels, and when access routers receive the request messages, they reply with response messages. Using these responses, mesh clients obtain access routers' signal strength and determine the new access routers. Compared to passive scan, active scan has

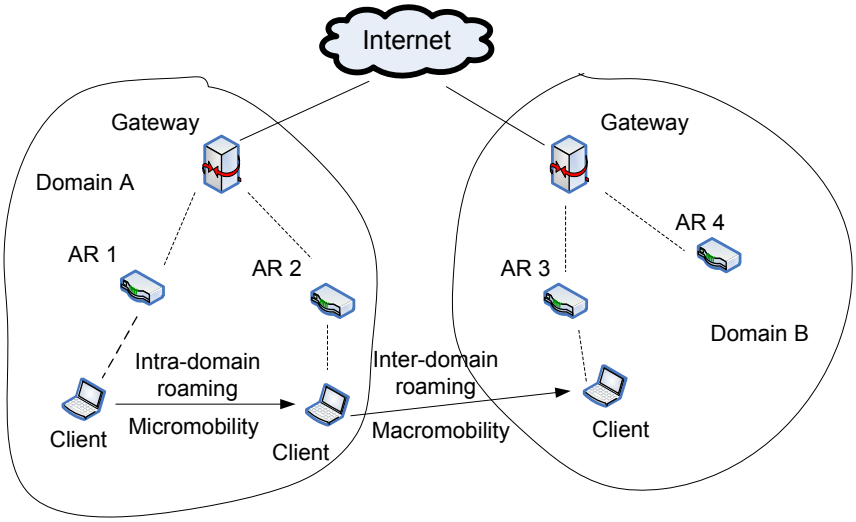


Figure 1.5: Intra-domain roaming and inter-domain roaming

less waiting time and is therefore more suitable for realizing smooth roaming in WMNs. Then, the authentication process is triggered between the mesh client and the new AR. If the authentication passes, the mesh client can connect to the new AR; otherwise, the association is denied.

After completing the MAC layer handoff, the network layer handoff is triggered. During the network layer handoff, the mesh client has to establish the logical connection with its correspondent nodes. There are two kinds of roaming in wireless mobile networks: intra-domain roaming and inter-domain roaming. For intra-domain roaming, the mesh client moves in the same domain; therefore, the mesh client can maintain the same IP address and it only needs to update the routing path between the gateway and itself. For inter-domain roaming, the mesh client moves to a new domain, and it has to receive a new IP address from the new server. As a result, the mesh client has to update its IP stack and establish new routing paths between correspondent nodes and itself in the network layer handoff. Figure 1.5 illustrates two kinds of roaming in wireless networks.

Unfortunately, the original handoff schemes introduce high handoff latency and packet loss ratio in the handoff process. It causes long delays in the network and degrades the quality of time-critical applications, such as VoIP. For instance, the network layer handoff

will spend more than 1 second which is not suitable for real-time applications in WMNs. Thus, in this thesis work, efficient handoff schemes are designed and implemented to support smooth mobility in WMNs.

1.3 Research Objective

The goal of this thesis work is to design and implement efficient handoff management schemes to support smooth mobility in WMNs. The research objectives are concentrated as follows:

- *Real-time:* Real-time applications are widely used in computer networks, such as VoIP, online games, etc. A long handoff process will cause large jitter and packet loss ratio, which would interrupt real-time applications and cause a bad user experience. Therefore, recovering data communication after the handoff immediately is an important objective of the thesis work. The best scenario is that during the movement, the user would not notice that the handoff process occurs, which means the mesh client switches its access router smoothly. To support smooth roaming under the real-time environment, the total handoff latency should be minimized, which includes the MAC layer handoff latency and the network layer handoff latency. The packet loss ratio should be reduced and the packet delay variation should also be controlled.

- *Security:* The WMN is an open environment, because wireless signals are easy to be recorded by malicious nodes. Thus, how does one prevent malicious nodes from viewing data packets? How does one allow authenticated users to access data information? And how can malicious nodes be prevented from modifying data packets? These points should be considered when implementing the handoff scheme. Therefore, the authentication process is needed to be completed during the handoff. However, the authentication process introduces extra latency to the

handoff. To support smooth mobility, a fast secure authentication scheme is needed to be implemented.

- *Micromobility and macromobility:* According to two kinds of roaming in wireless networks, there are two kinds of mobility: micromobility and macromobility. Micromobility refers to intra-domain roaming, which means the mesh client moves across different subnets in the same domain. Macromobility refers to inter-domain roaming, which means the mesh client moves among different domains. Figure 1.5 shows examples of the two kinds of mobility in wireless networks. For the network layer handoff, both micromobility and macromobility should be supported by the proposed handoff scheme.

- *Scalability:* When the mesh network scales up, multiple mesh clients will concurrently do the handoff process. In wireless environments, the handoff management packet transmission between the access router and the mesh client will block the communication between the access router and other mesh clients. Therefore, handoff latency reduction improves the scalability of the wireless mesh network. Secondly, the transmission overhead which is caused by handoff management packets also increases when the network grows. As a result, to have high scalability, the overhead introduced by the handoff scheme should also be reduced.

- *Accuracy:* Accuracy means the mesh client can select the optimized access router. The best case is when the mesh client selects the new access router which will maintain the wireless connection between the access router and the mesh client as long as possible. A bad sample scenario is when a mesh client completes the handoff process, it finds that the signal quality of the current access router drops quickly and it has to switch to another access router in a very short period. The high frequency of the handoff process has a negative impact on the quality of service provided by the wireless mesh network. The ping-pong effect should be erased.

- *High speed environment:* The VMN is a special kind of WMN in which mesh clients

have higher moving velocity compared with mesh clients in traditional wireless mesh networks. High speed introduces more frequent handoffs. Therefore, to support smooth mobility for VMNs, handoff schemes should be implemented based on the inherent characteristics of vehicular mesh networks, such as high speed, constrained moving paths, extra position systems, etc.

1.4 Contributions

This thesis proposes efficient handoff schemes to solve the problem of supporting smooth mobility in wireless mesh networks. The fundamental idea is to reduce handoff latency in the MAC layer and the network layer separately. Moreover, for vehicular mesh networks, handoff process is completed using a group-based strategy. The original contributions of this thesis are illustrated as follows.

- A self-configured scan scheme with dynamic adaptation is proposed to reduce handoff latency in the MAC layer. Using this scheme, the values of *MinChannelTime* and *MaxChannelTime* are changed for each handoff. The mesh client uses the scan result of the last scan or the available ARs' information in its neighborhood to set the value of *MinChannelTime*. If the mesh client receives any frame during *MinChannelTime*, it sets the value of *MinChannelTime* based on the signal strength of the captured frame. To reduce the number of scanned channels in the probe process, the scan can be terminated if the mesh client finds that an available AR has excellent signal quality to provide wireless connection. Therefore, both the waiting time for each channel and the number of scanned channels are minimized using our scheme.

- A fast authentication scheme is proposed to achieve smooth handoff in wireless mesh networks. A tunnel is introduced to forward data packets between the new access router being verified and the original reliable access router. The temporary

tunnel key is introduced to encrypt the wireless communication before the mesh client finishes a complete authentication process. The temporary tunnel key is generated based on the MAC addresses of the new access router and the mesh client. Upon receiving the temporary tunnel key, a secure tunnel is established between the mesh client and the new access router. The mesh client can use this tunnel to communicate with its correspondent nodes for maintaining the existing applications. As a result, the security of handoff is achieved without increasing overhead to authentication servers, and authentication latency can be minimized to support smooth roaming.

- To support smooth mobility in the network layer, a hybrid routing protocol for forwarding packets is proposed: this involves both the link layer routing and the network layer routing. Based on the hybrid routing protocol, a network layer handoff scheme for WMNs is presented. Both intra-domain and inter-domain roaming have been considered to support smooth roaming in WMNs. During intra-domain handoff, gratuitous ARP messages are used to provide new routing information, thus avoiding re-routing and location updating. In addition, using layer 2 routing can minimize the cost of packet relay among the mesh routers when compared with the layer 3 routing. Moreover, unlike tunnel-based solutions, the tunneling overhead at each hierarchy is removed. For inter-domain handoff, gratuitous ARP messages are also used to update routing information. Redundant tunnels are removed in order to minimize forwarding latency.

- For vehicular mesh networks, a novel multi-hop clustering scheme is presented to establish stable vehicle groups. To construct multi-hop clusters, a new mobility metric is introduced to represent relative mobility between vehicles in multi-hop distance. Vehicles broadcast beacon messages periodically, and they use the ratio of packet transmission delay between two successive beacon messages to represent the relative mobility between two vehicle nodes. In addition, the vehicle nodes which

have a low aggregate mobility will work as the cluster head nodes to construct multi-hop clusters. To the best of our knowledge, this is the first multi-hop clustering scheme for vehicular mesh networks. Based on the multi-hop clustering scheme, a fast handoff scheme is proposed to reduce handoff latency. Within each cluster, the network mobility solution is used to reduce the total number of handoff processes. Moreover, before the mobile routers start actual handoffs, assistant nodes help the cluster head nodes receive new care of addresses to reduce handoff latency.

- Extensive simulation experiments are run to illustrate the performance of proposed handoff management schemes. In addition, the comparison with other existing schemes is also presented to demonstrate the advantages of the proposed schemes.

1.5 Thesis Organization

This thesis presents our research work on the handoff management schemes in WMNs. It includes 6 chapters, and it is organized as follows.

- Chapter 1 introduces the background and the importance of this work. The research objectives and original contributions are also given in this chapter.

- Chapter 2 introduces related work published in past years. A comprehensive literature review is given on both the MAC layer handoff solutions and the network layer handoff solutions. The advantages and disadvantages of previous work are illustrated. Moreover, handoff solutions designed for vehicular networks are also reviewed.

- Chapter 3 presents the research work on the MAC layer handoff management. The work includes two parts: reducing scan latency by introducing a self-configured handoff scheme with dynamic adaptation and proposing a fast secure authentication scheme in the MAC layer.

- Chapter 4 proposes the research work on the network layer handoff management over traditional WMNs. A hybrid routing protocol is first introduced. Based on this hybrid routing protocol, both the intra-domain and inter-domain handoff management schemes are proposed.

- Chapter 5 illustrates the research work on the handoff management over VMNs. First, a multi-hop clustering protocol for VMNs is proposed. Based on the multi-hop cluster protocol, a fast handoff scheme is presented based on Network Mobility (NEMO). The simulation results are finally illustrated.

- Chapter 6 concludes the thesis work and lists the possible improvements for future work.

Chapter 2

Related work

In this chapter, the previous work on handoff management schemes over wireless mesh networks is reviewed. Due to the fact that the transmission range of the AR in WMNs is very limited, mesh clients need to switch among ARs to maintain their network connections, which is termed as the handoff process. The handoff process technically consists of two stages which occurs in the MAC layer and the network layer separately. Both the MAC layer handoff solutions and the network layer handoff solutions are reviewed in this part. The advantages and disadvantages of the solutions are illustrated. Finally, the handoff schemes designed for vehicular networks are also introduced.

2.1 MAC Layer Handoff over Traditional WMNs

To construct wireless mesh networks, a lot of transmission protocols can be used, such as IEEE 802.11 [31], Bluetooth [34], WiMax [35] and etc. Among these protocols, the IEEE 802.11 standard protocol is a promising option according to its broad bandwidth and low equipment cost. Therefore, we assume that IEEE 802.11 is used to provide wireless connections in wireless mesh networks in our thesis work. Mobile nodes in IEEE 802.11-based networks obtain network connections to the Internet through Access Points (APs), which construct the backbone of the network by wired/wireless commu-

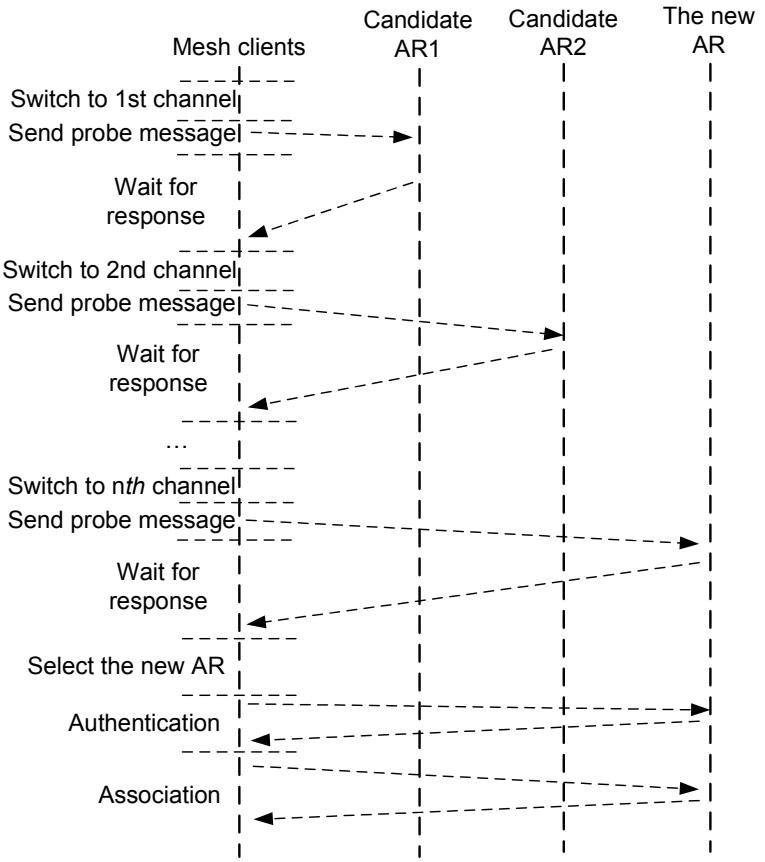


Figure 2.1: The MAC handoff procedure using active scan in IEEE 802.11 [47]

nication between each other. In an IEEE 802.11-based WMN, mobile nodes are mesh clients and access points are access routers.

For a standard MAC layer handoff in WMNs, it includes three steps: scan for available ARs, reauthentication and reassociation with the new AR [47]. To scan for available ARs, there are two approaches: passive scan and active scan. In a passive scan, the mesh client monitors beacon messages broadcasted by available ARs. Conversely, in an active scan, the mesh client switches to candidate channels, broadcasts probe request messages and selects the new AR according to the response messages received from ARs. Compared with the passive scan, the active scan is more appropriate for real-time applications since the passive scan has to synchronize with ARs and it spends more time to complete the handoff. The MAC layer handoff process using active scan in IEEE 802.11 networks is

Table 2.1: Comparison of the MAC layer handoff scheme with active scan

Solution	Category	Strategy
Mishra *et al.* [47]	Reducing waiting time	$MinChannelTime = 6.5\ ms$ $MaxChannelTime = 11\ ms$
Velayos *et al.* [73]	Reducing waiting time	$MinChannelTime = 1\ ms$ $MaxChannelTime = 10.24\ ms$
Chintala *et al.* [20]	Reducing waiting time	avoiding probe wait
Shin *et al.* [66]	Waiving unnecessary scans	neighbor graph
Shin *et al.* [67]	Waiving unnecessary scans	using channel mask to enable selective scan
Mustafa *et al.* [50]	Waiving unnecessary scans	using a new threshold to pre-scan
Liao *et al.* [45]	Waiving unnecessary scans	scanning channels by groups
Ramachandran *et al.* [58]	Waiving unnecessary scans	using two antennas

illustrated in Figure 2.1. Most improved handoff schemes proposed in literature adopt the active scan.

2.1.1 Handoff with Active Scan

When the mobile node senses that the signal quality of the current AR is below the threshold, it triggers the scan process. The mobile node switches to each candidate channel, broadcasts probe request messages, collects response messages and saves the ARs' information in the AR list. After the scan is finished, the mobile node selects the new AR according to the signal quality. Then the mobile node authenticates with the new AR. The handoff process is completed after the mobile node associates with the new AR and the physical link is established. Because the scan phase consumes most of time during the basic handoff process [47, 73], reducing the scan latency is one of most efficient approaches to minimize the handoff latency. The total scan time is affected by both the waiting time for each channel and the number of probing channels. In consequence, to reduce the scan latency, two approaches can be used: shortening the waiting time of channels and reducing the number of scanned channels. Table 2.1 illustrate the comparison of the MAC layer handoff schemes proposed in recent years.

Reducing waiting time for each channel is a reasonable solution to reduce scan latency. $MinChannelTime$ and $MaxChannelTime$ are two important parameters in active scan. When a mesh client starts the handoff process in the MAC layer, it switches to a selected channel and broadcasts a probe request message to detect available access routers in the channel. In the meantime, the mesh client initializes a timer to control the waiting time for response messages from access routers. If there is no response or data traffic detected during $MinChannelTime$, the channel is considered as an empty channel. Otherwise, the timer is extended to $MaxChannelTime$ and the mesh client must wait for $MaxChannelTime$ to collect probe responses. This said, waiting time in each channel is significantly affected by $MinChannelTime$ and $MaxChannelTime$. However, the values of $MinChannelTime$ and $MaxChannelTime$ are not given in IEEE 802.11 standard.

Mishra *et al.* [47] did lots of comparative tests in their testbed and suggested that the $MinChannelTime$ could be *6.5 ms* and the $MaxChannelTime$ could be *11 ms*. Velayos *et al.* [73] proposed a formula to calculate the MAC layer handoff time; and based on their experiments, the search time of the active scan can be reduced by 20% when the $MinChannelTime$ is set to *1 ms* and the $MaxChannelTime$ is set to *10.24 ms*.

Chintala *et al.* [20] proposed a fast handoff scheme by waiving the waiting time using the inter-AP communication, which refers to Fast Handoff by Avoiding Probe wait (FHAP). During the scan phase, after the mobile node broadcasts the probe request message, it switches to the next candidate channel directly without waiting for the response messages. Upon receiving the probe request message, instead of sending back the probe reply message, the neighboring AR sends the response message to the current served AR through the backbone. After the mobile node broadcasts probe request messages in all candidate channels, it switches its working channel to its original channel and connects to the original access point to collect response messages. Therefore, the main advantage of FHAP is that the mobile client does not need to wait for the reply, and it can directly switch to the following channels. Since the connections between access points in WMNs could be wireless, multi-hop inter-AP communication

introduces more latency and a higher loss ratio than the wired backbone. Especially, this scheme is not suitable for multichannel WMNs. In multichannel WMNs, to increase the capacity of the backbone, mesh routers equipped with multi-radios are allowed to work in different channels; and the mesh routers can send and receive packets in different frequencies simultaneously by utilizing non-interfering channels [60, 52, 38]. Therefore, in multichannel WMNs, less location distance does not mean fewer hops and sometimes communication between neighboring access point needs serval hops which introduces latency for collecting response messages.

Another means of reducing handoff latency is waiving unnecessary scans. Shin *et al.* [66] proposed an algorithm to obtain the relationship among ARs using neighbor graphs and non-overlapping graphs. Using the relationship, the mobile node can predict available ARs for the handoff. In addition, channel mask was introduced to predict the availability of the channel [67]. In the channel mask, the values of available channels are 1 and the values of empty channels are 0. An algorithm is introduced to update the channel mask, and the mobile node scans channels according to the corresponding value in the channel mask. This scheme is called selective scan. Mustafa *et al.* [50] introduced a new threshold to trigger the pre-scan phase before the handoff is initialized. Using pre-scan can shorten the interruption time during the handoff.

Moreover, the performance of active scan can be improved using other solutions. Liao *et al.* [45] proposed a smooth scan scheme. All of channels are categorized into groups, and during the scan phase, the mobile node scans all of channels by groups. After channels in a group are scanned, the mobile node returns to the former working channel to receive data frames buffered in the serving AR during the scan phase. Then, the mobile node scans the channels in the next group. This method can reduce packet loss ratio and lower the jitter obviously. Ramachandran *et al.* [58] introduced a scheme to reduce handoff latency by using two antennas. In the scan phase, one antenna is responsible for probing channels and another one is responsible for transmitting and receiving data packets. The drawback of this solution is that the extra physical equipment is needed.

2.1.2 Secure Authentication

In previous IEEE 802.11 networks, Wired Equivalency Protocol (WEP) is used to support authentication service. Using WEP, a key is shared with mobile clients and access points. In addition, two kinds of authentication schemes can be used to support secure connections: shared key authentication and open system authentication. Using shared key authentication, when the mobile terminal sends an authentication request to the new access point, the new access point sends a data packet to the mobile terminal. The mobile terminal uses its WEP key to encrypt the data packet and sends the result back to the access point. Upon receiving the result, the access point checks the encrypted message. If the result matches, the access point allows the mobile terminal to access the wireless network; otherwise, the connection request is denied. Open system authentication is another means of authentication. Using open system authentication, when the mobile terminal sends an authentication request to the access point, the access point sends the authentication response message back to the mobile terminal and allows the mobile terminal to access the wireless network by default. The challenge process is waived and the WEP key is used to encrypt all of the following data packets. Unfortunately, following studies show that WEP is vulnerable to many attacks [24].

IEEE 802.11i protocol [31] is then introduced to improve the security of wireless networks. In IEEE 802.11i networks, only open system authentication is supported. After selecting the new access point based on the signal quality, the mobile terminal sends an authentication request message to the access point and the access point sends back the authentication response message to the mobile terminal. Then, the Extensible Authentication Protocol (EAP) [2] process is triggered to generate and exchange keys. Usually, EAP-TLS [3] protocol is used for the authentication. Moreover, Remote Authentication Dial-In User Service (RADIUS) [62, 63] is used for communication between the access point and the authentication server which would provide Authentication Authorization Accounting (AAA) service. After completing the EAP process, the authentication server passes the Pairwise Master Key (PMK) and the Group Master Key (GMK) to the mo-

bile terminal. Finally, the four-way handshake process is used to derive the Pairwise Transient Key (PTK) and the Group Transient Key (GTK). However, the EAP-TLS protocol spends a lot of time in the handoff process because the mobile terminal should communicate with authentications server to generate keys. The whole process spend more than 1 sec [48]. Therefore, some approaches are proposed to reduce the authentication delay.

To reduce the authentication delay, IEEE 802.11i [31] proposed a preauthentication scheme. Before the mobile terminal starts the handoff process, it can start EAP-TLS authentication with the new access point through the connection with the old access point. The old access point will forward authentication messages between the mobile terminal and the new access point. This preauthentication process ends when the mobile terminal and the new access point receive the new PMK. As a result, when the handoff starts, only the four-way handshake is needed to complete the authentication. However, the way of how to predict the new access point is not defined in IEEE 802.11i. An inaccurate prediction will cause large resource waste.

Predictive authentication scheme is proposed in [54] using Frequent Handoff Region (FHR). Pack *et al.* used a statistical method to model the mobile terminal's mobility pattern. A set of access points are selected as the FHR access points with which are probably associated. Before the handoff, the mobile terminal sends the authentication request to the authentication server and the authentication server sends the authentication response back to the FHR access points with authentication information. Then, during the handoff, the mobile terminal needs to exchange fewer messages with the new access point.

Mishra *et al.* [48] used neighbor graph to extract the relationship among access points. Using neighbor graph, candidate access points with which would be associated can be determined before the handoff occurs. Therefore, the key materials can be distributed to the candidate access points before the handoff. In addition, PMK trees are adopted to generate new PMKs. This scheme is called proactive key distribution. Because the

access point receives key materials before the handoff, the communication between the access point and the authentication server can be waived and the handoff latency is reduced. However, this scheme introduces communication overhead between candidate access points and the authentication server.

To reduce the authentication delay further, proactive key distribution with anticipated four-way handshake is introduced in [39]. After the authentication server extracts the neighbor access points list, it sends MAC addresses of neighboring access points to the mobile terminal. As a result, the mobile terminal can generate PTKs before the handoff to waive the four-way handshake. Because this scheme is based on the proactive key distribution, it has the same problem with the proactive key distribution. It introduces extra load to the authentication server and some resources are wasted.

2.2 Network Layer Handoff over Traditional WMNs

The access routers involved in the network layer handoff may belong to the same network (intra-system or intra-domain) or to two different networks (inter-system or inter-domain) [4, 6]. Therefore, there are two kinds of network layer handoffs: intra-domain handoff and inter-domain handoff.

Mobile IP [55] is one of the most significant solutions for inter-domain handoff. It solves the mobility problem by using the tunnel technology. The original permanent IP address of the mobile node is called the home address, which is administrated by the home agent. When the mobile node moves to a foreign network, the foreign agent (usually a router) assigns the second temporary address known as the care of address to the mobile node. After receiving the care of address, the mobile node registers the secondary address to the home agent. Then all of the packets sent to the mobile node's home address are tunneled to the mobile node's care of address using IP encapsulation by the home agent. The architecture of Mobile IP is shown in Figure 2.2.

In terms of intra-domain handoff management, a variety of solutions has been pro-

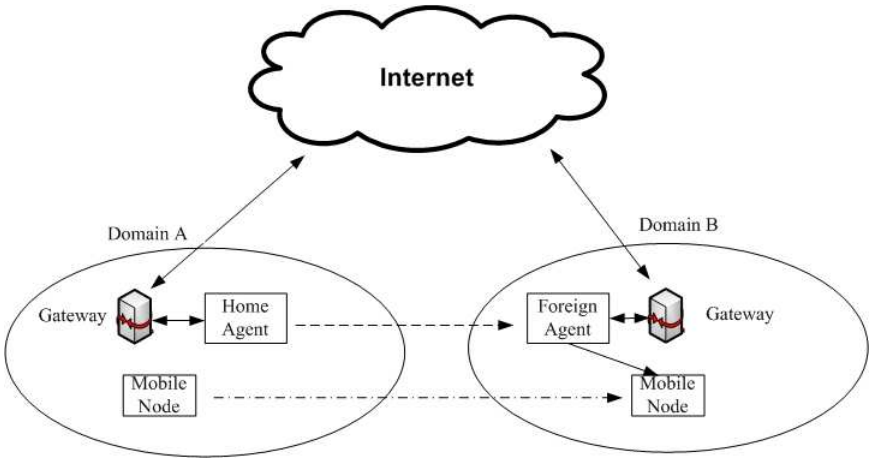


Figure 2.2: Architecture of Mobile IP

posed [17, 27, 49, 59, 69]. These solutions can be categorized into two types: tunnel-based and routing-based [18]. A hierarchical architecture is always adopted by tunnel-based solutions. The low level agent's address is encapsulated in an extra IP header by the high level agent. When the low level agent receives the packets, it decapsulates the packets to fetch the client address and forwards the packets to the client. For routing-based solutions, routing tables are updated in mesh routers to re-establish the connection after the handoff.

Hierarchical Mobile IPv6 [69] is one of the tunnel-based schemes, and it introduces a Mobility Anchor Point (MAP) between the home agent and the mobile node. When a mobile node moves between the ARs in the same MAP's domain, it only needs to register its address in the MAP instead of updating its new address at the home agent. This reduces the overhead of registering with the home agent when the mobile node moves within the same domain. HAWAII [59] is a routing-based solution. It uses a gateway called a domain root router to handle handoff management. When the mobile node roams within the same domain, if handoff occurs, the gateway should find a new route from the gateway to the mobile node.

2.2.1 Intra-Domain Handoff Management over WMNs

The intra-domain handoff management schemes that improve handoff performance over WMNs can be classified into three groups: tunnel-based, routing-based, and multicast-based. The features of the first two groups, tunnel-based and routing-based, are similar to those used in wireless networks. The multicast-based solution is a new kind of solutions that assigns the ARs of the same mesh client to two multicast groups, and the multicast is applied to support seamless handoff management for real-time applications such as VoIP and etc. The comparison of the solutions that are discussed in the following is shown in Table 2.2.

Huang *et al.* proposed the hierarchically structured Mesh Mobility Management (M^3) in [30]. There are three types of mesh routers employed: gateways, superior routers, and access points (APs). The architecture is shown in Figure 2.3. Superior routers are some special mesh routers which are used to gather the location information of the mesh clients in the vicinity of subordinate APs. When a new mesh client joins the mesh network, it registers its location information with the gateway, and the related superior router also keeps a copy of the location information of the new mesh client. To forward packets to the Internet, the APs send packets to the gateways using default routes. Tunneling is used for the reverse direction: packets flow from the Internet to mesh clients. Superior routers encapsulate packets with the destination APs' addresses in an extra IP header. The destination APs decapsulate the packets and send them to the mesh client. When the mesh client moves to another AP's vicinity, handoff is activated. The new AP sends a handoff request message to the former AP. In response to the request, the prior AP sends the subscriber's information of the mesh client to the new access point, and adds a temporary routing entry for forwarding the packets to the new AP. Therefore, all packets received by the previous AP will be forwarded to the new AP after handoff. Location information of the mesh client at the superior routers is updated after a certain period, and updating the location information at the gateway will also be delayed by the superior router.

Table 2.2: Comparison of intra-domain solutions

	M^3 [30]	Ant [74]	iMesh [51]	MEMO [61]	Mobile Party [64]	SMesh [8]
Group	tunnel-based	tunnel-based	routing-based	routing-based	routing-based	multicast-based
Layer	layer-3	layer-3	layer-2+3	layer-2+3	layer 3	layer 3
Location server	$\surd$	$\surd$	N/A	N/A	N/A	N/A
Routing update	N/A	N/A	$\surd$	$\surd$	$\surd$	N/A
Routing	not mentioned	OLSR	OLSR	AODV-MEMO	Mobile Party	not mentioned
Scalability	normal	normal	low	low	normal	normal
Overhead reason	hierarchical delay	location update	routing	routing	routing	group management
Overhead level	normal	normal	high	high	high	normal

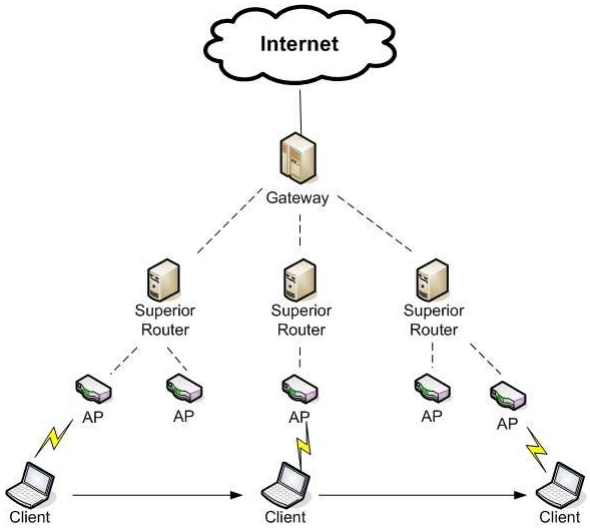


Figure 2.3: Architecture of the structured Mesh Mobility Management (M^3)

This location update scheme decreases the cost of location management compared to updating the location information at the gateway immediately. The hierarchical architecture decreases the delay of querying the client's location information at the gateway and increases the scalability of the mesh network; however, the overhead of encapsulation and decapsulation can not be waived.

Wang *et al.* [74] introduced a network-based intra-domain fast handoff scheme called Ant. When the mesh client decides to trigger the handoff, it sends a de-association message to the prior AR. Upon receiving the de-association message, the prior AR begins to buffer all the packets forwarded to the mesh client and sends handoff notification messages to all of its neighboring ARs containing the mesh client's ID. When another AR detects an association or re-association event from the mesh client determined by the same client's ID, it forwards a handoff confirmation message to the previous AR and it also sends a location update message to the location server in the mesh network. Then a temporary tunnel is set up to forward the buffered packets between the former AR and the new AR. Next the new data path for forwarding packets to the correspondent node's AR is established. All tunnels between the neighboring ARs are set up in advance to decrease the latency of establishing a temporary tunnel. After the new data path is

established, the whole process is finished. This approach can eliminate the amount of packet loss occurring during handoff if the buffer is big enough. The main cost is updating the location information immediately in the handoff process, and pre-establishing all bi-directional tunnels causes other costs.

Navda *et al.* [51] designed an infrastructure-mode wireless mesh network, which refers to iMesh, to provide wireless connection for mobile clients. It is a IEEE 802.11-based infrastructure WMN. In iMesh, the OLSR [21] routing protocol is used to search for data paths. When a mobile client decides to trigger the handoff, at link layer it broadcasts a probe message to all ARs in its neighborhood, receives the response messages from ARs, and selects the AR which has the best link quality as the next AR. Then, at network layer, the routing protocol is activated to update the routing tables in its relative ARs. The overhead from the routing table updates and control messages in OLSR constrains the scalability of the mesh network. Moreover, the mobile client only scans the mesh routers in one channel, if all of the channels are scanned in the MAC layer, the handoff latency would be very large.

MEMO (MEsh networks with MObility management) [61] is an applied wireless mesh network similar to iMesh. A modified AODV protocol, AODV-MEMO [61], is used to maintain the routing table. When the handoff is triggered, routing tables in ARs are updated to maintain the connections. Similar to iMesh, the overhead introduced by routing table updates constrains the scalability. Moreover, unchanged private IP addresses are assigned to mobile clients by a simple hash function, thus the possibility exists that the IP addresses might conflict.

Sabeur *et al.* [64] proposed a novel handoff management protocol, Mobile Party, which is integrated with a special routing protocol. It supposes that every mesh node has a unique ID, including mesh routers and mesh clients. When a mesh client joins the mesh network, it receives a temporary address from its parent node according to its location. The parent node of the mesh client, which can be a mesh router or a mesh client, determines the address of the mesh client. For example, the first mesh node is considered

as the root node, and its address is 000. The second mesh node is the first child node of the root node, thus its address is 100. The second child node of the root node should be assigned as 200. Moreover the first child node of 100 should be 110 and so on. These temporary addresses could indicate the mesh nodes' current location. When the mesh client wants to send a packet, it looks up the routing table to find the node which shares the longest prefix with the destination address, and forwards the packet to the next hop. Otherwise, the packet will be transmitted to the client's parent node. Each mesh node has a Rendezvous node which stores the mapping relationship between the node's unique ID and the node's temporary address. The mapping information will also be stored at the correspondent node and every node in the path to the Rendezvous node. All of the mapping information should be updated after the handoff. Thus the communication between the mesh client and its correspondent node can remain unchanged. The main overhead of mobile party is caused by the update of the mapping information. When the mesh network scales up, the overhead should be very large, because this protocol needs to maintain the mapping information. Thus its scalability is not very good.

Amir *et al.* [8] proposed a seamless 802.11 wireless mesh network, which refers to SMesh, to provide fast handoff for real-time applications. The Spines messaging system [7, 72] is used to support communications among mesh nodes. In SMesh, each mesh client receives a private address from the DHCP server by a hash function when it joins in the network, and sets the default gateway address to a single global IP address. All gateway nodes construct an anycast group, and transmit packets to Internet destinations using Network Address Translation (NAT) [23].

There are two multicast groups associated with a mesh client: Client Control Group and Client Data Group. A client's Control Group consists of ARs in the vicinity of the mesh client. When the mesh routers in a client's Control Group receive a broadcasting DHCP request, they calculate the link quality using the number of DHCP requests received in a past period and exchange the results with each other. Then the mesh router with the best link quality joins the client's Data Group. If a mesh router in the

Data Group receives a message notifying it that other ARs have better link qualities, it will leave the Data Group. This schema ensures that there is at least one AR in the Data Group and that few duplicate packets are introduced.

All of the packets sent to the mesh client are forwarded to its Data Group, and then the packets are sent to the client by the router in the Data Group. When an AR that has the best link quality changes, the AR forces the mesh client to change the associating AR instead of letting the mesh client decide. A gratuitous ARP [56] message with the new AR's MAC address is sent to the mesh client. Upon receiving the gratuitous ARP message, the client maps the new AR's MAC address to the default gateway IP address. Therefore, after finishing the handoff, the default gateway's IP address is still the same, but the MAC address has changed.

However, while using multicast groups of ARs ensures seamless handoff, maintaining the multicast group causes the extra overhead. A certain amount of bandwidth is consumed by the messages exchanged between ARs. Another drawback caused by multicast is that all ARs have to work in the same channel. This fact decreases the access capacity of the mesh network significantly.

To minimize packet loss during handoff, Wei *et al.* [76] introduce two caching mechanisms: En-route caching and promiscuous caching. Using Enroute caching, the ARs in the neighborhood of the original AR that are also along the route from the gateway to this AR cache the packets dispatched to the original AR. In promiscuous caching, all ARs in the vicinity of the original AR snoop and cache the packets forwarded to the original AR. An example of En-route caching and promiscuous caching is shown in Figure 2.4. Suppose En-route caching is used for sending packet 1 and promiscuous caching is used when packet 2 is sent. Packet 1 is forwarded following this path: Gateway $\rightarrow$ AR1 $\rightarrow$ AR3 $\rightarrow$ Client. Since AR1 is at the path from the gateway to the client, packet 1 should be cached in AR1's buffer using En-route caching. Packet 2 is forwarded following the same path of packet 1. But besides AR1, AR2 and AR4 will also store packet 2 in their buffer. Both of the two caching schemes can improve the performance of the handoff and

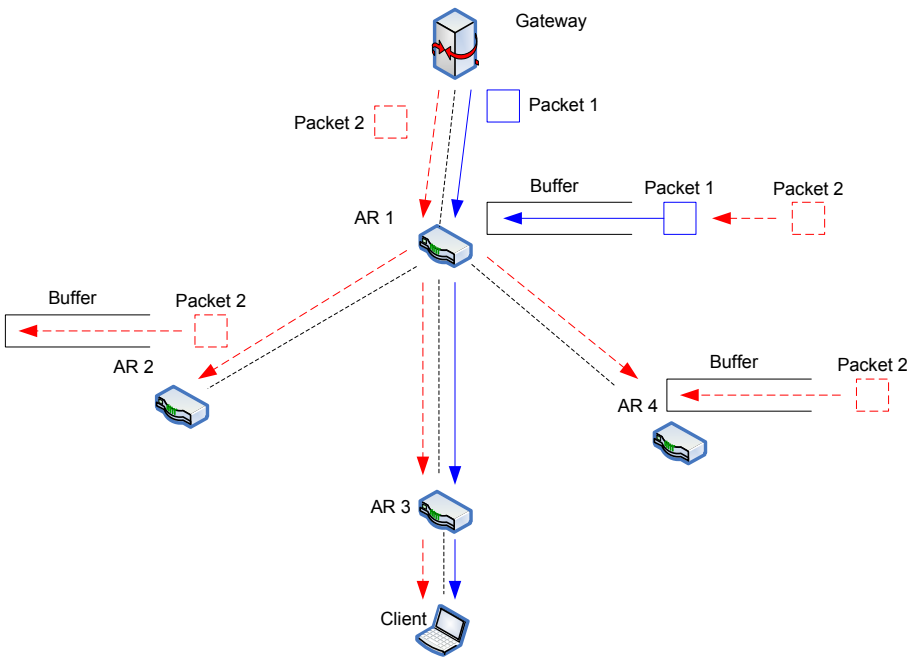


Figure 2.4: En-route caching and promiscuous caching

support seamless handoff in wireless mesh networks. Compared to En-route caching, promiscuous caching has a lower loss rate since it has a higher cache hit rate.

2.2.2 Inter-Domain Handoff Management over WMNs

In comparison with intra-domain handoff management, fewer inter-domain handoff management solutions considering the inherent characteristics of WMNs are proposed.

The original SMesh [8] architecture only supports intra-domain handoff; however, to enhance the basic SMesh, an inter-domain routing protocol is proposed in [9]. The idea is to incorporate a multicast group called Internet Gateway Multicast Group (IGMG), which is composed of Internet gateways to provide transparent inter-domain handoff. TCP and UDP applications are considered separately in the inter-domain handoff.

In TCP communications, when an Internet gateway receives a SYN packet that requests the setup of a new TCP connection, it regards this new TCP connection as belonging to itself, generates a new NAT entry in its NAT table, and forwards the SYN

packet to the destination. When an Internet gateway receives a TCP packet that is not a SYN packet from a mobile client and the gateway does not have any NAT entries regarding the TCP connection in its NAT table, it multicasts the packet to the IGMG. After that, the gateway which is the owner of that connection receives the packet, forwards the packet to the Internet, and sends an owner notification message to the IGMG. All members in the IGMG then know the ownership of that connection. After this notification, if any other gateways receive the TCP packets from that connection, they will forward the packets to the owner directly.

With regard to UDP communications, connection-less and connection-oriented communication are considered respectively. All connection-less UDP packets are sent to the destination directly, such as DNS. For connection-oriented UDP applications, when the gateway receives a new UDP packet from a mesh client, it relays the packet to both its destination and the IGMG, and sets a timer for waiting for a response regarding the ownership of that UDP connection. If another gateway in the IGMG finds that this UDP connection belongs to it, it sends an owner notification message to the IGMG. Then all future UDP packets in the same connection will be forwarded to the original owner. Otherwise, after timeout, if no owner notification has been received by the initial gateway, the gateway indicates that it is the owner of the UDP connection and forwards the subsequent UDP packets to the Internet destination without forwarding them to the IGMG.

To minimize re-routing delay for an inter-domain handoff, two routing protocol extensions are proposed: Pre-handoff discovery (PRD) for AODV [71] and FastSync for OLSR [70]. The basic idea of PRD for AODV is discovering the route in the new wireless mesh network before starting the handoff. This extension could be integrated with Mobile IPv4 Fast Handovers [41] or FMIPv6 [40]. Before handoff, the mesh client decides on the target AR according to a layer-2 scan. The mesh client sends a pre-handoff route discovery request to the new AR. When the new AR receives the request, it assigns the client a new care of address and initiates route discovery. After handoff, the mesh client

can use this route to continue communication. FastSync for OLSR decreases the time for updating routing information after inter-domain handoff. It enables the mesh client to establish the links quickly according to the network topology.

2.3 Handoff Management over Vehicular Mesh Networks

To implement handoff management for maintaining wireless connections during the mobile nodes' movement in vehicular mesh network, some of handoff schemes designed for traditional wireless networks can be immigrated [55, 59, 40]. However, there are still some problems needed to be solved, such as long handoff latency, high handoff frequency, etc, because vehicles have high mobility compared with regular mobile nodes. Therefore, for vehicular mesh networks, special handoff schemes should be implemented. In general, handoff schemes used in vehicular networks can be categorized to two groups: predication-based and cluster-based.

2.3.1 Prediction-based schemes

For vehicular mesh networks, mesh clients have to move according to the existing roads. Therefore, it is possible to predict the location of mesh clients and use the location information to improve the handoff. Kousalya *et al.* [42] proposed a predictive handoff scheme to reduce handoff latency for real-time mobility tracking. They divided the radio coverage of access points into two zones based on the signal quality. The location of mobile terminal can be predicted based on the serving zones. Using the location information and capacity of the access point, the mobile node can reserve computing resources for the following communication. The pre-handoff process is introduced to receive the handoff region notifications and discuss about QoS parameters. Moreover, to provide various QoS parameters, the dynamic channel scanning is adopted to improve

the handoff performance. The problem is that the prediction is not precise enough.

Harri *et al.* [28] proposed a new location-aware framework, called kinetic graphs, to capture the location information of mobile vehicles. Based on the kinetic graphs, the handoff process in vehicular networks can be optimized. The kinetic graph framework includes four functions: representing the trajectories, posting the trajectories, building the kinetic graphs and maintaining neighborhood. To represent the trajectory, they used the current position and the moving velocity of vehicles to predict the next possible position. Then, to post the trajectory information, the data format of the trajectory information is specified. It includes the position coordinates, the speed vector and a time stamp. To build the kinetic graphs, two link weights are used: distance-based weight and nodal degree weight. The distance-based weight uses the Euclidean distance and it is useful for routing protocols; and the nodal degree weight is popular to model mobility for broadcast protocols. Finally, to maintain neighborhood, some heuristics are proposed to detect aperiodic neighborhood changes. Using the kinetic graph framework, we can adapt the graph theories to handoff management solutions in vehicular networks. Thus, this framework can be used to reduce handoff latency by predicting the possible location of vehicles.

Huang *et al.* [29] presented a packet forwarding control scheme to improve the handoff performance by reducing packet transmission time over multi-hop vehicular networks. Using the proposed control scheme, a new relay point called the common ahead point is selected to forward data packets to the new access router with less hops. In original scheme, when the data center sends a packet to the vehicle, the packet is sent to the old access router and the old access router sends the packet to the new access router. Upon receiving the data packet, the new access router then forwards the packet to the vehicle. The tunnel technology used here introduces extra packet forwarding latency. Using the packet forwarding control scheme, the common ahead point intercepts the packets sent from the data center to the old access router and it forwards the packet to the new access router directly. Then, the packet forwarding delay during the handoff is reduced.

However, the main problem of the scheme is that how to select the common ahead point. A bad choice of the common ahead will hazard the packet forwarding performance.

2.3.2 Cluster-based schemes

In vehicular networks, some vehicles would move in a group formation, because the movement of vehicles are constrained by the roads and the speed cap. Therefore, the solution for group mobility can be adopted in vehicular mesh networks to reduce the quantity of handoffs. One of the possible solutions is Network Mobility (NEMO) [22]. Using NEMO, mobile nodes are divided into different groups, and one group is called a mobile network. Moreover, according to different functionality provided in a mobile network, mobile nodes are divided into four categories, mobile routers, local mobile nodes, visiting mobile nodes and local fixed nodes. Mobile routers are the logical center of mobile networks. They work as routers in other nodes' view. All of other mobile nodes have to communicate with the backbone through their registered mobile routers. Local mobile nodes and visiting mobile nodes are two kinds of mobile nodes that will change their current associated mobile routers. In a mobile network, local mobile nodes' home agents are located in the current mobile network; and visiting mobile nodes' home agents are located in other mobile networks. Local fixed nodes are mobile nodes that always belong to their current mobile network. When the mobile router switches its access points during the handoff process, the whole mobile network is associated with the new access point. In consequence, only the mobile router has to complete the handoff for connecting to the new access point. Other mobile nodes do not need to do handoff during their movement; because they are always associated with the current mobile router and no handoff is triggered. Consequently, only one handoff is completed for a mobile network and the total handoff overhead is reduced.

Baldessari *et al.* [10] introduced two approaches to implement NEMO in vehicular networks: the MANET-centric approach and the NEMO-centric approach. Using MANET-centric approach, a new ad hoc routing protocol is introduced to support the

data communication between the vehicle node and the roadside infrastructure. The ad hoc routing protocol is executed by vehicle nodes distributively, and the routing protocol implements address configuration and gateway selection function. The NEMO scheme is then implemented on top of the routing protocol. Therefore, the routing functionality includes two layers: the MANET layer and the NEMO layer. In the MANET layer, the ad hoc routing is used to forward data packets between vehicles. In the NEMO layer, mobility routing is provided to forward packets between mobile networks. Conversely, using the NEMO-centric approach, if a vehicle node wishes to communicate with other roadside units through multi-hops, the routing path includes at least one mobile router of another node. In this solution, the mobile routers construct a logical backbone to forward data packets. According to the analysis, the MANET-centric approach provides better V2V and V2I communications in vehicular networks. Moreover, the MANET-centric approach has easier implementation than the NEMO-centric approach. However, the time spent by one handoff is still too long. The handoff latency should be reduced to the tolerant level.

To reduce handoff latency, a real bus scheme is proposed in [19]. In this scheme, buses are selected as mobile router nodes, and two routers are installed in each bus. One router called the front mobile router is installed in the front of the bus, and the other one is installed in the rear of the bus, which is called the rear mobile router. The rear mobile router is used to connect to the Internet through roadside units. When the bus moves to the new network, it triggers the pre-handoff process and updates its IP address. In the pre-handoff process, the front mobile router helps the rear mobile router to acquire the new IP address. Then, the rear router can receive the new IP address and the information of the access point from the front router. To acquire the new IP address, the front router will communicate with vehicles on the lanes of the opposite direction or vehicles in the same direction to reduce the handoff overhead. The advantage of the real bus scheme is that the rear mobile router can receive its IP address before the handoff process is triggered. However, the real bus scheme needs extra routers to support its

pre-handoff scheme. Moreover, the front router is only used in pre-handoff process and it does not help the data communication. Based on the real bus scheme, Chen *et al.* proposed a virtual bus scheme. Using the virtual bus scheme, two vehicles are grouped as a virtual bus with two mobile routers. The idea used to reduce handoff latency by the virtual bus is the same with that of the real bus scheme. However, the virtual bus scheme introduces extra communication overhead because it has to maintain the virtual structure.

For cluster-based solutions, the performance of the clustering algorithms used will also affect the handoff performance. Therefore, the question of how to design an efficient clustering is important. The lowest ID clustering algorithm [25] is one of the easiest way to cluster mobile nodes for wireless networks. Using this algorithm, all of the wireless nodes broadcast beacon messages in which the node IDs are encapsulated. Moreover, these nodes IDs are assigned uniquely. The node which has the lowest ID in its neighborhood is selected as the cluster head node; and other nodes are selected as the cluster member nodes. The lowest ID algorithm proposed a basic strategy to cluster mobile nodes. First, we need to define a metric to model the property of wireless nodes; and then we can use the predefined metric to group nodes based on some rules. The following clustering schemes are all based on this idea. The difference of various clustering scheme is the metrics used for modeling. The question of selecting the metrics is very important for clustering schemes.

Because the lowest ID clustering algorithm did not take into account the mobility, to cluster mobile nodes, a mobility metric is needed to represent the property of mobile nodes. Basu *et al.* [12] proposed a new mobility metric, which refers to MOBIC, to represent relative mobility between nodes in one-hop distance. Using MOBIC, mobile nodes broadcast beacon messages every broadcast interval. When a mobile node receives two consecutive beacon messages from its neighbor node, it measures the relative mobility between two nodes as the ratio of the received signal strength of the new beacon message and the received signal strength of the old beacon message. The mobile nodes then

calculate the aggregate mobility metric based on the relative mobility. After that, the mobile nodes which have the smallest aggregate mobility value are selected as the cluster head nodes. Therefore, using MOBIC we can construct clusters in which the maximum distance between the cluster head nodes and other cluster members nodes is one hop.

Shea *et al.* [65] presented a clustering scheme using affinity propagation for VANETs. Affinity propagation is first proposed to solve data clustering problem and it is demonstrated that this algorithm can generate clusters more efficiently compared with traditional solutions. In this solution, the idea of affinity propagation is used to cluster vehicle nodes in a distributed way. The vehicle nodes exchange messages with their neighbor nodes to transmit availability and responsibility, and make the decision based on the availability and responsibility values to construct clusters. The simulation results demonstrate that the performance of the clustering scheme using affinity propagation is better than MOBIC in terms of stability.

Density Based Clustering (DBC) algorithm is proposed in [43]. Using DBC, connectivity level, link quality and traffic conditions are taken into account completely to cluster vehicle nodes. The mobile network is divided into dense part and sparse part. A node which has links more than a predefined value is considers as in the dense part; otherwise, it is in the sparse part. During the clustering process, link quality is estimated to make re-clustering decision. According to the experiment results, the cluster head change ratio is less than the lowest ID algorithm [25].

Maslekar *et al.* [46] presented a direction based clustering algorithm for vehicular networks. Authors assume that each vehicle has a digital map and can calculate the moving path based on the source address and the destination address. The vehicle clusters are formed based on the directions which vehicles take at the intersection of the roads. Therefore, the vehicles which have the same direction will form a cluster. However, this scheme only takes into account the direction; as a result, the cluster will change frequently.

The Distributed and Mobility Adaptive Clustering protocol (DMAC) is proposed

in [11]. Using DMAC, each node is assigned a weight parameter; and the weight can be computed based on link quality, mobility and etc. The mobile nodes which have the largest weight are selected as cluster nodes. To adapt DMAC to vehicular environment, a modified DMAC protocol is presented in [77]. Wolny *et al.* added the consideration of moving direction when two groups of vehicle meet. If the groups of vehicle are moving in the different direction, the re-clustering process is avoided. Therefore, the algorithm can increase the stability of vehicle clusters.

Moreover, Kwon *et al.* [44] proposed the Passive Clustering (PC) scheme to reduce the messages exchanged for clustering. Using PC, each node maintains its cluster status based on the traffic flows locally; and no cluster control messages are needed to construct clusters. Therefore, the maintenance cost of clusters is decreased. Based on PC, Wang *et al.* [75] proposed three clustering algorithms which took into account different metrics to construct vehicle clusters. These clustering algorithms are called VPC. Compared with the PC, VPCs can increase performance of packet delivery and reduce the end-to-end delay.

In summary, to generate vehicle clusters, a lot of clustering schemes are proposed considering the inherent characteristic of VMNs. However, all of them can only construct clusters in which the maximum distance between the cluster head node and the cluster member nodes is one hop. Therefore, we want to have some preliminary research work on how to construct multi-hop clusters in vehicular networks.

Chapter 3

MAC Layer Handoff Management over Traditional WMNs

In this chapter, our work on the MAC layer handoff management over traditional wireless mesh networks is presented. It includes two parts: a self-configured MAC layer handoff scheme with active scan and a fast secure authentication scheme.

3.1 Introduction

In recent years, Wireless Mesh Networks (WMNs) have presented wide deployment due to their strengths compared with traditional wired networks, such as low cost, easy to deploy and easy to configure. To deploy wireless mesh networks, there are a lot of standards can be used in the physical layer and the MAC layer to support wireless connections, such as IEEE 802.11 [31], IEEE 802.16e [35] and etc. Compared with other protocols, IEEE 802.11 standard protocol is a promising option to construct wireless mesh networks, according to its broad bandwidth and low equipment cost. Therefore, in this chapter, we focus on the question of how to reduce the MAC layer handoff latency for IEEE 802.11 protocol.

Mesh clients in IEEE 802.11-based WMNs obtain network connections to the Internet

through Access Routers (ARs); and, ARs construct the backbone of the network using wireless communication between each other. Mesh clients are mobile terminals which can move freely in the wireless environment or join different networks, such as laptops, iPhones, mobile sinks and etc [13]. Due to the fact that the transmission range of antennas in the IEEE 802.11 networks is very limited, mesh clients have to associate with the new AR during their movements to maintain their network connections, which is termed as the handoff process. In the MAC layer handoff, the mobile nodes select the new AR according to the signal quality and re-associate with it to establish the physical connection.

For IEEE 802.11-based wireless mesh networks, the MAC layer handoff includes three phases: scan, reauthentication and reassociation [47]. The handoff latency can be calculated using Equation 3.1.

$$T_{handoff} = T_{scan} + T_{authentication} + T_{association} \tag{3.1}$$

During the scan phase, the mesh client can detect available ARs using two methods: passive scan and active scan. Using passive scan, the mesh client collects beacon messages from ARs to determine the signal quality of ARs and selects the new AR. However, the waiting time for beacon messages is hard to be specified, and synchronization among ARs is needed; therefore, passive scan is not suitable for real-time applications. Conversely, using active scan, the mesh client switches to different channels and broadcasts probe messages to detect available ARs. Upon receiving the probe messages, ARs reply with response messages which can be used by the mobile node to determine the new AR. Most improved handoff schemes proposed in literature adopt active scan and in this thesis, active scan is also used in our handoff scheme. The scan phase is responsible for most of the handoff latency in a standard MAC layer handoff process with shared key authentication, and usually the scan time T_{scan} is more than 150 ms [47, 73]. Therefore, the time spent for the scan process has to be minimized.

On the other side, shared key authentication cannot meet the requirement of the secure handoff. Therefore, recent secure authentication schemes use open system authentication. Using open system authentication, after the mobile terminal sends an authentication request to the access point, the access point sends the authentication response message back to the mobile terminal and allows the mobile terminal to access the wireless network by default. Then, all of the following data packets are encrypted by the security key. To generate and distribute secure keys, the Extensible Authentication Protocol (EAP) [2] process is introduced. However, the generation and distribution of security keys between mobile terminals and access points spends more than 1 sec [48]. Thus, besides reducing scan latency, the authentication latency is also needed to be minimized to support smooth roaming.

3.2 A Self-configured MAC Layer Handoff Scheme

The basic idea of the proposed active scan scheme is that the MAC layer scan latency is minimized by reducing both the waiting time in each scanned channel and the number of scanned channels. To shorten the waiting time for the scanned channel, dynamic adaptation is adopted to adjust *MinChannelTime* and *MaxChannelTime* based on the previous scan information and the response messages received in the current scan. To reduce the number of scanned channels, a threshold is predefined and the active scan process can be terminated before all of candidate channels are scanned. The notations used in this section are illustrated in Table 3.1.

3.2.1 Dynamic Adaptation

For an IEEE 802.11 wireless mesh network, *MinChannelTime* and *MaxChannelTime* are two essential components of the handoff process in the MAC layer. *MinChannelTime* is the minimal time spent to wait for an empty channel and *MaxChannelTime* is the maximal time used to wait for an available channel. Therefore, the scan time is controlled

Table 3.1: Notations used in the proposed self-configured scan scheme

Notation	Description
$Channel(i)$	channel i
$Pr(i)$	the probability of channel i is available
$AR(j)$	access router j
$MinChannelTime(i)$	*MinChannelTime* of channel i
$MaxChannelTime(i)$	*MaxChannelTime* of channel i
$Difference(i)$	the difference between *MinChannelTime* and *MaxChannelTime* of channel i
$MinChannelTime_{min}$	the minimal *MinChannelTime*
$MinChannelTime_{max}$	the maximal *MinChannelTime*
$MaxChannelTime_{max}$	the maximal *MaxChannelTime*
$RSS_{highest}(i)$	the highest Received Signal Strength sensed during *MinChannelTime(i)* in $Channel(i)$
$RSS_{current}$	the Received Signal Strength of the current AR
$RSS_{AR(j)}$	the Received Signal Strength of the AR(j)
$AR_{candidate}(i)$	the candidate AR of channel i

by *MinChannelTime* and *MaxChannelTime*. The scan time for each channel can be modeled using Equation 3.2.

$$T_{scan}(i) = T_{probe}(i) + \theta * MinChannelTime + (1 - \theta) * MaxChannelTime \quad (3.2)$$

θ is the parameter to indicate whether the channel scanned is empty or not. If during *MinChannelTime*, there is no frame sensed in the scanned channel, θ is 1 and the scan time for the scanned channel is the sum of the time used to send the probe message and *MinChannelTime*; otherwise, θ is 0 and the scan time is the sum of the time spent to send the probe message and *MaxChannelTime*. Therefore, the scan time can be controlled by modifying *MinChannelTime* and *MaxChannelTime*.

According to [47, 73], the total time spent on active scan is the most significant part of the MAC layer handoff latency with a simple authentication, which is the sum of channel switching time and the scan time for each channel. If we can shorten the total scan time, the handoff performance can be improved. Since the channel switching time

is dependent on the hardware and the time of sending probe messages is pretty small compared with *MinChannelTime* and *MaxChannelTime*, we can reduce the MAC layer handoff latency by controlling *MinChannelTime* and *MaxChannelTime*. The total scan time can be calculated using Equation 3.3.

$$T_{scan-total} = \sum_{i=1}^{N} (T_{switching} + T_{scan}(i)) \tag{3.3}$$

The question of how to set *MinChannelTime* and *MaxChannelTime* properly is very important for supporting smooth roaming. If the values of *MinChannelTime* and *MaxChannelTime* are too large, the handoff latency is too long to provide smooth roaming for real-time applications, and if the values of *MinChannelTime* and *MaxChannelTime* are too small, it is possible that ARs can not receive probe request messages or mobile stations can not receive probe response messages, and the handoff success ratio will be decreased in consequence. Both of the cases lower the quality of service and cause unnecessary handoffs. Unfortunately, IEEE 802.11 standard does not specify the values of *MinChannelTime* and *MaxChannelTime*. Mishra *et al.* [47] suggested that *MinChannelTime* could be set to *6.5 ms* and *MaxChannelTime* could be set to *11 ms*, and Velayos *et al.* [73] recommended that *MinChannelTime* and *MaxChannelTime* could be set to *1 ms* and *10.24 ms* separately. However, using the same fixed *MinChannelTime* and *MaxChannelTime* for all channels is not a good solution, because we can predict the next available channels according to the former scan or receive neighboring ARs' information from the current AR using inter-AR communication. Thus, we can dynamically adapt *MinChannelTime* and *MaxChannelTime* to reduce the handoff latency.

MinChannelTime determines the final scan time of an empty channel and affects the handoff success ratio. For an empty channel in which there is no AR works, *MinChannelTime* should be reduced as much as possible. The optimal case is that *MinChannelTime* of the empty channel is 0; in consequence, the scan time of the empty channel is waived. In the meantime, *MinChannelTime* should be long enough to detect the channel is busy

for an available channel in order to prevent considering an available channel as an empty channel by mistake. Thus, *MinChannelTime* should be adapted based on the prediction information of AR channels. If the neighboring ARs' information can be received from the current AR in some mesh networks, where ARs can exchange their information through inter-AR communication or ARs use the second antenna to receive bacon messages of the neighboring ARs, before the mesh client starts probing channels, the mesh client updates the probability of each channel according to the result of the last scan and the neighboring information received from the current AR. Otherwise, the probability of each channel is changed only based on the result of the last handoff process. Algorithm 1 illustrates the procedure of dynamic adaptation of *MinChannelTime*. $Pr(i)$ is the variable to specify the probability of that channel i is an available channel, which has the value between 0 and 1. α is the unit value of probability to increase or decrease. If the current AR can provide neighboring ARs' information, the probability of the channel where neighboring ARs work is increased with a large factor, since we assume that the information of neighboring ARs provided by the current AR is always correct. Moreover, the probability of the scanned channel is increased if the mesh client found there were some ARs provide wireless connection in the last scan. The probability of the scanned channel is decreased if the channel is empty in the last scan as well. Finally, *MinChannelTime* is adapted between the minimal *MinChannelTime* and the maximal *MinChannelTime*, according to the probability of the scanned channel.

For an available channel where there are some ARs provide wireless connections, the scan time is determined by *MaxChannelTime*, therefore we want to reduce *MaxChannelTime* when the handoff success ratio is kept at a high level. To probe the available channel, after the probe request message is broadcasted, if the mesh client senses that the scanned channel is busy during *MinChannelTime*, it extends the *ProbeTimer* to *MaxChannelTime* for collecting response messages. When the mesh client finds that the scanned channel is busy during *MinChannelTime*, there are two possible cases: the mesh client receives the response messages or the beacon messages from ARs, or the

Algorithm 1 Dynamic adaptation of *MinChannelTime*

```
1: if receives information of neighboring ARs then
2:    for all Channel(i) where a neighboring AR works do
3:       Pr(i) = Pr(i) + 7 * α;
4:    end for
5: end if
6:
7: for all Channel(i) do
8:    if Channel(i) is not empty in the last scan then
9:       if Channel(i) is a non-interfering channel then
10:         Pr(i) = Pr(i) + 2 * α;
11:      else
12:         Pr(i) = Pr(i) + α;
13:      end if
14:   else
15:      Pr(i) = Pr(i) - 2 * α;
16:   end if
17:   if Pr(i) > 1.0 then
18:      Pr(i) = 1.0;
19:   end if
20:   if Pr(i) < 0.0 then
21:      Pr(i) = 0.0;
22:   end if
23:   MinChannelTime(i) = (MinChannelTime_max - MinChannelTime_min) * Pr(i)
      + MinChannelTime_min
24: end for
```

mesh client receives other data frames. For the first case, the mesh client processes response messages or beacon messages intermediately after *MinChannelTime*. In our work, RSS is adopted to present the signal of quality. Therefore, after processing the probe response messages, we get RSSs of available ARs which reply with response messages. All of these ARs are saved in the response list with their RSSs. The highest RSS is used to adapt *MaxChannelTime*. The algorithm of adapting *MaxChannelTime* is shown in Algorithm 2. The difference of *MaxChannelTime* and *MinChannelTime* is changed according to the highest RSS. If the highest RSS is greater than the RSS of the current AR, the difference is reduced; otherwise, the difference is unchanged. The deduction of the difference is calculated according to the ratio of the highest RSS and the current

Algorithm 2 Dynamic adaptation of *MaxChannelTime*

1: $RSS_{highest}(i) = 0.0$;
2: $Difference(i) = MaxChannelTime_{max}$ - $MinChannelTime_{max}$;
3: **for all** $AR(j)$ in the response list **do**
4: **if** $RSS_{AR(j)} > RSS_{highest}$(i) **then**
5: $RSS_{highest}(i) = RSS_{AR(j)}$;
6: $AR_{candidate}(i) = AR(j)$;
7: **end if**
8: **end for**
9: **if** $RSS_{highest}(i) > RSS_{current}$ **then**
10: $Factor = RSS_{highest}(i)$ / $RSS_{current}$;
11: **if** $Factor < \beta$ **then**
12: $Difference(i) = (1 - \log\beta(Factor)) * Difference(i)$;
13: **else**
14: $Difference(i) = 0.0$;
15: **end if**
16: $MaxChannelTime(i) = MinChannelTime(i) + Difference(i)$;
17: **else**
18: $MaxChannelTime(i) = MinChannelTime(i) + Difference(i)$;
19: **end if**

RSS. The higher the ratio, the more deduction we achieve. In a special scenario, if the ratio is greater than β that is the threshold to switch to the next channel, we think the candidate AR is good enough to reassociate with; therefore, the mesh client can stop waiting for other response messages and switch to the next channel to scan for ARs. For the second case, the scanned channel is marked as busy since some data frames are captured or the scanned channel is interfered by neighboring channels, and the mobile client needs enough time to collect response messages and determines candidate ARs; and, default difference value is used to control the value of *MaxChannelTime*.

3.2.2 Handoff Procedure with Dynamic Adaptation

To provide smooth handoff for real-time applications, reducing scan time for each channel only is not enough. Because the handoff latency is controlled by both the scan time of each channel and the number of the scanned channels in the handoff, the number of the

scanned channels is also needed to be reduced. The total handoff time can be calculated using Equation 3.4 approximately.

$$T_{handoff} = \sum_{i=1}^{N}(T_{switching} + T_{scan}(i)) + T_{reauthentication} + T_{reassociation} \tag{3.4}$$

N is number of the scanned channels which is 11 in North American if all of channels are scanned [32], and $T_{switching}$ is the time spent to switch from the current channel to the next channel which is around 5 ms. Therefore, if all of channels are scanned, the total handoff time is more than 50 ms even the scan time for each channel is 0. In consequence, the number of channels scanned should be minimized for real-time applications.

To reduce the number of scanned channels, we allow the mesh clients to terminate the scan phase before all of the channels are scanned. A new threshold of RSS, denoted as RSS_{req}, is defined to specify the RSS which the mesh client needs to establish the wireless communication with the new AR in high quality condition. When the RSS of a certain AR is higher than RSS_{req}, the active scan process is terminated and the mesh client tries to authenticate and associate with the new AR.

The whole procedure of our handoff scheme with dynamic adaptation is shown in Figure 3.1. When a new mesh client joins the wireless network, probability of all channels are initialized to 0.5, since there is no information about ARs in the network and the probability of a channel is empty is the same with the probability of that channel is available. A full scan is triggered to find the AR which has the best signal quality and the mesh client connects to the new AR for the wireless communication. During the movement, the RSS of the current served AR would be decreased and the handoff is triggered by sensing the event that the RSS of the current AR is below the predefined threshold. First, the mesh client sends a scan request message to the current AR. Upon receiving the scan request message, the current AR starts to buffer all of packets received during the handoff, which are forwarded to the mesh client. Then, a scan confirmation message is sent to the mesh client in which the information of neighboring

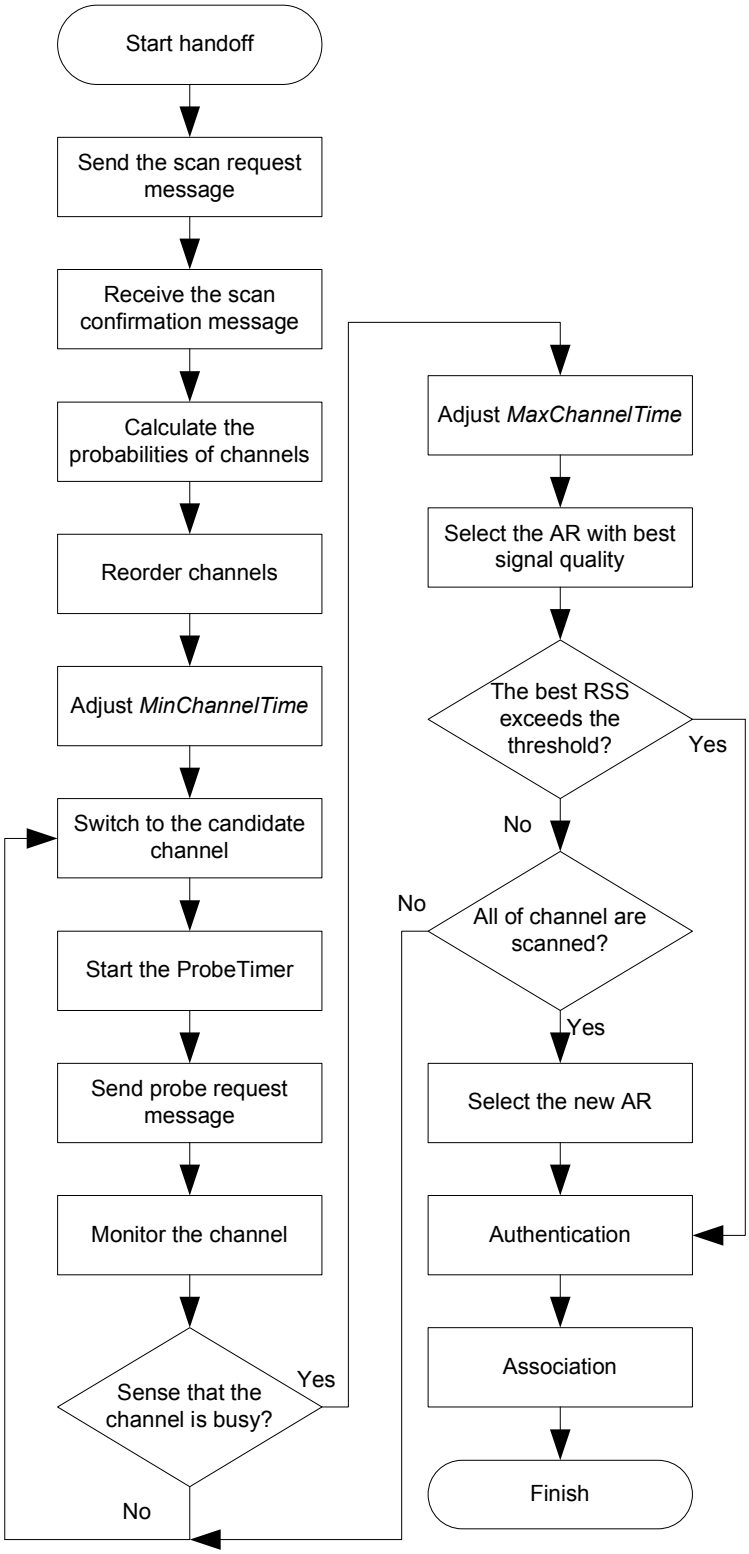


Figure 3.1: Procedure of the MAC layer handoff scheme with dynamic adaptation

ARs is piggyback if the current AR can find neighboring ARs. When the mesh client receives the confirmation message, it enters into the scan phase. The probabilities of all channels are calculated according to the last scan and the information provided by the current AR if it is available, and *MinChannelTime* of each channel is adapted using the probability. In addition, all of channels are ordered according to the probabilities, and the mesh client scans channels in the order from the channel with higher probability to the channel with lower probability. For each channel, the mesh client broadcasts the probe request message and initializes the *ProbeTimer*. After the *ProbeTimer* reaches *MinChannelTime*, if there is no frame captured, the mesh client thinks that the scanned channel is empty and switches to the next channel for probing. Conversely, the mobile node adjusts *MaxChannelTime* according to the frames captured during *MinChannel-Time*. When the *ProbeTimer* reaches *MaxChannelTime*, the mesh client calculates RSSs of ARs that reply with probe response messages. If the RSS of the AR which has the best signal quality is greater than RSS_{req}, the mesh client terminates the scan process and authenticates and associates with the new AR; otherwise, the mesh client continues to scan the next candidate channel. Until all of candidate channels are scanned, the mesh client selects the AR that has the best signal quality as the new AR and completes the authentication and association.

3.3 Fast Authentication

In this section, a fast authentication scheme for handoff in WiFi-based wireless networks is proposed. In general, during the handoff process, the new access router can accept the connection request from the mesh client conditionally, and a tunnel is established between the old access router and the new access router to forward data packets. In the meantime, a complete authentication such as EAP-TLS is processed in the background. Therefore, in parallel with the authentication of the new access router, the mesh client can still communicate with its correspondent node via the old access router. After the EAP-

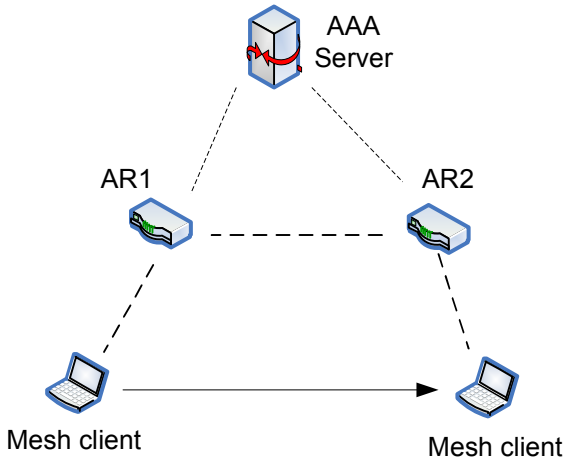


Figure 3.2: A sample scenario to illustrate the secure authentication

TLS authentication is finished, the communication between the mesh client and the new access router recovers to the usual mode. Using this scheme, the complete authentication is implemented without introducing interruption time to the handoff process, and thus both security and performance is achieved.

3.3.1 Detailed Design

In IEEE 802.11i, the equipments involved in the authentication process can be categorized into three groups: suppliants (SAs), authenticators (AAs) and authentication servers (ASs). SAs can be any kind of mesh clients access to the wireless networks, such as iPods, laptops, etc. AAs are physical equipments that provide wireless connections for mesh clients, such as access routers. ASs are some special servers that can implement a complete authentication process, such as AAA servers. To describe our solution simply and generally, we introduce a scenario illustrated in Figure 3.2. This simple scenario consists of one AAA server, two access routers and one mesh client.

When the mesh client joins the network, it should finish the complete authentication process for the first time. After the authentication, it receives the Pairwise Master Key (PMK) and the Pairwise Transient Key (PTK) to communicate with the access

router. In our scenario, the mesh client associates with AR1. As the mesh client keeps moving, it finds out that the signal quality of the current access router is below a predefined threshold, and it triggers the handoff process. The mesh client broadcasts the probe request messages in different channels and collects probe response messages from available access routers in its neighborhood. It then selects the access router which has the best signal quality as the new access router and starts the authentication. AR2 is selected as the new access router in this case.

In our authentication scheme, when the mesh client starts authentication, it sends an authentication notification message to the old access router. In this notification message, the MAC address of the new access router (MAC_{nAR}) and a random number ($RANDOM_{MC}$) are piggybacked. Upon receiving the notification message, the old access router also generates a random number ($RANDOM_{oAR}$) which will be used as a parameter in the hash function. Because the mesh client's MAC address (MAC_{MC}) is already known by the old access router, the old access router can use the mesh client's MAC address, the new access router's MAC address, the random number generated by the mesh client and the random number generated by itself to create the temporary tunnel key (TTK). A hash function is used to generate the temporary tunnel key as shown in Equation 3.5.

$$TTK = H(MAC_{nAR}, MAC_{MC}, RANDOM_{oAR}, RANDOM_{MC}) \tag{3.5}$$

After the temporary tunnel key is generated, the old access router sends an authentication confirmation message to the mesh client, in which the temporary tunnel key is piggybacked. In the meantime, the old access router also sends a tunnel setup message to the new access point. In this tunnel setup message, the mesh client's MAC address and the temporary tunnel key are encapsulated. To make sure the tunnel setup message is not used by some malicious nodes, this message should be encrypted by a shared key which is known by both the old access router and the new access router. We assume that the

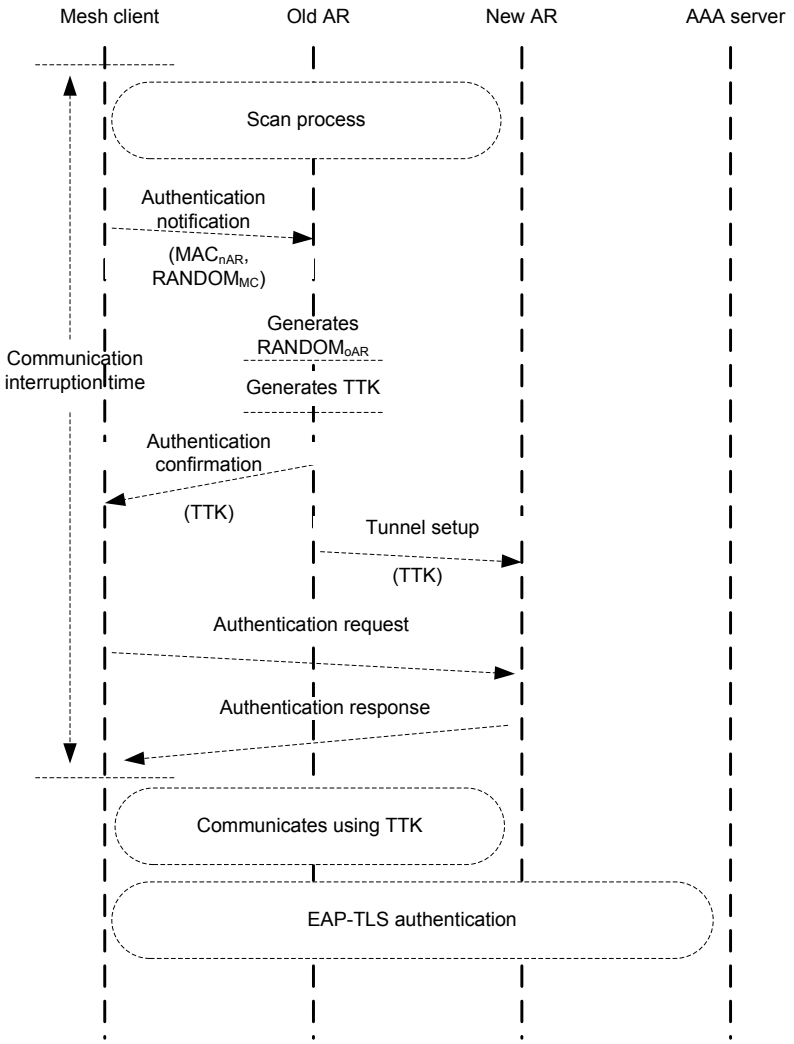


Figure 3.3: Sequence diagram of the proposed authentication scheme

shared key is already distributed when the network is established and the communication between the old access router and the new access router is secure.

Upon receiving the authentication confirmation message from the old access router, the mesh client starts open system authentication with the new access router. The mesh client sends an authentication request message to the new access router. The new access router determines whether to accept the authentication request based on its policy, such as the MAC address filter and etc. If the new access router passes the authentication, the mesh client can associate with the new access router. A complete EAP-TLS process is

then triggered to authenticate both the new access router and the mesh client using the authentication server. The PMK and PTK are generated and distributed during EAP-TLS. In the original IEEE 802.11i scheme, before the new access router and the mobile terminal receive the PTK, the data transmission is blocked; therefore, the communication between the mesh client and its correspondent node is interrupted. Because the complete EAP-TLS takes more than 1 sec, the interruption time is too long to maintain the communication for real-time applications. In our scheme, before the complete EAP-TLS authentication process is finished, the new access point can still receive and transmit data packets for the mobile terminal conditionally. The whole procedure of our authentication scheme is illustrated in Figure 3.3.

Before finishing the complete EAP-TLS authentication scheme, if the mesh client wants to send data packets, the data packets can be sent via the following steps. First, the mesh client encrypts the data packet (PKT) using the PTK shared with the old access point (PTK_{oAR}) as shown in Equation 3.6.

$$PKT_{oPTK} = ENCRYPT(PKT, PTK_{oAR}) \tag{3.6}$$

Then, the encrypted data message is encrypted using the temporary tunnel key again as shown in Equation 3.7.

$$PKT_{TTK} = ENCRYPT(PKT_{oPTK}, TTK) \tag{3.7}$$

After being encrypted twice, the data packet is sent to the new access point.

When the new access router receives the encrypted data packet, it decrypts the data packet using the temporary tunnel key. Then it checks the integrity of the data packet to make sure the data packet is not modified by other unauthorized nodes. If the data packet is modified, the new access router drops the packet; otherwise it sends the decrypted message to the mesh client's old access router. However, the new access router cannot read the contents of the data packet. The pseudocode of the operations executed by the

Algorithm 3 Operation of the new access router while receiving data packets

1: PKT_{TTK}: the data packet received
2: Decrypts the data packet received
3: PKT_{oPTK} = DECRYPT(PKT_{TTK}, TTK)
4: **if** PKT_{oPTK} is modified **then**
5: drops PKT_{oPTK};
6: **else**
7: finds the old access router's address.
8: sends PKT_{oPTK} to the old access router.
9: **end if**

Algorithm 4 Operation of the old access router while receiving data packets

1: PKT_{oPTK}: the data packet received
2: Decrypts the data packet received
3: PKT = DECRYPT(PKT_{oPTK}, PTK_{oAR})
4: **if** PKT is modified during the transmission **then**
5: drops PKT;
6: **else**
7: finds the routing path according to the destination address of the data packet.
8: sends PKT to the correspondent node.
9: **end if**

new access router is shown in Algorithm 3.

When the old access router receives the data packet, it can use the PTK shared with the mesh client to decrypt the data packet and test the integrity of the data packet. If the data packet is not modified by other nodes, the old access router forwards the data packet according to its destination address. The pseudocode of the operations executed by the old access router is shown in Algorithm 4.

3.3.2 Advantages

Compared with the original authentication scheme defined in IEEE 802.11i, our scheme can reduce the communication interruption time to a tolerable level for real-time applications. The communication interruption time is the time interval when the mobile terminal cannot send or receive data packets from its correspondent nodes. Using our

scheme, the communication interruption time is minimized by allowing the mesh client to process EAP-TLS and communicate with its correspondent nodes simultaneously. After the open system authentication and association are completed, the mobile terminal is allowed to send and receive data packets conditionally. Therefore, the communication interruption time is reduced by waiving the time spent for EAP-TLS.

Compared with the predictive authentication and the proactive key distribution, our scheme has lower overhead and the same security level. Using the predictive authentication or the proactive key distribution, a group of candidate access routers are selected and the new PMK should be distributed before the handoff. Therefore, extra communication overhead is introduced to authenticate with all of candidate access routers even only one access router will be the new access router. Using our scheme, there is no extra communication needed for unassociated access routers after the handoff. This is because our authentication scheme is a reactive process and it is executed after the new access router is selected. Moreover, prediction success ratio is another impact to effect the performance of the predictive authentication and the proactive key distribution. If the prediction misses, a complete authentication is still needed to make sure the communication with the access router is secure. The communication interruption time is too large to maintain real-time applications in this situation. Conversely, our scheme is not effected by the prediction success ratio and the handoff would be smooth.

3.3.3 Security Analysis

In general, there are three requirements for secure wireless communications: confidentiality, availability and integrity. All of these requirements are considered in the design of our authentication scheme.

(a) Confidentiality

Confidentiality is preventing malicious nodes from viewing data packets. Using wireless communications, data packets can be caught by malicious nodes easily, because wireless

signals are easy to be recorded. Malicious nodes can intercept and analyze the data packets to get the information in which they are interested, such as important emails, financial information and etc. Therefore, providing confidentiality during wireless communications is a very important issue.

In our authentication scheme, there are two means of data transmission, transmitting data packets via temporary tunnels and transmitting data packets as usual. The confidentiality of both two ways are provided by the encryption. Before the complete EAP-TLS authentication process is finished, the new access router allows the mesh client to transmit data packets conditionally. During these time period, the data packets are encrypted using the temporary tunnel key and the PTK shared with the old access router. Although other malicious nodes can record the wireless signals of data packets, they cannot read the content of data packets because malicious nodes cannot decrypt the data packets without the temporary tunnel key and the old PTK. After the EAP-TLS process is completed, the mobile terminal and the new access router receive the new PTK. For the following communication, they can use the new PTK to encrypt the data packets, instead of using the temporary tunnel key and the old PTK. Similar to the previous period, the confidentiality of information is also provided by using encryption. If malicious nodes cannot get the new PTK, they cannot view the content of data packets.

(b) Availability

Availability is allowing authenticated users to access to data information. Because the data communication after association can be divided to two phases, before EAP-TLS is completed and after EAP-TLS is completed, we also consider the availability issue through two steps. Before the complete EAP-TLS is finished, data packets are encrypted using the temporary tunnel key and the old PTK. The new access router can access to the data packet encrypted with the old PTK, because it holds the temporary tunnel key. The old access router can access to the original data packet which is supposed to be sent to the correspondent node, because the old access router has the old PTK. After

the EAP-TLS authentication is finished, the new access router and the mesh client can access to data information because they have the new PTK.

(c) Integrity

Integrity is preventing malicious nodes from modifying data packets. In our scheme, before the EAP-TLS is completed, both the data information and the checksum are encrypted using the temporary tunnel key. When the new access router receives the data packet, it decrypts the data packet using the temporary tunnel key and tests the checksum. If the checksum matches, the new access router forwards the data packet to the old access router. Otherwise, the data packet will be dropped, and the data communication can be blocked until the EAP-TLS is completed.

(d) Resistance to Attacks

To attack a wireless connection, there are a lot of misbehaviors. Unauthorized analysis is one of the most common security attacks. Because wireless communication is based on wireless signals which are easy to be intercepted, recorded and analyzed, prevention from unauthorized analysis is the basic requirement of the secure wireless network. Using our authentication scheme, although malicious nodes can capture data packets, they cannot view the content of data packets, because data packets are encrypted using the temporal tunnel key and the old PTK before EAP-TLS is completed, and after EAP-TLS is completed, data packets are encrypted using the new PTK.

Malicious access routers are some access router deployed to get clients' data information illegally. It is supposed that the complete EAP-TLS authentication can recognize malicious access points. Before completing EAP-TLS authentication, data packets are encrypted using both the temporary tunnel key and the old PTK shared between the mesh client and the old access router. Even malicious access routers can get the temporary tunnel key via some means, they cannot get the old PTK. Therefore malicious access points cannot view data packets and the data communication is still secure.

Spoofing is to impersonate an authorized access router to receive data information from the mesh client. Similar to the scenario of malicious nodes, spoofing is prevented because fake access routers do not have the old PTK to decrypt the data packets.

3.4 Experimental Results

In this section, the experimental results are presented and the discussion of the performance is given. Both the self-configured MAC layer handoff scheme and the fast authentication scheme are implemented by modifying the Network Simulator - ns2 [53].

3.4.1 The Self-configured MAC Layer Handoff Scheme

To compare with other schemes, extra three handoff schemes are implemented: standard handoff scheme using full scan [31], handoff scheme using selective scan [67] and FHAP [20]. Common simulation parameters are defined in Table 3.2. For our scheme, the minimal *MinChannelTime* is set to 1 ms, the maximal *MinChannelTime* is set to 6 ms, the maximal *MaxChannelTime* is set to 12 ms, *alpha* is set to 0.1 and β is set to 20. For other schemes, *MinChannelTime* is set to 5 ms and *MaxChannelTime* is set to 11 ms.

Table 3.2: Simulation parameters for the self-configured handoff scheme

Parameters	Value
Simulation time	1200 sec
Area	800 m * 600 m
Number of ARs	30
Mobility model	Random waypoint model
Maximal velocity	10 m/s
Pause between movements	1 sec
Channel switch delay	5 ms
Probe delay	0.1 ms
Propagation model	Two-ray ground reflection

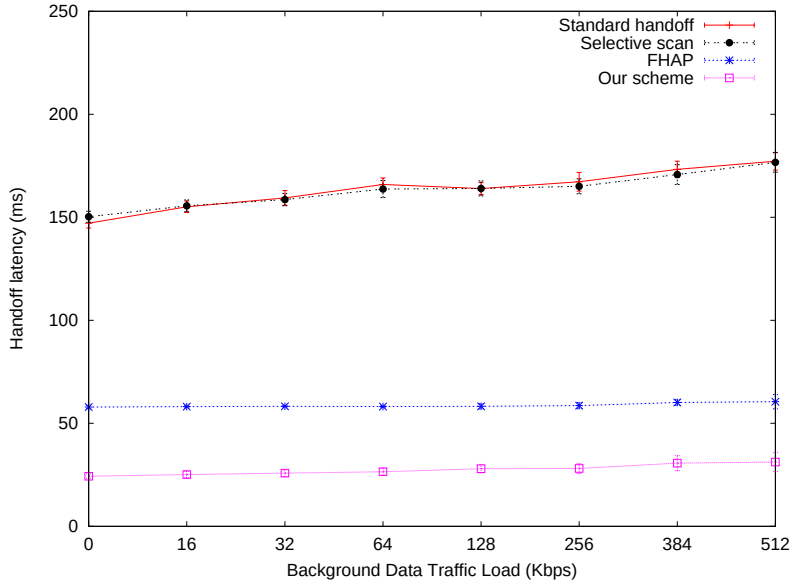


Figure 3.4: The MAC layer handoff latency using 1 AR channel

(a) Handoff latency

Handoff latency is one of the most important metrics to evaluate the performance of handoff schemes. Handoff latency is measured as the time spent from when the mesh client starts to probe ARs to when the mesh client associates with the new AR. To measure the handoff latency completely, three scenarios are constructed to execute the simulation experiments. In these scenarios, 1 AR channel, 3 AR channels and 11 AR channels are used to provide wireless connections separately. Moreover, background traffic is introduced in the experiments to illustrate how the traffic affect the handoff latency, the 95% confidence range is calculated for the handoff latency.

In the first scenario, all of ARs use channel 3 as their AR channels. The handoff latencies of four schemes in the first scenario is shown in Figure 3.4. In this scenario, the standard handoff using full scan and the handoff scheme using selective scan have similar performance. Both of these two schemes spend more than 140 ms to complete the handoff in MAC layer. For the standard handoff, all of 11 channels are needed to

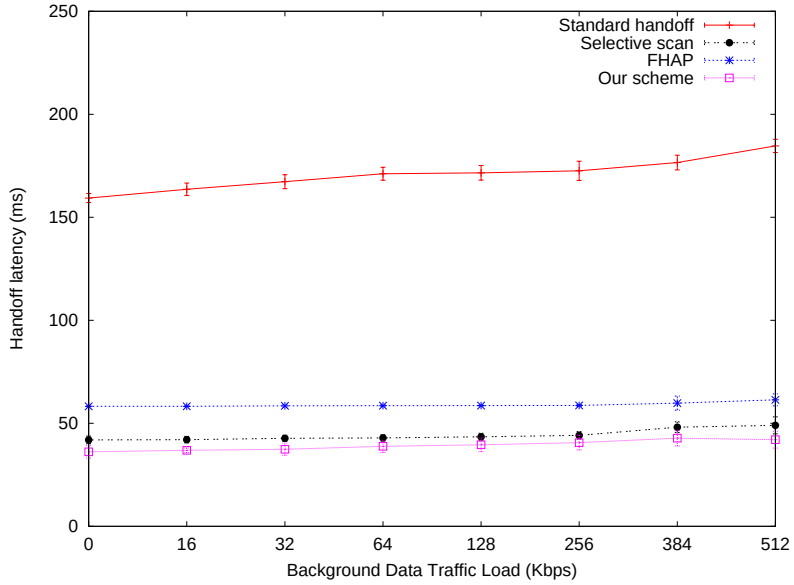


Figure 3.5: The MAC layer handoff latency using 3 AR channels

be scanned, although in this scenario only 1 AR channel is used; therefore, the standard handoff spends a lot of time. For the selective scan, since the mask of the current channel is always set to 0, the predication of available channels is incorrect and the selective scan also has to scan all of channels. FHAP can waive waiting time by forcing the candidate AR sends the response message to the current AR; however, the mobile node still needs to broadcast the probe request message in all of channels, and the handoff latency is more than 50 ms. Our scheme spends the least time in this scenario, which is less than 35 ms; because our scheme can predict the available channel correctly for the first scenario, and *MinChannelTime* and *MaxChannelTime* are set properly.

In the second scenario, channel 1, 6 and 11 are adopted as AR channels. These three channels are non-interfering channels in IEEE 802.11 wireless networks, and most of ARs provide wireless connections using one of the three channels. In our opinion, this scenario is the most common scenario in a real IEEE 802.11 wireless network. ARs are assigned AR channels using one of the three channels randomly. Figure 3.5 illustrates the experiments results of this scenario. Standard handoff spends similar time with that

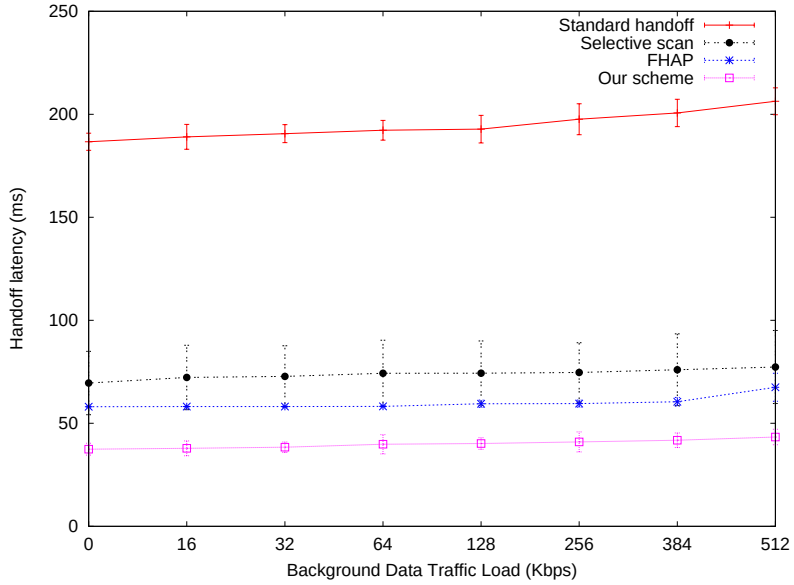


Figure 3.6: The MAC layer handoff latency using 11 AR channels

of the first scenario, because it has to scan all of channels to select the new AR. FHAP also has similar performance with its performance in the first scenario; since FHAP does not optimize the handoff process according to different scenarios. The selective scan shortens handoff latency using the channel mask. For the second scenario, using the channel mask can predict the available channels and waive unnecessary scans for empty channels. Our scheme has the best performance in this scenario again. Compared with the first scenario, the handoff latency of our scheme increases a little bit; however, the latency is still less than 50 ms.

In the third scenario, all of 11 channels can be used as AR channels, and ARs select a channel to support wireless connections randomly. This scenario is the most complex scenario, because it is hard to predict the next available channels. Figure 3.6 shows the results of this scenario. The standard handoff uses more than 170 ms, since the number of non-empty channels increases and the mesh client has to wait for response messages in these non-empty channels. For the same reason, the performance of the selective scan is decreased and it uses more than 60 ms in this scenario. The performance of FHAP

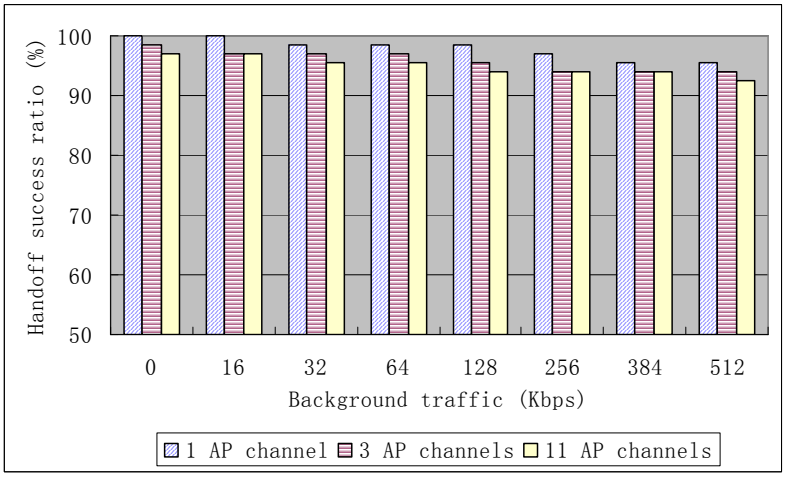


Figure 3.7: The MAC layer handoff success ratio in three scenarios

is not decreased in this scenario, because the waiting time for each channel is waived. Our scheme can reduce the number of scanned channels and the waiting time for each scanned channel in this scenario; in consequence, it still has the best performance in this scenario, and the latency is less than 50 ms.

The handoff success ratios of our scheme in three scenarios are shown in Figure 3.7. The handoff success ratio of our scheme is greater than 90% in all of three scenarios.

(b) Inter-frame delay and frame loss ratio

In addition, we establish a real-time communication between two mobile nodes to test the effect of our handoff scheme for real-time applications. To simulate the G.711 encoded/decoded VoIP stream [37], a CBR traffic is established between two mobile nodes. One mobile node sends a 160-byte UDP [57] packet to another one every 20 ms. APs use 3 AP channels: channel 1, 6 and 11, to support wireless connections for mobile nodes.

Table 3.3: Average frame loss ratio for the self-configured scan scheme

Scenario	Frames sent	Frames lost	Loss ratio
1 AP channel	10000	3.652	0.036%
3 AP channels	10000	6.298	0.063%
11 AP channels	10000	8.409	0.084%

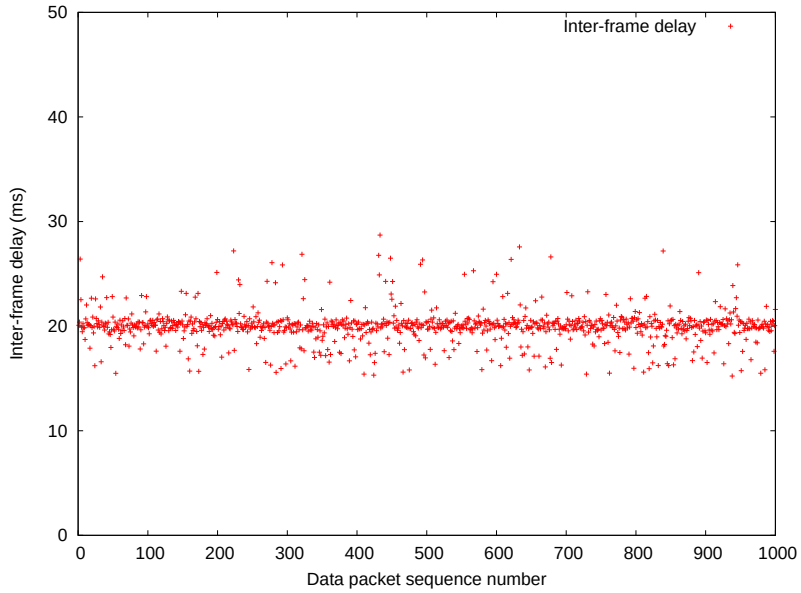


Figure 3.8: Inter-frame delay using the proposed MAC layer handoff scheme

The inter-frame delay of 1000 data packets of the CBR traffic is shown in Figure 3.8. Most frames can arrive at destination in time, and the inter-frame delay is around $20ms$. Some frames arrive late because of the jitter which is introduced by the wireless connection and the interruption caused by the handoff. The average frame loss ratio is illustrated in Table 3.3. The frame loss ratio using our handoff scheme is less than 0.1%. Thus, our handoff scheme meets the requirements of real-time application.

3.4.2 Secure Authentication

This section presents the experiments and simulation results for the proposed fast authentication scheme. To test our authentication scheme, four mobility models are used to generate the mobile client's movement path: (i) Bounded Random Mobility Model (BRMM), (ii) Brownian Motion Mobility Model (BMMM), (iii) Random Direction Mobility Model (RDMM), and (iv) Random Waypoint Mobility Model (RWMM). The simulation time is 1000 *sec* and the maximal speed is set to 10 m/s. Other specific parameters are set

(a) Bounded Random Mobility Model

(b) Brownian Motion Mobility Model

(c) Random Direction Mobility Model

(d) Random Waypoint Mobility Model

Figure 3.9: Authentication latency using different mobility models

as follows: in the case of BRMM, the incremental time interval is 1 *sec*, the maximal angular change in direction is 90 degrees, the maximum acceleration is 1 m/s^2, and in the case of RWMM, the pause between movements is 1 sec.

In our experiments, authentication latency is defined as the time period from the time when the mesh client selects the new access point to the time when the communications between the mobile terminal and its correspondent nodes are resumed. To determine how the background data traffic load affects handoff latency, data traffic is established in the background, using a different data rate. The authentication latencies introduced by the proposed scheme using four mobility models are illustrated in Figure 3.9. The authentication latency spent by our solution grows when the background traffic increases. However, in the worst case, the authentication latency using our scheme is still less than 30 ms, because our authentication scheme decreases communication interruption

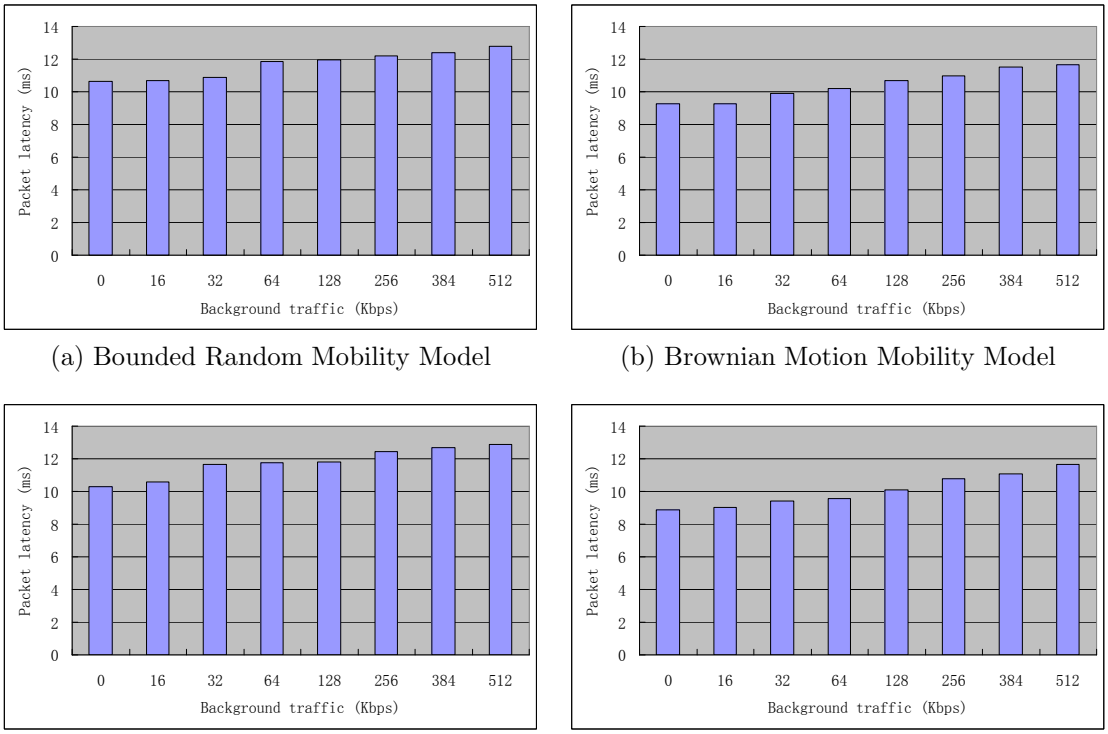
time by introducing temporary tunnels between the old access router and the new access router. Before the complete EAP-TLS process is finished, the mesh client and the new access router can use the temporary tunnel key to encrypt the data packets. Using the temporary tunnel guarantees that the communication between the mesh client and the new access point is secure. Moreover, using tunnel technology can prevent attacks from malicious nodes. After the complete EAP-TLS process is finished, the temporary tunnel will become obsolete, because the mesh client and the new access router receive the new PTK. After that, they can communicate with each other using the new PTK, and the new access point can forward the data packets to the mesh client's correspondent nodes directly.

To test two other important performance parameters, the packet delay and the packet loss ratio, we set up a CBR traffic between two mobile nodes. To simulate a G.711 [1] encoded/decoded VoIP stream [37], one mobile node sends a 160-byte packet every 20 *ms*, by using the UDP protocol [57]. In this experiment, a total of 10000 packets are sent. Table 3.4 shows the packet loss ratio using four mobility models. The average loss ratio is less than 0.12%, which is suitable for real-time applications. The best case occurs with BRMM, where the loss ratio is 0.093%. The worst case happens with RWMM, where the loss ratio is 0.116%. The main reason of packet loss is that the handoff causes the communication interruption.

Figure 3.10 illustrates the packet latencies using four mobility models. Using our scheme, before the complete EAP-TLS is finished, the data packets will be transmitted to their destination via the temporary tunnel between the new access router and the old access router; therefore, the packet latency of the transmission during this time

Table 3.4: Average packet loss ratio and overhead for the authentication

Mobility model	Packets sent	Packets lost	Loss ratio	Overhead (Kbps)
Bounded Random	10000	9.237	0.093%	0.011
Brownian Motion	10000	9.594	0.096%	0.010
Random Direction	10000	10.638	0.106%	0.010
Random Waypoint	10000	11.623	0.116%	0.013



(a) Bounded Random Mobility Model

(b) Brownian Motion Mobility Model

(c) Random Direction Mobility Model

(d) Random Waypoint Mobility Model

Figure 3.10: Packet latency using different mobility models

period would be larger than the packet latency of the transmission after the EAP-TLS is completed. Although more hops would be introduced during this time period, the packet latency is still reasonable, because the time period of conditional connection is very short. Moreover, the packet latency increases when the background traffic grows up. This is because the background traffic increases the probability of the collision during the wireless transmission. As the traffic rate increases, the probability of the collision would also be increased and more retransmission would happen. More retransmission introduces extra packet latency.

The communication overhead introduced by our authentication includes three parts, authentication notification messages, authentication confirmation messages and tunnel setup messages. The authentication notification message is the message sent from the mesh client to the old access router to return the information of the new access point. In

this message, the address of the new access router and a random number generated by the mesh client are encapsulated. The authentication confirmation message is sent to the mesh client for transmitting the temporary tunnel key. Many algorithms can be used to generate the temporary tunnel key for encryption and decryption. In our experiments, an AES key is used as the temporary tunnel key which is 128 bits. The tunnel setup message is sent to the new access router to transmit the temporary tunnel key; therefore it has the same size with the authentication confirmation message. The average overhead of our scheme is around 0.01 $Kbps$. The communication overhead introduced by our scheme is quit low.

In summary, our authentication scheme reduces the communication time during handoff by introducing very low communication overhead. For transmission of data packets, the packet loss ratio is low and the the packet latency is reasonable. Thus, our scheme is ideal for smooth secure handoff in real-time scenarios.

3.5 Summary

For IEEE 802.11 wireless networks, long MAC layer handoff latency degrades the quality of time-critical applications, such as VoIP. The question of how to provide fast MAC layer handoff and smooth roaming for real-time applications is one of important topics for constructing a wireless network. In this chapter, a MAC layer handoff with dynamic adaptation is presented. *MinChannelTime* is adapted according to the neighboring information received from the served AP or the result of the last scan, and *MaxChannelTime* is configured according to the scan result during *MinChannelTime*. Moreover, the scan can be terminated when a certain AP meets the requirement defined by the mobile node. Thus, the scan latency is minimized by reducing the waiting time of each channel and the number of scanned channels to meet the requirements of real-time applications. In addition, to provide fast and secure MAC layer handoff, a novel MAC layer authentication scheme that uses tunnel technology to minimize interruption time during handoff

process is presented. Instead of completing a whole authentication process defined in IEEE 802.11i, we allow the access routers to accept connections conditionally. The temporary tunnel keys are used to encrypt communication between the mesh clients and their new access routers, and all packets are forwarded to the old access router before the complete authentication is finished. As a result, the mesh client can communicate with its correspondent node via the old access router and implement the authentication simultaneously. Simulation results show that the authentication latency can be minimized, which is required by real-time applications in a wireless network.

Chapter 4

Network Layer Handoff Management over Traditional WMNs

In this chapter, our work on the network layer handoff management is proposed. A hybrid routing protocol is first introduced. Based on this hybrid routing protocol, both the intra-domain and inter-domain handoff management schemes are presented.

4.1 Introduction

The performance of WMNs is affected significantly by how the network manages the movements of mesh clients. Therefore, handoff management is one of the most important problems of WMNs. Although many existing network layer handoff management solutions for conventional wireless networks can be applied to mesh networks, new handoff management solutions should be designed and implemented specifically for WMNs, considering their differences.

There are two kinds of roaming: *inter-domain roaming*, which refers to movement across different domains, and *intra-domain roaming*, which means movement among dif-

ferent access routers in the same domain. Accordingly, the network layer handoff management requires both inter-domain and intra-domain handoff management. Inter-domain handoff means the access routers involved in the handoff belong to different domains. Conversely, intra-domain handoff means the access routers involved in the handoff belong to the same domain.

This chapter proposes an innovative scheme to provide both intra-domain and inter-domain handoff management within WMNs. The proposed solution uses a hybrid routing protocol, which integrates the network layer routing and link layer routing to forward packets and achieves easier handoff. For intra-domain handoff, our scheme avoids location updating in the centralized location server, while also decreasing the time for re-routing after the handoff. In addition, our scheme can provide inter-domain handoff with low overhead by minimizing redundant tunnels. It provides smooth roaming with high scalability for real-time applications such as VoIP.

4.2 Hybrid Routing

In recent years, some link layer routing protocols have been designed to improve routing performance in WMNs [26, 33]. In [26], a WDS-based link layer (layer 2) routing protocol for wireless mesh networks is proposed. This protocol introduces a layer 2 routing table which consists of the MAC address pairs that describe the routing information in the link layer. However, since the layer 2 routing tables in mesh routers and the centralized controller involved in this solution grow in proportion to the number of mesh clients, scalability is not well supported.

4.2.1 Detailed Design

A hybrid routing protocol involves both link layer (layer 2 in the OSI model [81]) routing and network layer (layer 3) routing. In essence, the mesh clients reply to an ARP message with their access router's MAC address. Therefore, packets forwarded among

mesh routers can use link layer routing to decrease the encapsulation and decapsulation delay of IP datagrams, and the packets communicated between access routers and mesh clients can use network layer routing. This method avoids centralized control, so that the size of the layer 2 routing table is related to the number of mesh routers, and not the mesh clients. Thus, the layer 2 routing table will be small; the number of mesh routers in a domain is usually less than one hundred.

Table 4.1: The MAC address based routing table

Destination MAC address	Next hop MAC address
00 15 58 83 DF 86	00 15 58 83 DF 88
... ...	... ...

An example of a routing table is shown in Table 4.1. In this example, if a mesh router wants to forward a frame to the destination with MAC address 00 15 58 83 DF 86, it searches the layer 2 routing table and finds that the MAC address of the next hop is 00 15 58 83 DF 88. The frame is then forwarded to 00 15 58 83 DF 88. The layer 2 routing table can be maintained with either a proactive routing method or a reactive routing method, such as the Hybrid Wireless Mesh Protocol (HWMP), which is specified in 802.11s [33]. Every mesh router knows the routing path to any other router.

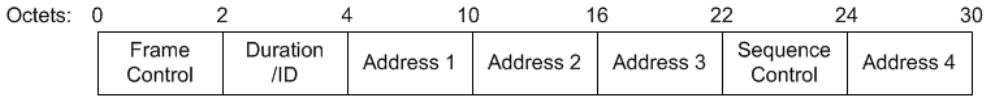


Figure 4.1: IEEE 802.11 MAC frame header

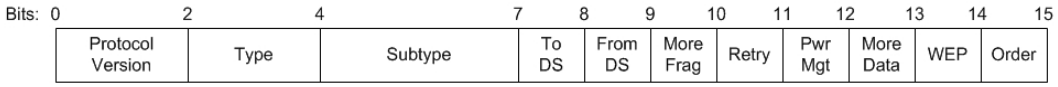


Figure 4.2: Frame control field

The format of the 802.11 MAC frame header is shown in Figure 4.1 [31]. The detailed information of the 11 subfields in the frame control field is given in Figure 4.2 [31]. There are four address fields in the frame header. The different combinations of ToDS, FromDS

Table 4.2: To/From DS fields and address fields

To DS	From DS	Address 1	Address 2	Address 3	Address 4
0	0	Destination	Source	BSSID	N/A
0	1	Destination	BSSID	Source	N/A
1	0	BSSID	Source	Destination	N/A
1	1	Receiver	Transmitter	Destination	Source

and address fields are shown in Table 4.2 [26, 31]. In the various combinations, the source address or destination address can be located in different address fields. In our mesh network, data frames are set by ToDS = 1 and FromDS = 1. The address 1 is the receiver address, which indicates the MAC address of the next hop. Address 2 is the transmitter address, which is the MAC address of the current node. The destination address and source address are in the fields of address 3 and address 4, respectively. This scheme enables the implementation of layer 2 routing in 802.11-based mesh networks.

When a mesh client joins the mesh network, it selects the AR that has the best link quality of all candidates. Many parameters can be used to measure the quality of links, such as signal-to-noise ratio, response delay, and so on. We adopt response delay as the link quality metric in our scheme. The mesh client then sends an association request message to the selected AR. Upon receiving the request, the AR obtains a private IP address from the DHCP server, and returns an association confirmation message with the client's IP address, the AR's MAC addresses, and the gateway's MAC address. Each AR maintains a table containing the IP address and MAC address pairs of all clients registered in it.

Two instances of communication are considered: communication between mesh clients, and communication between a mesh client and the Internet. In the first case, when a mesh client wants to transmit a packet to another client, and already knows the destination's MAC address, it constructs the packet and forwards it to its AR. Otherwise, the mesh client uses an ARP message to obtain the destination's MAC address (which is the MAC address of the destination's AR). The sender's AR then forwards the packet to the

Algorithm 5 Pseudocode of layer 2 routing algorithm

```
1: Extract destination MAC address M
2: if M matches current router's MAC address then
3:    decapsulate the frame
4:    transmit the datagram to the network layer and use layer 3 routing
5: else if the table contains a specified route for M then
6:    forward the frame to the next-hop according to the layer 2 routing table
7: else
8:    report the layer 2 routing error
9: end if
```

Algorithm 6 Pseudocode of layer 3 routing algorithm

```
1: Extract destination IP address D
2: if D matches any mesh client's IP address, which registers in this mesh router then
3:    map D to its MAC address
4:    encapsulate the MAC frame header for the datagram
5:    forward the datagram to its destination
6: else if the table contains a specified route for D then
7:    forward the datagram to the next-hop according to the layer 3 routing table
8: else
9:    report the layer 3 routing error
10: end if
```

receiver's AR, based on the layer 2 routing path. When the receiver's AR receives the packet, it forwards this packet to the receiver, using the destination's IP address. In the second case, in order to send a packet to the Internet, the mesh client must forward the packet directly to the gateway, which transmits the packet to its destination. To send a packet to the mesh client, the gateway forwards the packet according to the layer 2 routing table, provided it knows the MAC address of the mesh client (the MAC address of the client's AR). Otherwise, it sends an ARP message to obtain the MAC address. The client's MAC address is maintained in the table temporarily. After a period of no communication between the mesh client and the Internet, the temporary entry is deleted. The pseudocode of the layer 2 routing is shown in Algorithm 5, and the pseudocode of the layer 3 routing is shown in Algorithm 6.

4.2.2 Sample Scenario

An example of our hybrid routing is described in this subsection. Table 4.3 shows the notation used in this example. The topology of the mesh network is shown in Figure 4.3. One gateway and five access routers construct the relay backbone in our scenario. When client A joins the mesh network, this client selects AR3 as its access router. All packets sent from client A should be sent to AR3; AR3 forwards these packets to the destination. Client A initially registers with AR3 and then obtains an IP address, denoted as IP_A. It also receives the MAC addresses of AR3 and the gateway, which are denoted as MAC_{AR3} and MAC_{GW}.

The following steps are taken when client A desires to send packets to client B. Suppose that A does not know B's MAC address, A sends an ARP message in order to determine it. AR2 then receives the ARP message and replies with its MAC address, MAC_{AR2}. Upon receiving this reply message, client A adds a mapping entry with IP_B and MAC_{AR2} into its mapping table. Using this approach, client A generates frames which are sent to client B, setting ToDS = 1, FromDS = 1, Address 1 = MAC_{AR3}, Address 2 = MAC_A, Address 3 = MAC_{AR2}, Address 4 = MAC_A, and forwards the frames to AR3. AR3 receives the packets, and finds that the destination is AR2. Then, AR3 searches the layer 2 routing table, which is shown in Table 4.4, to find the next hop. In this case, the next hop is AR4. AR2 therefore sets Address 1 = MAC_{AR4} and Address 2 = MAC_{AR3}. Because the frames come from AR3's clients, AR3 sets Address 4 = MAC_{AR3}. In the next step, the frames are forwarded to AR4. Similarly, AR4 forwards

Table 4.3: Notation used to illustrate the hybrid routing

Notation	Description
ARi	access router i
A, B	mesh client A, mesh client B
CN	correspondent node
IP_n	the IP address of node n
MAC_n	the MAC address of node n
GW	gateway

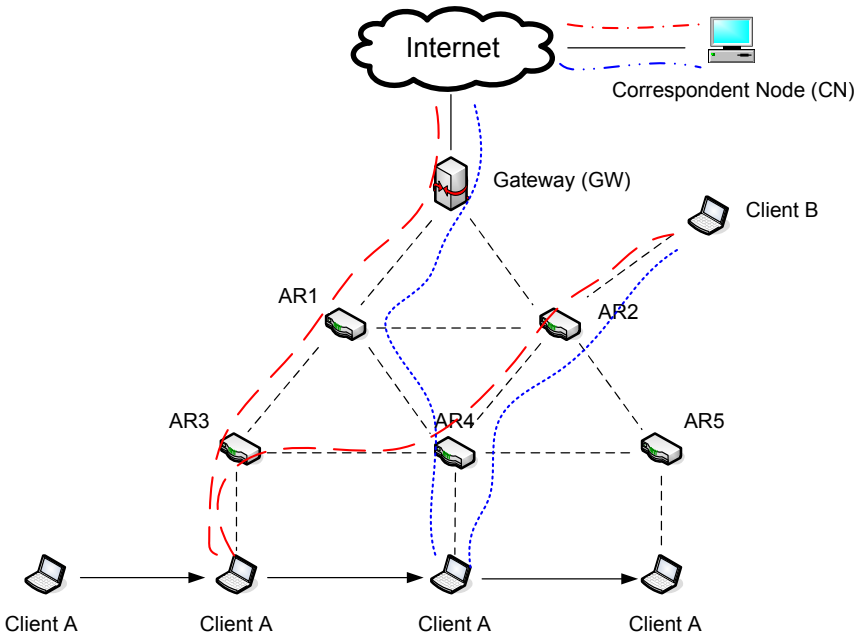


Figure 4.3: A samples scenario during intra-domain roaming

Table 4.4: Layer-2 routing table of AR3

Destination MAC address	Next hop MAC address
MAC_{GW}	MAC_{AR1}
MAC_{AR1}	MAC_{AR1}
MAC_{AR2}	MAC_{AR4}
MAC_{AR4}	MAC_{AR4}
MAC_{AR5}	MAC_{AR4}

the frames to AR2. After this, AR2 finds that the MAC address of the destination is the same as its own MAC address. Therefore, the frames are decapsulated and transmitted to the network layer. In the network layer, AR2 finds that the IP address of the destination is IP_B, determines that the MAC address of client B is MAC_B, sets the frame header as ToDS = 1, FromDS = 1, Address 1 = MAC_B, Address 2 = MAC_{AR2}, Address 3 = MAC_B, Address 4 = MAC_{AR2}, and transmits the packets to client B directly. In this way, client B can also send packets to client A. To do so, client B sets Address 3 = MAC_{AR3} in the frame header.

4.3 Intra-domain Handoff Management

When a mesh client roams in the mesh network, the AR that has the best link quality changes. Therefore, the mesh client has to associate with a new AR. The handoff is triggered by sensing that the response delay of the original AR is below a threshold. The mesh client then selects an AR that has the best link quality as its new AR. This probing process is done in all channels.

Next, the mesh client sends an association request message to the new AR, in which the client's IP address and MAC address pair are encapsulated. When the new AR receives this request, it adds the client's IP address and MAC address pair to its ARP table, and sends back an association confirmation message to the client. The new address pair is used to transmit packets to the client after the handoff. Upon receiving the confirmation message, the mesh client must reply to all of the ARP messages with the new AR's MAC address.

During the handoff, if communication exists between the mesh client and another client in the same domain, the mesh client sends a gratuitous ARP (GARP) message, to the correspondent node, with the new AR's MAC address. Upon receiving this gratuitous ARP message, the correspondent node maps the new AR's MAC address to the mesh client's IP address. All following packets are forwarded directly to the new MAC address without re-routing. The client then sends a disassociation message to the former AR with the new AR's MAC address. The former AR can therefore tunnel the following packets that are sent to the client to the new AR. This temporal tunnel is deleted when the original AR no longer receives any packets sent to that client. This entire handoff process is shown in Figure 4.4.

Conversely, if a mesh client is communicating with an Internet destination, when the mesh client receives the association confirmation message, the gratuitous ARP message is sent to the gateway and not the correspondent node. Then the gateway maps the new AR's MAC address to the client's IP address, and the packets from the Internet can be

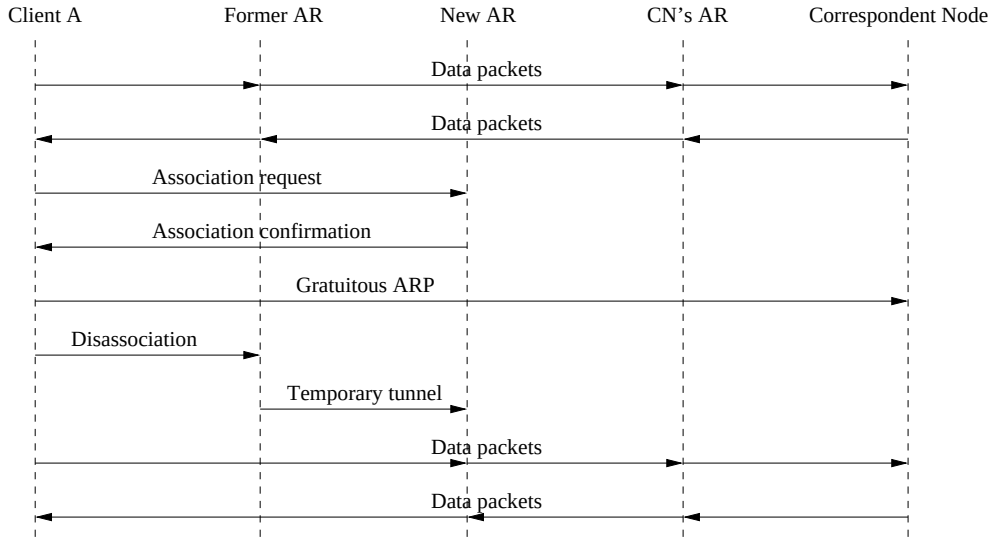


Figure 4.4: Intra-domain handoff if there is a connection in the same domain

correctly forwarded to the client after handoff. However, if a new correspondent node wishes to establish a connection with the mesh client right after handoff but before the gateway is informed of these changes, the connection may not be established. To address this problem, uplink traffic and downlink traffic are considered separately. In case of uplink traffic, the client may accurately send packets to the correspondent node, because the correspondent node's address remains unchanged. For downlink traffic, if the correspondent node sends packets to the mesh client before the gateway has been informed of the changes, all packets received by the gateway will be forwarded to the old AR. And yet, because the prior AR receives the disassociation message which encapsulates the mesh client's new MAC address, the former AR will forward the packets to the new AR. In this way, the connection can still be established even if a new connection request is received before the gateway is informed of changes introduced by handoff. The handoff process in this situation is provided in Figure 4.5. In our solution, the IP address of the mesh client remains the same after handoff. Both UDP and TCP use a combination of IP address and port number to determine the destination. Therefore, the unchanged IP address permits both UDP and TCP communication to remain after handoff.

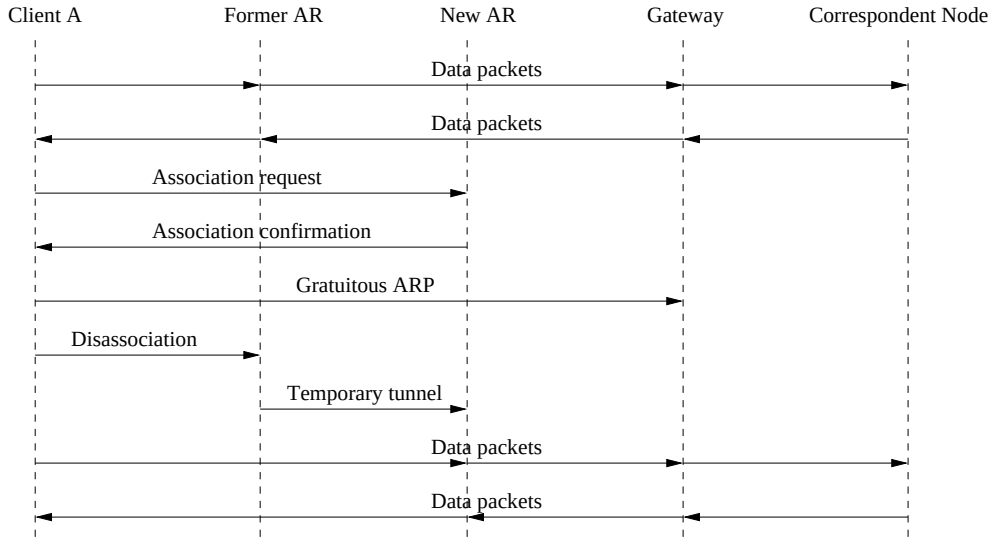


Figure 4.5: Intra-domain handoff if there is a connection through Internet

4.3.1 Sample Scenario

We now consider the same topology of the mesh network, which is shown in Figure 4.3. Client A is a mobile node, and it moves in the network following the arrows shown in Figure 4.3. The dashed lines show the routing path before handoff, while the dotted lines show the routing path after handoff. As client A moves, it senses the signal strength of AR3 weakening. When the signal strength drops below a threshold, client A associates with a new mesh router. In this example, client A chooses to associate with AR4. Initially, client A sends an association request message with its IP address, IP_A, and MAC address, MAC_A, to AR4. AR4 replies with an association confirmation message containing the IP address and MAC address of AR4, IP_{AR4} and MAC_{AR4}. Client A then sends a gratuitous ARP message to client B with MAC_{AR4}, and client B maps the IP_A to MAC_{AR4}. All of the frames sent to client A are set by Address 3 = MAC_{AR4}. Finally, client A sends a disassociation message to AR3 for disconnection. After finishing these operations, the handoff process is finished; and the IP address of client A remains the same after handoff.

The main difference between communicating with a correspondent node (CN) through

the Internet and communicating with a mesh client in the same domain is as follows. When the frame is generated, the field of Address 3 in the frame header is set to MAC_{GW}. Instead of sending a gratuitous message to the correspondent mesh client, when the handoff occurs, mesh client A sends the gratuitous ARP message to the gateway and claims that its new MAC address is MAC_{AR4}.

4.3.2 Advantages

Compared with other solutions, our solution has the following advantages:

- This solution does not require the central location database, and so does not incur the cost of updating and querying the location database is waived. All mesh clients use their ARs' MAC addresses as their MAC addresses; the locations of mesh clients can therefore be obtained according to clients' MAC addresses;
- The use of layer 2 routing can minimize the cost of packet relay among the mesh routers when compared with layer 3 routing. The delay of encapsulation and decapsulation of IP packets in the network layer is avoided. In addition, the IP address of the mesh client remains the same throughout roaming;
- Unlike traditional routing-based solutions, no routing table updating is necessary after handoff. The packets can be transmitted to the mesh client correctly by changing the MAC address of the mesh client;
- Unlike tunnel-based solutions, the tunneling overhead at each level of the hierarchy is removed.

4.4 Inter-domain Handoff Management

The inter-domain handoff is triggered when the mesh client switches to a AR in a different domain. As seen with intra-domain handoff, the mesh client broadcasts a probe message,

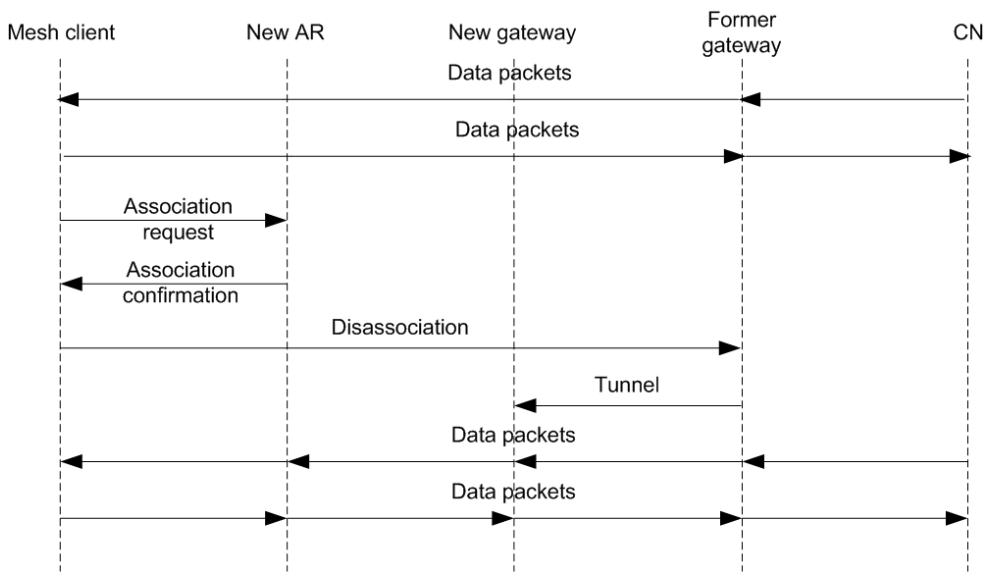


Figure 4.6: The general inter-domain handoff process

gathers response messages, and selects the AR with the best signal strength as the new AR. An association request message is then sent to the new AR. Upon receiving the request, the new AR replies with an association confirmation message that contains the new AR's MAC address and the new gateway's IP address. The main difference between intra-domain roaming and inter-domain roaming is that the IP address of the mesh client changes, since the server in the new domain assigns a new private address to the mesh client. Therefore, in order to maintain the existing data traffic after moving into a new domain, a temporary tunnel is established between the prior gateway and the new gateway through which data packets are forwarded.

After the new AR is selected, the mesh client sends a disassociation message, which includes the new gateway's information, to the prior gateway. The gateway maintains a tunnel list, which records temporary tunnels used to forward packets. For each temporary tunnel entry in the tunnel list the mesh client's address, paired with the new gateway's address, is used to represent a tunnel. When the old gateway receives the disassociation message, it registers the new tunnel in its tunnel list. Therefore, a temporary tunnel is established. All packets sent to the mesh client are then forwarded to the new gateway.

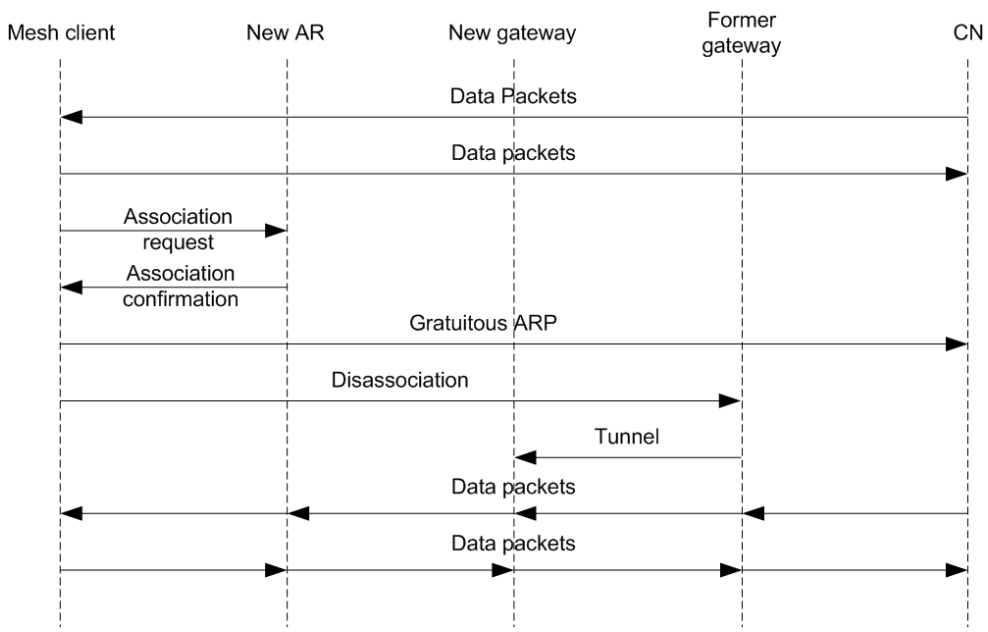


Figure 4.7: Inter-domain handoff if there is a communication in the prior domain

Additionally, to maintain the tunnel list efficiently, a tunnel entry that is not used for a certain period will be removed. This inter-domain handoff process is illustrated in Figure 4.6.

During inter-domain handoff, if the mesh client is communicating with the correspondent node of a former domain, further operations are required to guarantee that the correspondent node is able to find the new data path. In this case, a gratuitous ARP message is sent to the correspondent node by the mesh client before it moves to the new domain. The old gateway's MAC address is encapsulated in the gratuitous ARP message. When the correspondent node receives the gratuitous ARP message, it updates its ARP table and maps the mesh client's address to the gateway's MAC address. After that, according to our hybrid routing protocol, if the correspondent node wants to send packets to the mesh client, it will send packets to the gateway first; the gateway then forwards packets to the mesh client through the new gateway following the temporary tunnel. Figure 4.7 shows this special inter-domain handoff process.

Moreover, when the mesh client moves across different domains, some redundant tun-

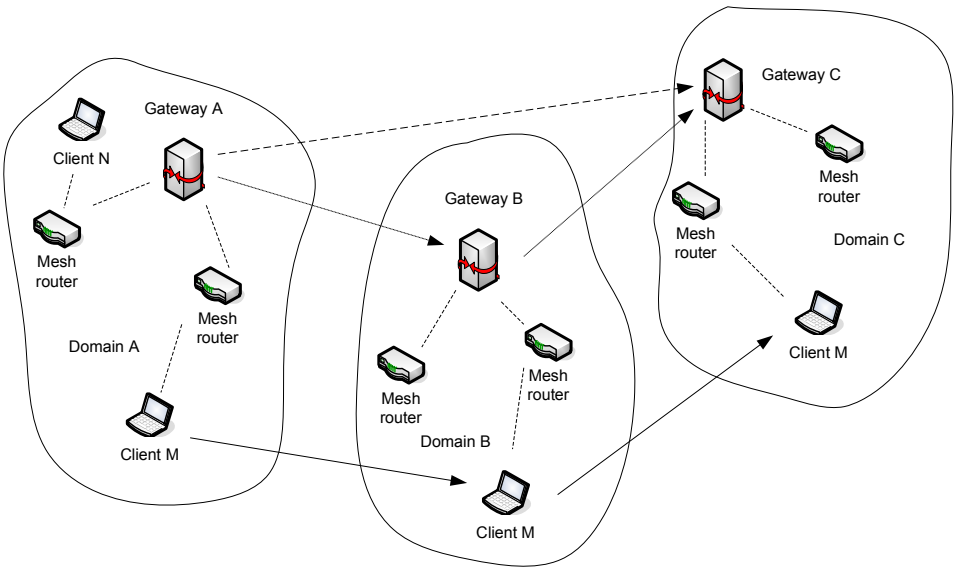


Figure 4.8: An example of inter-domain handoff

nels exist. To save network bandwidth and to reduce forwarding latency, these redundant tunnels should be removed. If a gateway receives data packets sent to a mesh client which moved to a new domain, it sends a notification message encapsulating the new gateway's IP address. The old gateway then updates its tunnel list and maps the mesh client's address to the new gateway's address. As a result, the data packets can directly be sent from the old gateway to the new gateway and redundant tunnels are merged.

An example is shown in Figure 4.8. Suppose that, mesh client M switches from domain A to domain B. It sends a disassociation message to gateway A (GWA) in which the information of gateway B (GWB) is encapsulated. A tunnel is established between GWA and GWB. Because M has a communication with correspondent node N, a gratuitous ARP message containing GWA's MAC address is sent to N. Upon receiving the gratuitous ARP message, N first forwards data packets to GWA; then GWA forwards packets to M through GWB. A new tunnel between gateway B and gateway C (GWC) is established to forward packets after M moves into domain C. The new data path is shown by dotted arrows in Figure 4.8. GWA first forwards packets to GWB, and GWB then

forwards packets to GWC. Though it appears two tunnels are needed to send packets from GWA to M, only one tunnel is truly needed. To reduce the number of tunnels, after M moves into domain C, GWB sends a notification message to inform GWA that the following packets sent to M can be forwarded to GWC directly. In Figure 4.8, the new tunnel is represented by the dashed arrow. The old path $GWA \rightarrow GWB \rightarrow GWC$ is replaced by $GWA \rightarrow GWC$. As a result, the packet latency is reduced.

4.5 Experimental Results

The Network Simulator - ns2 [53] is used to simulate the proposed solution. Intra-domain roaming and inter-domain roaming are evaluated separately. To simulate a G.711 [1] encoded/decoded VoIP stream, a CBR flow, which sends a 160 byte UDP packet every 20 ms at a rate of 64 $Kbps$, is established between the mobile client and the correspondent node in our experiments. Mesh clients move according to four mobility models: (i) Bounded Random Mobility Model (BRMM), (ii) Brownian Motion Mobility Model (BMMM), (iii) Random Direction Mobility Model (RDMM), and (iv) Random Waypoint Mobility Model (RWMM). The general parameters are shown in Table 4.5. Other specific parameters are set as follows: in the case of BRMM, the incremental time interval is 1 sec, the maximal angular change in direction is 90 degrees, the maximum acceleration is 1 m/s^2, and in the case of RWMM, the pause between movements is 1 sec.

Table 4.5: Simulation parameters for the network layer handoff scheme

Parameters	Value
Simulation time	400 sec
CBR packet size	160 $bytes$
CBR packet interval	20 ms
Probe message interval	2 sec
Radio range	250 m
Maximal velocity	5 m/s

4.5.1 Intra-domain Roaming

The performance of our management scheme during intra-domain roaming is discussed in this subsection. The mesh relay backbone is composed of twenty fixed mesh routers, and the mesh routers are deployed randomly in each simulation. However, we guarantee that all mesh routers are connected. There are thirty mobile nodes in our scenarios, and we establish one CBR flow between two mobile nodes. Moreover, the 95 % confidence intervals are demonstrated in our results.

(a) Packet latency and loss ratio

In these experiments, the correspondent node sends the CBR packets every 20 ms to the mobile mesh client; therefore, 20000 UDP packets can be sent in 400 sec. The mobile client roams under a different mobility model and changes its access router. These experiments are used to test how our solution affects the data traffic. The average packet loss ratio and latency is shown in Table 4.6. According to the experimental results, although the mesh client changes its access router frequently, the loss ratio remains at a low level. The best results occur with the Random Direction Mobility Model, where the loss ratio is 0.085%. The worst results occur with the Brownian Motion Mobility Model with a loss ratio of 0.102%. As the table demonstrates in the simulations, the mesh client successfully changes its access router during the roaming process with a low loss ratio. Thus, our solution is proven to realize seamless handoff. The average latency during intra-domain roaming is around 10 ms, and the average latency per hop

Table 4.6: Average packets-loss ratio and latency during intra-domain roaming

Mobility model	Packets sent	Packets lost	Loss ratio	Latency (ms)	Latency per hop (ms)
Bounded Random	20000	18.636	0.093%	9.947	2.458
Brownian Motion	20000	20.394	0.102%	9.936	2.423
Random Direction	20000	16.970	0.085%	9.705	2.464
Random Waypoint	20000	19.788	0.099%	10.633	2.530

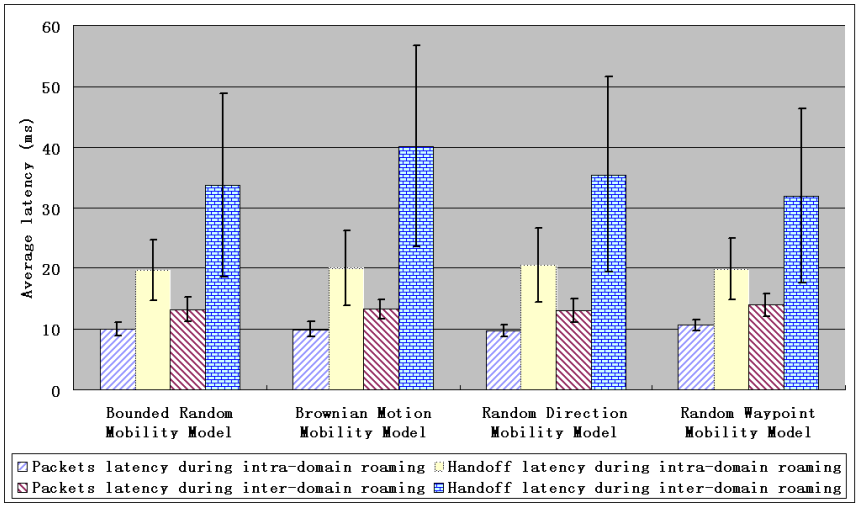


Figure 4.9: Average packet and handoff latency during intra-domain roaming

is around 2.5 ms. The 95% confidence range of the average packet latency is shown in Figure 4.9. We are 95% confident that the true packet latency will be between 8.73 ms and 11.57 ms.

The relationship between packet latency and the associated access router, in a given simulation of each mobility model, is illustrated in Figures 4.10 - 4.13. The left y axis shows the latency of each packet, and the right y axis represents the access router with which the mobile client associates. In Figures 4.10 - 4.13, most packets arrive in time at the mobile client. However, some exceptions do occur, since, according to the CSMA/CA-based medium access approach in IEEE 802.11, the packets should be retransmitted when the mesh routers experience a collision. The average latency during intra-domain roaming is around 10 ms, and the average latency per hop is around 2.5 ms. The 95% confidence range of the average packet latency is shown in Figure 4.9. We are 95% confident that the true packet latency will be between 8.73 ms and 11.57 ms.

(b) Handoff overhead and latency

In intra-domain roaming, there are six kinds of overhead traffic which can be categorized into three groups:

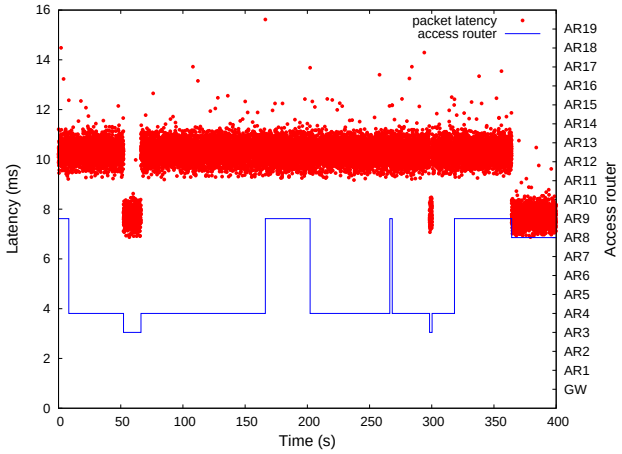


Figure 4.10: Packets received in the case of BRMM during intra-domain roaming

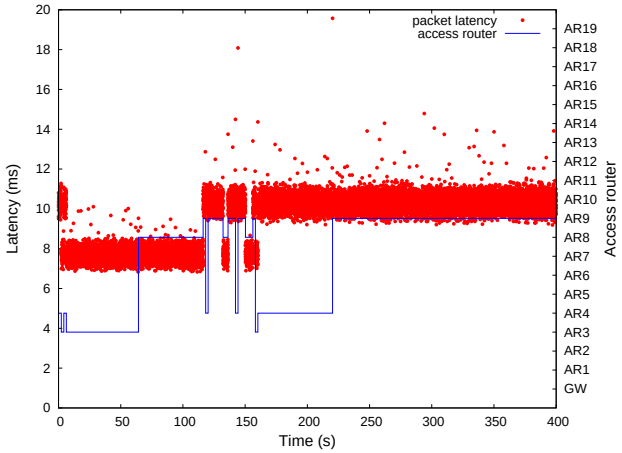


Figure 4.11: Packets received in the case of BMMM during intra-domain roaming

1. *Link control.* There are two kinds of link control traffic: probe messages and reply messages. The mobile client broadcasts a probe message to all mesh routers in its neighborhood every 2 seconds, and waits for the reply messages to calculate the link quality of the mesh routers. The probe message and the reply message are both 40 bytes long. The link control traffic depends on the number of mesh routers in the mobile client's neighborhood. More mesh routers cause higher link control traffic.

2. *Association control.* Three types of messages are included: association request

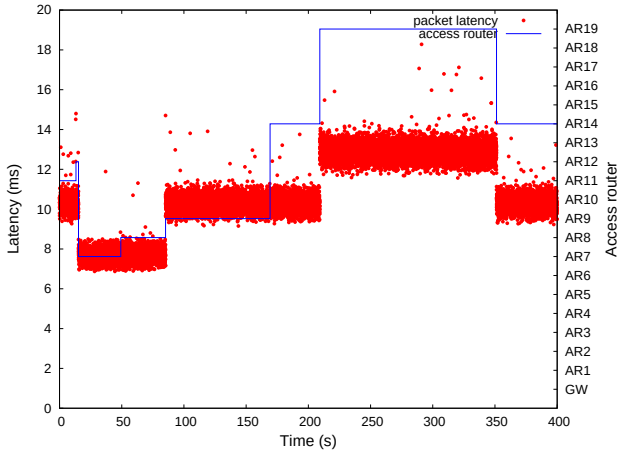


Figure 4.12: Packets received in the case of RDMM during intra-domain roaming

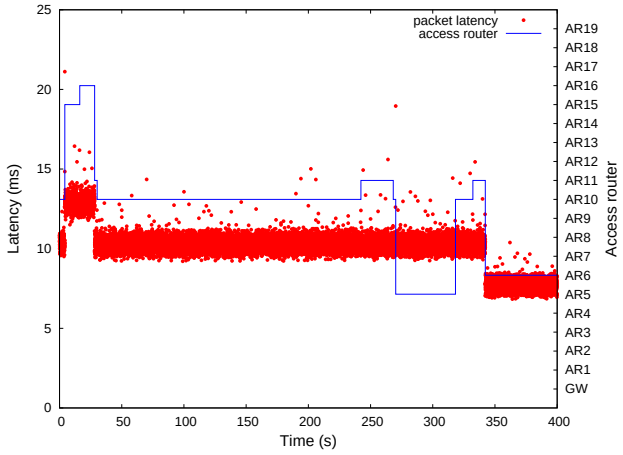


Figure 4.13: Packets received in the case of RWPM during intra-domain roaming

messages, association confirmation messages and disassociation messages. The 144-bytes request message is composed of the mobile client's IP address and MAC address, the old access router's IP address and MAC address, and the certification information. The confirmation message is used for the new AR authorizing the mesh client to connect. The size of a confirmation message is 144 bytes. The disassociation message is 64 bytes long. It includes the mobile client's IP address, and the new AR's IP address and MAC address. The association control overhead increases in accordance with the number of handoffs.

3. *Gratuitous ARP.* A gratuitous ARP message is 28 bytes long. The gratuitous ARP traffic grows with the number of handoffs and the number of correspondent nodes.

The average throughput rates for the three categories of overhead are shown in Table 4.7. The total overhead is quite low, and the negative impact on CBR traffic throughput is minimized. The link control traffic consumes the most bandwidth, which is around 1 Kbps per client. Association control traffic is only affected by the number of handoffs. For each handoff, the total overhead of association is 352 bytes. Thus, the maximum association traffic rate is 1.375 Kbps per mobile node. GARP traffic also consumes little bandwidth; even if the mesh client associates with a new AR every 2 seconds, the GARP traffic rate is 0.112 Kbps per CN. However, the GARP traffic grows linearly with the number of correspondent nodes; it is still quite small. Therefore, our solution has high scalability.

In addition, Figure 4.14 presents a quantitative comparative study of intra-domain handoff overhead. Our scheme is compared with three other schemes: M^3 [30], MEMO [61] and SMesh [8]. These three schemes belong to different types. According to the simulation results, our scheme has the lowest overhead. MEMO, the routing-based solution, has the highest overhead. This is because routers need to update their routing tables during the handoff to find a new routing path for the mesh client. Moreover, SMesh has to maintain multicast groups for forwarding packets and M^3 needs to update the mesh client's location information on the location server; therefore, these two schemes introduce more overhead than ours.

Table 4.7: Average handoff overhead and latency during intra-domain roaming

Mobility model	Handoff number	Overhead (Kbps)	Latency (ms)
Bounded Random	11.545	1.102	19.682
Brownian Motion	10.788	1.192	20.068
Random Direction	5.273	1.009	20.496
Random Waypoint	5.909	1.009	19.899

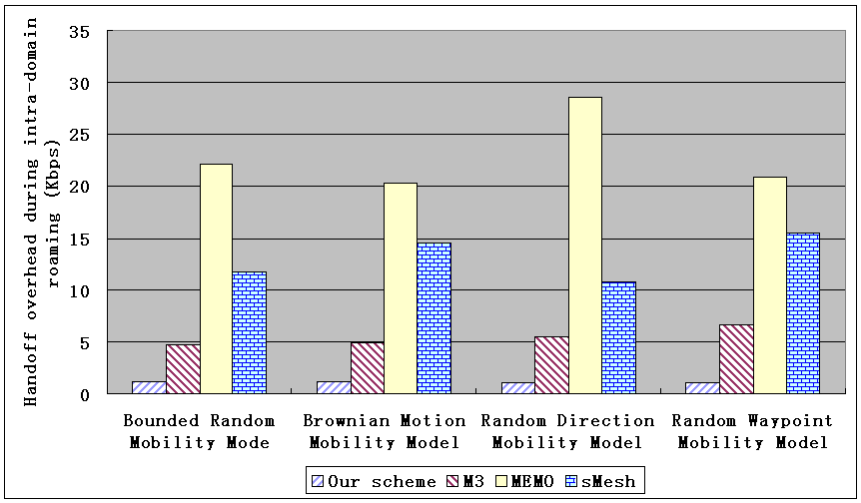


Figure 4.14: Average handoff overhead during intra-domain roaming

Total intra-domain handoff latency consists of two parts: MAC layer handoff latency, which refers to the time for associating with the new AR in the MAC layer, and network layer handoff latency which is defined as the period that begins when the mobile client starts handoff and ends when the correspondent node receives the GARP message. Using our fast MAC layer handoff scheme, the average MAC layer handoff latency should be around 40 - 50 ms. The network layer handoff latency of our solution is shown in Table 4.7, and the 95% confidence range of the average packet latency is shown in Figure 4.9. In our multi-hop mesh networks, we are 95% confident that the true network layer handoff latency would be between 13.84 ms and 26.54 ms. Therefore, the total handoff latency would be around 53.84 - 76.54 ms. Figure 4.15 illustrates the network layer intra-domain handoff latency of three schemes: our scheme, M^3 and MEMO. SMesh uses multicast to eliminate the handoff latency, and is not included in the figure. Our scheme and M^3 take a similar amount of time to complete handoff in the network layer, which is around 20 ms. And MEMO requires more time to complete the network layer handoff, which is around 45 ms.

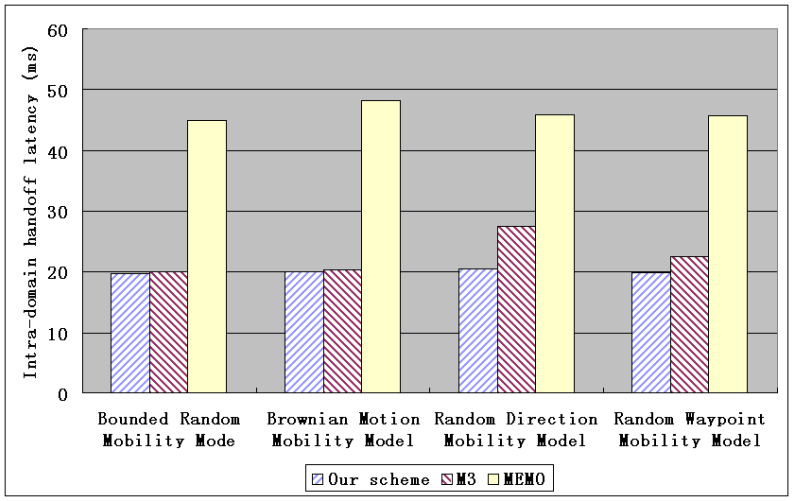


Figure 4.15: Average intra-domain handoff latency in the network layer

4.5.2 Inter-domain Roaming

In this subsection, the performance of our scheme during inter-domain roaming is illustrated. In these experiments, thirty access routers were deployed randomly and three domains were constructed by these mesh routers, and we changed the topology of the mesh network in each test. As with the intra-domain roaming scenario, there were thirty mobile nodes in each simulation experiment.

(a) Packet latency and loss ratio

To evaluate our proposed scheme during inter-domain roaming, we randomly select two mobile nodes to establish a CBR flow between them, and both can move among different domains. In each simulation, 20000 UDP packets were sent in 400 sec. Table 4.8 shows the average packet loss ratio, latency and the 95% confidence range of the latency during inter-domain roaming. Although a few packets were lost due to handoffs and conflicts in wireless transmission, most of the packets were received by the nodes correctly. Compared to intra-domain roaming, inter-domain roaming has a high loss ratio, because paths typically require more hops, causing higher loss probability.

The average packet-forwarding latency during inter-domain roaming is still quite

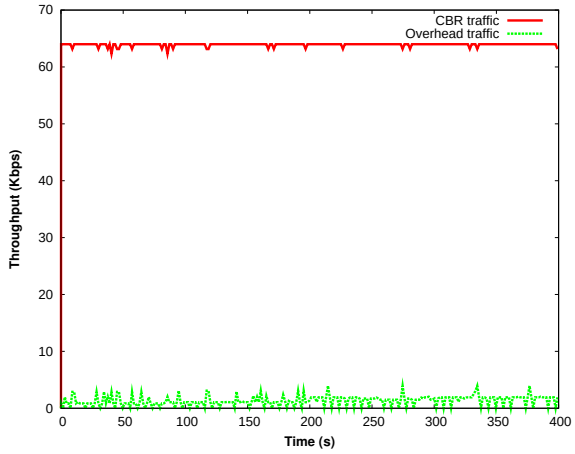


Figure 4.16: Traffic throughput and overhead throughput in the case of BRMM

low. The best case occurs with the Random Direction Mobility Model, with a latency is 12.997 ms; and the worst case is that of the Random Waypoint Mobility Model, with a latency is 13.968 ms. In addition, we are 95% confident that the real packet latency throughout would be between 11.494 ms and 15.277 ms. The average packet latency is also greater here than in the case of intra-domain roaming, due to extra hops introduced by the tunnels.

(b) Handoff overhead and latency

In these experiments, the average overhead during inter-domain roaming is assessed. We classify the overhead into three categories.

1. *Link control.* During inter-domain roaming, link control traffic is the same as during intra-domain roaming. It depends on the size of the link control message and the

Table 4.8: Average packets-loss ratio and latency during inter-domain roaming

Mobility model	Packets lost	Loss ratio	Latency (ms)	Confidence range
Bounded Random	107.317	0.536%	13.236	11.186 - 15.287
Brownian Motion	120.866	0.604%	13.342	11.735 - 14.949
Random Direction	105.960	0.529%	12.997	11.023 - 14.971
Random Waypoint	122.571	0.613%	13.968	12.034 - 15.902

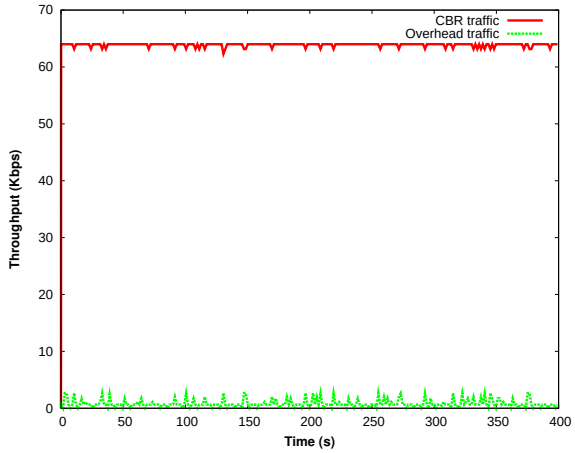


Figure 4.17: Traffic throughput and overhead throughput in the case of BMMM

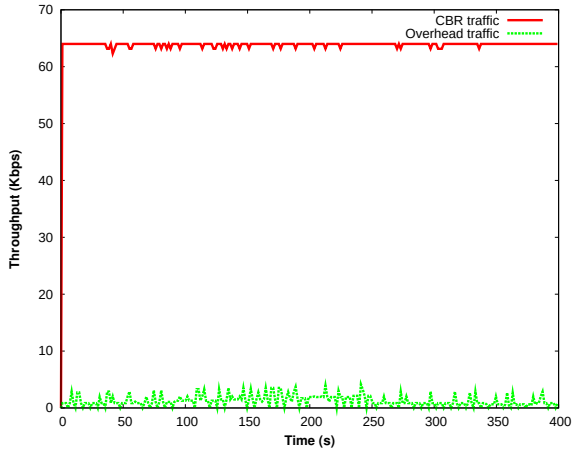


Figure 4.18: Traffic throughput and overhead throughput in the case of RDMM

probe interval. Both the probe message and the reply message are 40-bytes long. The probe interval is still set to 2 *sec*.

2. *Handoff overhead.* The handoff overhead includes both the intra-domain handoff overhead and inter-domain handoff overhead. The intra-domain handoff traffic includes association control traffic and GARP traffic, as discussed above. The inter-domain handoff traffic also includes association request messages, association confirmation messages, gratuitous ARP messages and disassociation messages. However, the association request message is 160 bytes, the association confirma-

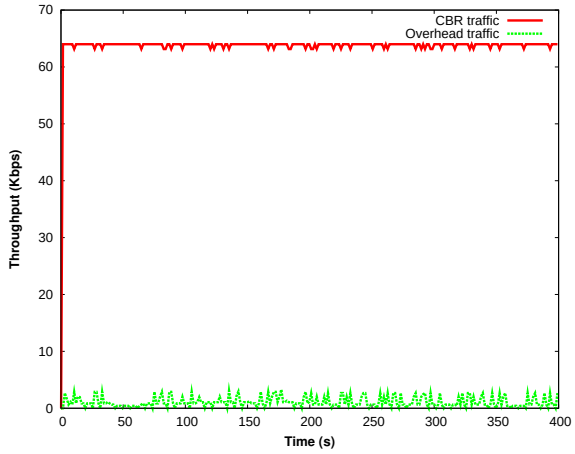


Figure 4.19: Traffic throughput and overhead throughput in the case of RWPM

tion message is 176 bytes, and the disassociation message is 96 bytes long. The gratuitous ARP message is still 28 bytes long. When the number of the handoffs increases, the handoff overhead increases.

3. *Tunnel control.* To reduce the number of hops during inter-domain roaming, tunnel control messages are sent to remove redundant tunnels. A tunnel control message is 64 bytes long, and it costs the lowest amount of bandwidth in overhead.

Table 4.9 shows overhead traffic during inter-domain roaming under four mobility models. As with intra-domain roaming, link control traffic consumes the most bandwidth. The total overhead traffic introduced by our scheme during inter-domain roaming is very low which is around 1.2 Kbps. The link control traffic consumes the most bandwidth which is around 1 Kbps. Figures 4.16 - 4.19 show a comparison between CBR traffic throughput and the overhead traffic throughput in a simulation of each mobility model.

Table 4.9: Average handoff overhead and latency during inter-domain roaming

Mobility model	Link (Kbps)	Handoff (Kbps)	Tunnel (Kbps)	Latency (ms)
Bounded Random	1.023	0.102	0.002	33.682
Brownian Motion	1.132	0.101	0.003	40.068
Random Direction	0.997	0.118	0.003	35.496
Random Waypoint	0.983	0.093	0.002	31.899

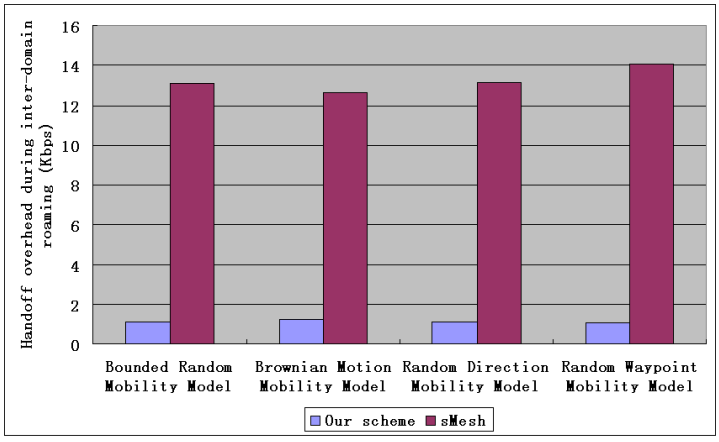


Figure 4.20: Average handoff overhead during inter-domain roaming

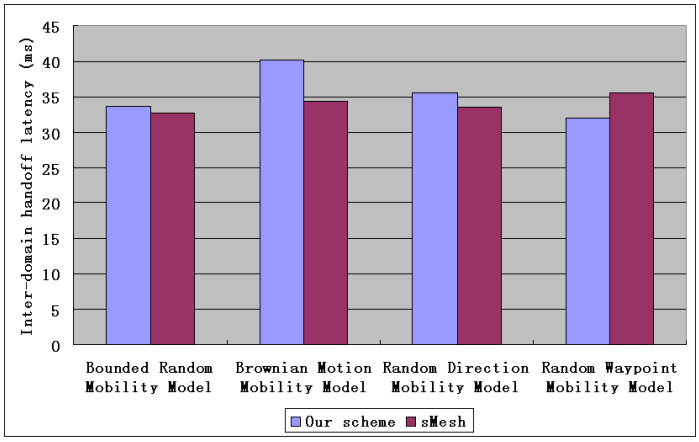


Figure 4.21: Average inter-domain handoff latency in network layer

In addition, the comparison of our scheme and SMesh [9] on handoff overhead is shown in Figure 4.20. Our solution causes less overhead than SMesh, because SMesh introduces large overhead for multicast group management.

In our experiments, the network layer handoff is measured as the time period between the triggering of the association request by the mobile client and the establishing of a tunnel between the original gateway and the new gateway. Average network layer handoff latency under four mobility models is shown in Figure 4.21, and the 95% confidence range of the network layer handoff latency is shown in Figure 4.9. Figure 4.21 also presents the comparison of our solution with SMesh on inter-domain handoff latency. According to

our experiments, our scheme and SMesh perform similarly in terms of handoff latency. The inter-domain handoff latency is around 42.57 - 96.523 ms, and it successfully achieves the requirements of a real-time application, which is 120 ms. Thus, during inter-domain roaming, our scheme can still be used to support smooth real-time applications.

4.6 Summary

In this chapter, a hybrid routing algorithm, which routes with both a MAC address at the link layer and an IP address at the network layer, is introduced. With this routing algorithm, during the intra-domain handoff process, location updates in the centralized location server are avoided, and re-routing after the handoff is not required. Moreover, a fast inter-domain handoff management scheme is also realized in WMNs. According to the experiment results, smooth handoff can be achieved with a lower overhead cost, and with the packet latency and loss ratio remaining at a lower level in both intra-domain roaming and inter-domain roaming. Our scheme thus can be used to support smooth real-time applications.

Chapter 5

Handoff Management over Vehicular Mesh Networks

In this chapter, our research work on the handoff management over Vehicular Mesh Networks (VMNs) is presented. First, a multi-hop clustering protocol for vehicular mesh networks is introduced. Then, based on this multi-hop cluster protocol, a fast handoff scheme for vehicular mesh networks is proposed.

5.1 Introduction

Vehicular mesh networks emerge as an innovative wireless network architecture in recent years. Based on VMNs, a lot of new applications have been introduced, such as intelligent transportation systems and mobile entertainment environment. A typical vehicular network includes two kinds of communication nodes, roadside nodes and mobile nodes. Roadside nodes are the access points or base stations. They can communicate with each others through wired or wireless connections and provide relay or gateway functions for mobile nodes. Roadside nodes are usually fixed and have strong capacity of communication. Conversely, mobile nodes are vehicles equipped with antennas for wireless communication. Therefore, vehicles can communicate with each other in ad-hoc

mode or access Internet through roadside nodes. The communications among vehicles are called vehicle to vehicle (V2V) communication; and the communications between vehicles and roadside nodes are called vehicle to infrastructure (V2I) communication. To construct VMNs, different wireless technologies can be used to provide physical links, such as IEEE 802.11 [31], WiMax [35], UMTS, and etc, and we will not address them in details in this chapter.

Compared with traditional wireless networks, VMNs have a lot of advantages, such as sufficient energy and enough space. However, the high speed of vehicles introduces more frequent handoffs. Therefore, the question of how to reduce handoff latency and overhead is very important. Handoff is the process of switching access points of mobile nodes during their movement. In traditional wireless networks, because the transmission radio range of antennas is limited, the mobile nodes have to associate with different access points to maintain their wireless connections while they are moving. Consequently, in vehicular environment, vehicles also have to switch access points if they want to maintain the connections with the infrastructure backbone. In this chapter, a network layer handoff scheme which is based on Mobile IP for VMNs is proposed. This scheme is independent of specific MAC layer protocols and has been proved to significantly reduce handoff latency.

In our scheme, the Network Mobility (NEMO) [22] scheme, which can reduce the total number of handoffs, is adopted to provide basic handoff functions in VMNs. To implement NEMO in vehicular environment, a multi-hop clustering scheme is proposed. Using the clustering scheme, vehicles are divided into clusters and each cluster constructs a mobile network. In each cluster, one vehicle is selected as the cluster head node and works as the mobile router in the mobile network. Other vehicles communicate with roadside units through the mobile router within the same mobile network. However, NEMO does not provide a solution to reduce the time of each single handoff, which still may cost long time. In our scheme, the communications among cluster nodes are used to reduce the time cost for each handoff. Cluster nodes in the front of the cluster respecting to the moving direction are selected as the assistant nodes, they will detect the signal

strength of access points in their neighborhood. When one assistant node finds that the signal quality of the current access point with which the mobile router associates is below the threshold, it starts to scan channels to find the new access point, and gets the new care of address for the mobile router. After that, it sends a message which contains its current position, the information of the new access point and the new care of address, to the mobile router. The mobile router keeps these information in its cache, and when the mobile router wants to switch its access point, it selects the new access point from the cache based on its position. Then, the mobile router tries to associate with the new access point directly. If the association is successful, the handoff latency can be reduced significantly; otherwise, the mobile router will start a complete handoff process.

5.2 The Multi-hop Clustering Scheme

In this section, our multi-hop clustering scheme for VMNs is presented. The clustering scheme has been well studied in wireless ad-hoc networks in recent years [79]. However, considering the characteristic of VMNs, such as high speed, sufficient energy and etc., the clustering schemes proposed for conventional wireless ad-hoc networks are not suitable for VMNs. Therefore, clustering schemes for VMNs should be designed specifically.

5.2.1 Definitions

The basic idea is that we allow the vehicle nodes broadcast beacon messages periodically. Upon receiving two consecutive beacon messages, the vehicle nodes can calculate relative mobility with other vehicle nodes in its N-hop neighborhood. The relative mobility metrics are then used to calculate the aggregate mobility metric; the vehicle nodes which have the lowest aggregate mobility are selected as cluster head nodes. Other vehicle nodes will join the cluster when they receive the messages from cluster head nodes. The notations will be used in this section are illustrated in Table 5.1. In addition, the definitions of some items are shown as follows.

Table 5.1: Notations used for the multi-hop clustering scheme

Notation	Description
$V(i)$	Vehicle i
$Dis(i,j)$	The distance between $V(i)$ and $V(j)$ in hops
$Cluster(m,n)$	A n-hop cluster and the cluster head node of the cluster is $V(m)$
$PktDelay(i,j,n)$	The packet delay of a packet sent from $V(i)$ to $V(j)$ using n hops
$RelM(i,j,n)$	The n-hop relative mobility metric between $V(i)$ and $V(j)$
$AggM(i,N)$	The N-hop aggregate mobility metric of $V(i)$
N	The maximum distance allowed between the cluster head node and the cluster member nodes in hops

If the vehicle node V_i can receive a data packet from the vehicle node V_j in N hops, the vehicle node V_i is the N-hop neighbor of the vehicle node V_j; and the distance between the vehicle node V_i and the vehicle node V_j is N in hops. The property of N-hop neighbors is symmetrical. We assume that the wireless transmission is bi-directional. If the vehicle node V_i can receive a data packet from the vehicle node V_j in N hops, the vehicle node V_j can also receive the data packet from the vehicle node V_i in N hops. Therefore, the vehicle node V_j is also the N-hop neighbor of the vehicle node V_i, if the vehicle node V_i is the N-hop neighbor of the vehicle node V_j.

A N-hop cluster is a cluster in which the cluster head node is the N-hop neighbor of all cluster member nodes. As a result, the maximum distance in hops between the cluster head node and cluster member nodes is N. Moreover, for a N-hop cluster, the diameter in hops should be less than 2N, which means in a N-hop cluster, any two cluster nodes can communicate with each other using less than 2N hops in ad-hoc mode .

The N-hop relative mobility metric is used to represent the relative mobility between two vehicle nodes which are N-hop neighbors in the vehicular network. If absolute value of the N-hop relative mobility metric is low, the relative position of two vehicle

nodes is stable. The N-hop aggregate mobility metric is used to represent the summary relative mobility between the vehicle node and all of other vehicle nodes in its N-hop neighborhood.

5.2.2 Mobility Metrics

Mobility metrics are used to represent the mobility level of mobile nodes. To calculate relative mobility between two vehicle nodes, a lot of solutions are proposed to represent the metric in the literature, such as relative speed, relative position, ratio of received signal strength from two consecutive packets and etc. Unfortunately, these metrics are not suitable to calculate the N-hop relative mobility for constructing N-hop clusters. Relative speed, which is one of the most common mobility metrics for mobile nodes, can be used to represent and predict the relative position of mobile nodes. However, it is not good to represent N-hop relative mobility for vehicle nodes, because in the vehicular environment, fading effects caused by obstacles cannot be ignored. Even two vehicle nodes have similar speeds, it is possible that wireless connections could be extremely weak. For relative position, it is possible that two nodes are in the radio range of each other physically and they cannot communicate with each other in one hop because of obstacles. Moreover, the ratio of received signal strength from two consecutive packets proposed in [12] can only be used to represent relative mobility in one hop. Consequently, a new relative mobility metric should be implemented to construct N-hop clusters.

In this section, a new mobility metric is proposed to represent the N-hop relative mobility between two vehicle nodes. The ratio of packet deliver delay of two consecutive packets is used to calculate the N-hop relative mobility. Every vehicle node is allowed to broadcast a beacon message in its neighborhood for every beacon interval. In the beacon message, the time when the vehicle broadcasts the messages is encapsulated. When the neighbor node receives the beacon message, it calculates the packet transmission delay and saves the packet delay in a data structure called neighbor list. If a vehicle node receives two consecutive beacon messages from the same node, it can compute

the relative mobility between them. The formula used to compute the relative mobility metric is shown in Equation 5.1.

$$RelM(i,j,n) = 10log_{10}\frac{PktDelay_{new}(i,j,n)}{PktDelay_{old}(i,j,n)} \tag{5.1}$$

In the beacon message, the number of maximum hops allowed is also encapsulated. The maximum number of hops is used to control the distance between the cluster head node and the cluster member nodes. Therefore, if we are going to create N-hop clusters, the maximum hop number should be set to N. Besides the maximum hop number, the current hop number is also included in the beacon message. When a neighbor node receives the beacon message, it increases the current hop number by one and checks whether the current hop number is less than the maximum hop number. If the current hop number is less than the maximum hop number, the neighbor node will broadcast the beacon message again. Otherwise, the neighbor node just calculates the packet transmission delay and updates its neighbor list according to the beacon message. In the new beacon message, besides the sending time of previous hops, the time to send the new beacon message is also appended. Therefore, when a vehicle node receives a beacon message, it can calculates the packet delay for all the nodes which the beacon message passed. Moreover, to reduce the number of beacon messages, if the vehicle node finds that it receives or forwards the packet before, the vehicle node drops the packet.

Based on the relative mobility metrics for neighbors in N-hop distance, the vehicle node can compute the aggregate mobility value using Equation 5.2.

$$AggM(i,N) = \sum_{Dis(i,j)\leq N} \|RelM(i,j,n)\| \times \frac{n}{N} \tag{5.2}$$

The aggregate mobility metric equals the summary of the relative mobility times a weight value for all neighbor nodes in N-hop. The weight metric is used to represent the contribution of different relative mobility to the whole aggregate mobility. Because the vehicle node which can access in less hops is prone to stay in the N-hop neighborhood longer,

the weight value of that vehicle node should be assigned a small value. The vehicle nodes which have higher hops are more possible to change the clusters. After calculating the aggregate mobility metric, vehicle nodes broadcast their aggregate mobility value in the N-hop neighborhood. The vehicle node which has the smallest aggregate mobility value is selected as the cluster head node; and other vehicle nodes work as the cluster member nodes.

5.2.3 Multi-hop Clustering

The detail steps of our multi-hop clustering scheme is illustrated in this subsection. The basic idea of our clustering scheme is that the cluster head node is selected based on the mobility metric defined in the previous subsection; and other vehicle nodes will join a cluster in its neighborhood. In general, there are three kinds of states for a vehicle node, Cluster_Undecided, Cluster_Member and Cluster_Head. When the a vehicle node joins the vehicle network, it initializes the state status and the state is set to Cluster_Undecided. Then the vehicle node can switch its status among these states during the movement. For a vehicle node which is in the undecided state, if it has the lowest aggregate mobility value in its N-hop neighborhood, it will go to the cluster head state; otherwise, it changes to the state Cluster_Member. For a cluster member node, if it loses the connection or the cluster is destroyed, it switches its state to Cluster_Undecided. Moreover, for a cluster head node, if it meets another cluster head node and the re-clustering process is triggered, it switches its state to Cluster_Undecided if it does not have the lowest aggregate mobility value. To control the clustering process, a parameter called beacon interval is predefined to specify the time interval between two consecutive beacon process.

In the first step, every vehicle node which wants to work in the cluster mode should broadcast a beacon message in its neighborhood. The broadcasting process will be repeated every beacon interval. In this beacon message, the sender's address, the transmitted time and the current hop number are encapsulated for calculating the relative

Algorithm 7 The process of receiving a beacon message

```
1: for all V(i) in the path list of the beacon message do
2:    if V(i) == V(j) then
3:       drop the beacon message;
4:    end if
5: end for
6: for all V(i) is in the path list of the beacon message do
7:    PktDelay(i, j, n) = Now - TT(i);
8:    if V(i) is in the neighbor list then
9:       calculate the relative mobility metric RelM(i, j, n);
10:   else
11:      add V(i) into the neighbor list;
12:   end if
13: end for
14: Nhop = Nhop + 1;
15: if Nhop < N then
16:   TT(j) = Now;
17:   add V(j), TT(j) into the path list of the beacon message;
18:   broadcast the beacon message;
19: end if
```

mobility metric. When a vehicle node receives a beacon message, it checks the path list in the beacon message to make sure the beacon packet is not received before. Then it calculates the packet transmission delay between itself and the vehicle nodes appeared in the path list. After that, the vehicle node computes the relative mobility value and updates the neighbor list. Moreover, the vehicle node checks whether the current hop number of the beacon message equals the maximum hop number. If the current hop number is smaller than the maximum hop number, the vehicle node should add itself into the path list of the beacon message and broadcast it again. The process of receiving a beacon message is shown in Algorithm 7. After the vehicle node starts the broadcasting process, it initializes a timer using the predefined value. When the timer expires, the vehicle node stops receiving the beacon messages.

After that, the vehicle node initializes the second timer for receiving aggregate mobility messages and it starts to calculate the aggregate mobility metric using Equation 5.2. The vehicle node then broadcasts its aggregate mobility value in its neighborhood. The

neighbor nodes, which receive the aggregate mobility message, update their neighbor lists and save the aggregate mobility values. For the aggregate mobility message, a parameter called current hop number is also encapsulated to control the forwarding process. When the vehicle node receives the aggregate mobility message, the current hop number is increased by one; and the current hop number is compared with the maximum hops allowed. If the current hop number is less than the maximum hop number, the vehicle node adds its own aggregate mobility value into the aggregate mobility message and broadcasts the message in its neighborhood again. After the timer expires, the vehicle node compares its aggregate mobility value with other vehicle nodes' aggregate mobility values in its neighbor list. If the vehicle node has the smallest value, it changes its state to cluster head node and broadcasts vehicle cluster information message in N hops. Otherwise, the vehicle node will be a cluster member node and listens to the cluster information message. The vehicle node will join a cluster if it hears the cluster information message from the cluster head node of that cluster. A special case is that a vehicle node receives multiple cluster information messages. In that case, the vehicle node selects the cluster head node which is the closest one in hops, and joins the cluster led by that cluster head node. If several cluster head nodes have the same hops to the vehicle node, the vehicle node joins the cluster led by the cluster head node which has the lowest relative mobility value. The process of selecting the cluster is shown in Algorithm 8.

Using our clustering scheme, at the beginning, the distance between two cluster head nodes should be more than N in hops. However, after moving, it is possible that two cluster head nodes meet each other during the moving path. When two cluster head nodes can contact with each other using less than N hops. The re-clustering process will be triggered. However, to reduce the re-clustering cost, the re-clustering process is deferred. Instead of starting the re-clustering process immediately, the re-clustering process is started when the two cluster head nodes are in the contact range for several broadcast intervals. Therefore, the re-clustering time is reduced and the cluster head duration time is increased.

Algorithm 8 The process of selecting the cluster

1: $SNhop$: the smallest hop distance between the vehicle node and a cluster head node
2: $V(j)$: the ID of the vehicle
3: $V(c)$: the cluster head node candidate
4: $SNhop = N$
5: **for all** $V(i)$ in the neighbor list **do**
6: **if** $V(i)$ is a cluster head node **then**
7: **if** $SNhop > Dis(i,j)$ **then**
8: $SNhop = Dis(i,j)$;
9: $V(c) = V(i)$
10: **else if** $SNhop == Dis(i,j)$ AND $RelM(i,j) < RelM(c,j)$ **then**
11: $V(c) = V(i)$
12: **end if**
13: **end if**
14: **end for**
15: $V(j)$ joins the cluster which is led by $V(c)$;

5.3 Fast Handoff for VMNs

Based on the proposed multi-hop clustering algorithm, the detailed design of our fast handoff scheme for VMNs is presented in this section. The basic idea is that vehicles are divided into different clusters and each cluster works as a mobile network to forward data packets. And specific vehicles are assigned to work as assistant nodes which help mobile router nodes reduce handoff latency.

To present our proposed scheme without loss of generality, a sample scenario is illustrated in Figure 5.1. In this example, a vehicle network consisted of vehicle nodes and roadside units is presented. Vehicle nodes are mobile vehicles with wireless antennas, and roadside units are access points. Vehicles are divided into different clusters based on the proposed multi-hop clustering algorithm. After dividing, each cluster works as a mobile network to forward data packets. In the cluster, one vehicle is selected as the cluster head node. This cluster head node works as the mobile router in the vehicle mesh network; and other mobile nodes in the same cluster have to communicate with the backbone through the cluster head node. For example, in Figure 5.1, vehicle A5 is the cluster head node of the mobile network A. Moreover, a special node in the cluster is

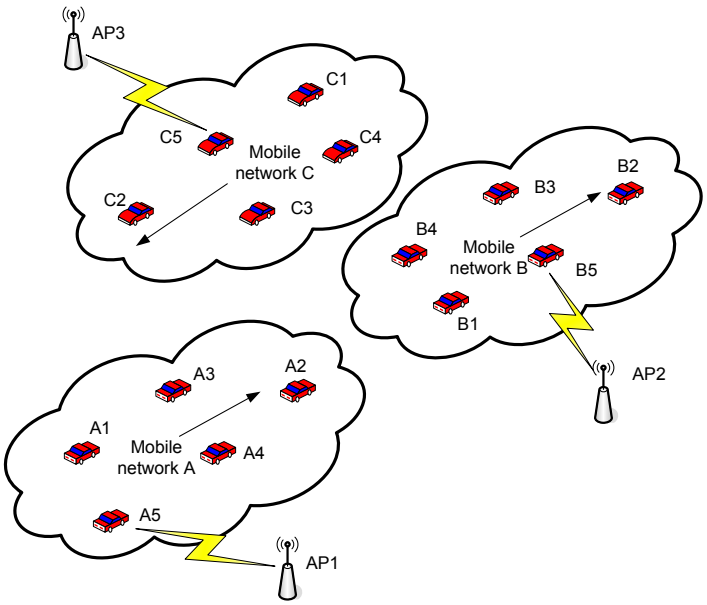


Figure 5.1: A sample scenario of VMNs in the urban environment

selected as the assistant node. The assistant node is the vehicle which is in the front of the cluster concerned with the moving direction of the cluster. In our sample scenario, vehicles A2, A3 and A4 can be the assistant node. Therefore, the mobile node should have a position system to detect its position and moving direction. We suppose all of the vehicles are equipped with position systems such as GPS. The assistant node will help the cluster head node probe and select the new access point in our handoff scheme. To gain the advantages of using assistant nodes, the cluster head node and the assistant nodes cannot be the same node except there is only one node in the group; and the cluster head node should not be in the front of the cluster.

While vehicles are moving, the connection between the cluster head node and the associated access point will be interrupted due to the limited transmission range. To maintain the wireless connection with the backbone, the cluster head node should find the new access point and establish the connection with the new access point. Therefore, when the cluster head node senses that the signal quality of the current access point is below the acceptable level, it should start the handoff process. However, a complete

handoff process would cost a lot of time, and we want to reduce the handoff latency for minimizing the packet loss ratio and the jitter. To reduce the handoff latency, the idea of using assistant nodes is introduced in our scheme to provide smooth handoff.

Assistant nodes are vehicles in the front of the cluster in respect of the moving direction. They are determined by the cluster head nodes when the cluster is created. The assistant nodes will be changed when some nodes join or leave the cluster, or the relative position of vehicles changes greatly. Although assistant nodes send and receive data packets through the cluster head node, they can sense the signal quality of access points in their neighborhood. When a mobile vehicle is selected as the assistant node, it sends a query message to the cluster head node to get the information of the current associated access point of the cluster head node. This query message can be sent periodically by the assistant node. Upon receiving the query message, the cluster head node sends the information of the current associated access point to the assistant node. Then, the assistant node starts to monitor the signal quality of the current access point. When the assistant node finds that the signal of the current access point is below the predefined threshold, it broadcasts probe messages to detect available access points. Access points in the assistant node's vicinity will send probe response messages to the assistant node. The assistant node then selects the access point which has the best signal quality as the candidate access point. After the candidate access point is selected, the assistant node sends a handoff preparation message to the candidate access point to trigger the pre-handoff process. Upon receiving the preparation message, the candidate access point determines whether to accept the connection request from the cluster head in future; and sends a preparation confirmation message to the assistant node to return the result. If the request is denied, the assistant node should start to search for and negotiate with another candidate access point. Otherwise, the candidate access point assigns a new care of address to the mobile network; and this care of address is piggybacked in the confirmation message.

When the assistant node gets the new care of address for the incoming handoff, it

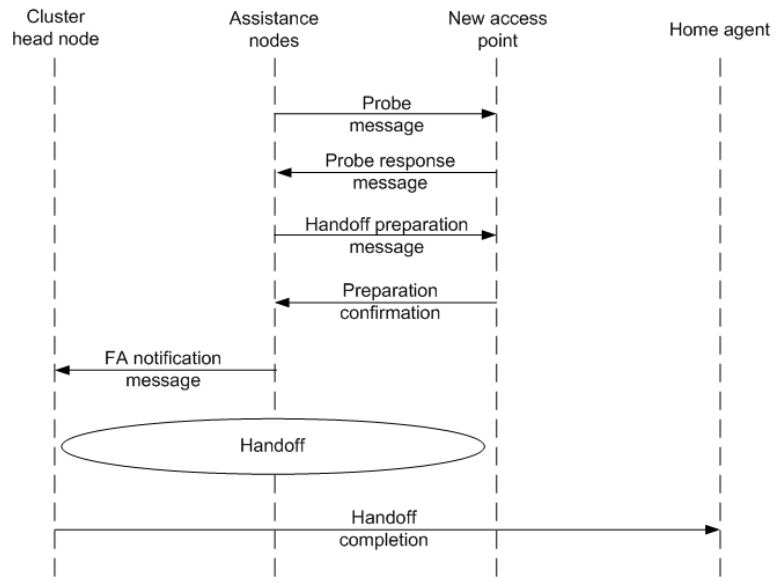


Figure 5.2: Messages exchanged for the proposed handoff scheme

sends a foreign agent notification message to the cluster head node. In this notification message, the new care of address, the information of the new access point and the position of the assistant node when it starts the probe process are included. Upon receiving the notification message, the cluster head node updates its candidate access point table. A candidate access point table is a data structure used to save available access points' information for incoming handoffs. The data entry of the candidate access point table includes the address of the access point, the care of address, the high probability area and the alive time. The high probability area is determined by the position of the assistant node, which is included in the notification message, and the predefined range. When the cluster head node starts handoff in that area, it means that it is probable that the cluster head node associates with the corresponding access point after completing the handoff process. The alive time is used to determine the effective time of the table entry information. After the alive time is up, the table entry will be deleted to reduce the overhead.

While the cluster head finds the signal quality of the current access point is below the handoff threshold, it searches the candidate access point table to check if there is any

available access point close to its current position. If it finds a possible access point in its table, it tries to associate with the access point directly. If the association is successful, the cluster head node sends its home agent the new care of address that is recorded in the table. Therefore, the time used to find the new access point in different channels and negotiate with the new access point to have the new care of address can be waived. All of the messages exchanged for the handoff are illustrated in Figure 5.2.

5.4 Experimental Results

In this section, we will present our experimental results of our multi-hop clustering scheme and the handoff management scheme in vehicular mesh networks.

5.4.1 The Multi-hop Clustering Scheme

To simulate our multi-hop clustering scheme, the Network Simulator - ns2 [53] is used to implement the clustering protocol. To generate the movement path file, two different mobility models are used to evaluate our clustering scheme extensively: the freeway mobility model and the Manhattan mobility model [80]. For the Manhattan mobility model, the probability of moving forward is set to 0.5, and the probability of turning left or right is set to 0.25 separately. The other general simulation parameters are illustrated in Table 5.2. For each test, the simulation runs 600 seconds; however, the clustering process starts at 300 seconds. To evaluate our multi-hop clustering protocol, three metrics are selected to demonstrate the performance: cluster head duration, cluster member duration and cluster head change number. These performance metrics can illustrate the stability of our clustering scheme.

The cluster head duration is the time interval from when the vehicle node becomes the cluster head node to when the vehicle node gives up the cluster head role. The average cluster head duration of our scheme with two mobility models are shown in Figure 5.3 and Figure 5.4 separately. In our simulation tests, the maximum number of hops in the

cluster is changed to find how the number of the maximum hops effect the performance. Three different values are used in the simulation, N = 2, 3 and 5. Figure 5.3 illustrates the average cluster head duration when we used Manhattan mobility model to generate moving path files. We did the tests using different maximum velocity. According to Figure 5.3, the average cluster head duration will decrease when the maximum velocity of vehicles increases. This is because when the vehicles move faster, the topology of the vehicle network is more dynamic. When the maximum speed changes from 10 m/s to 35 m/s, the cluster head duration is reduced about 20%. Moreover, the number of the maximum hops used to cluster vehicles will also affect the performance significantly. If we assign a large value for N, the cluster head duration is larger than the one when we use a small value for N. This is because the value of the maximum hops determines the distance between two cluster head nodes. If a large maximum hops number is used, the distance between two cluster head nodes should be far. Therefore, the time interval that two cluster head nodes spent to meet each other is long, and the time interval between two re-clustering processes is also long. As a result, the cluster head duration becomes long.

The average cluster head duration of scenarios in which the freeway mobility model is used is shown in Figure 5.4. Similar to the results of the Manhattan mobility model, the average cluster head duration is decreased when the maximum velocity of the vehicles is increased; and when the value of maximum hops increases, the average cluster head

Table 5.2: Simulation parameters for the clustering scheme

Parameters	Value
Simulation time	600 s
Area range	1000 m * 1000 m
Maximum velocity	10 - 35 m/s
Number of vehicles	100
Transmission range	120 m
Transmission rate	54 Mbps
Propagation model	Two-ray ground model

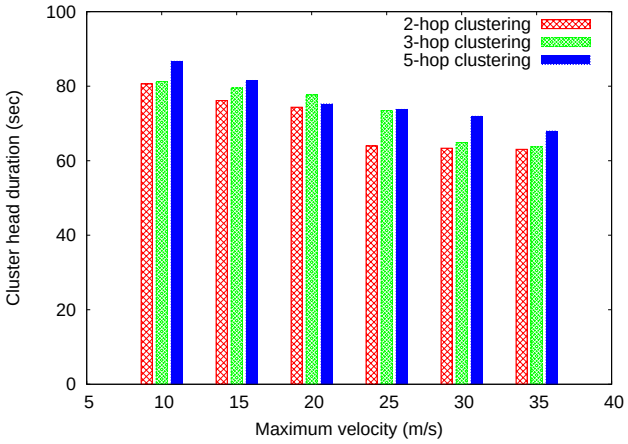


Figure 5.3: Average cluster head duration using Manhattan model

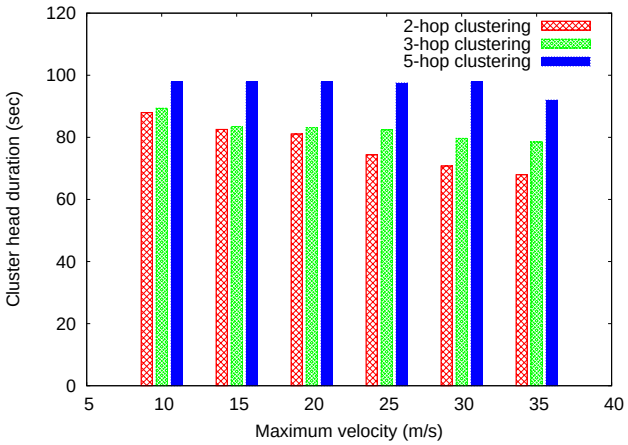


Figure 5.4: Average cluster head duration using the freeway model

duration also increases. Comparing with the results of two mobility models, we can see that the average cluster head duration using freeway mobility model is larger than the value using the Manhattan mobility model. This is because using the freeway mobility model, the vehicles have stronger connections with each other and the mobility is lower than the case where we use the Manhattan mobility model.

The average cluster member duration of our scheme under two mobility scenarios are shown in Figure 5.5 and Figure 5.6 separately. The cluster member duration is the time interval from the time when a vehicle node joins a specified cluster to the time when the vehicle node leaves the cluster. The cluster member duration also demonstrate the

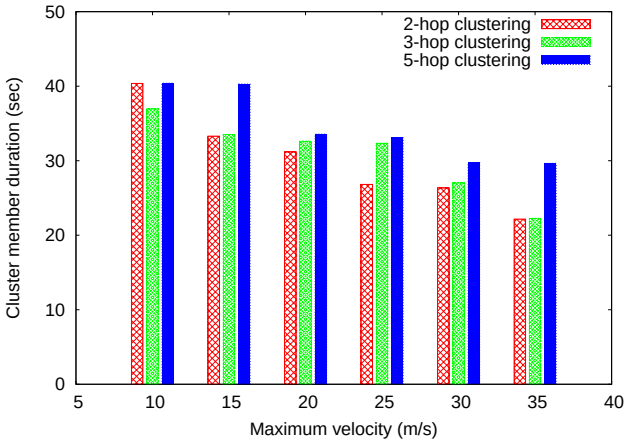


Figure 5.5: Average cluster member duration using Manhattan model

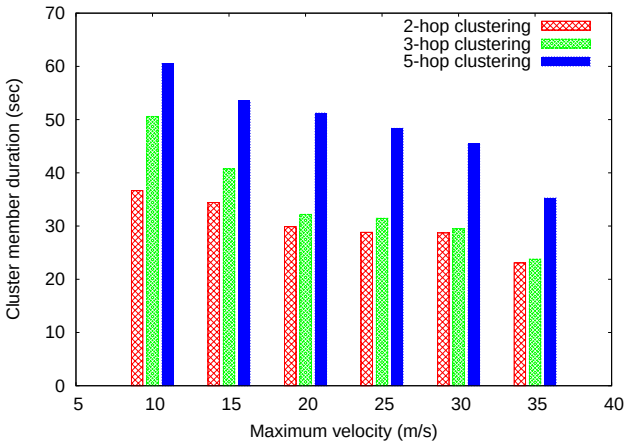


Figure 5.6: Average cluster member duration using the freeway model

stability of the clusters. Similar to the average cluster head duration, we use different maximum speed and maximum hops to show the relation between the parameters and the cluster member duration. In Figure 5.5, the average cluster member duration under the scenario which uses the Manhattan mobility model is illustrated. According to Figure 5.5, we can conclude that the average cluster member duration decreases when the maximum velocity of the vehicle nodes increases and the average cluster member duration increases when the maximum hops increases. When the maximum speed increases, the average vehicle speed will also increase. The vehicles which have high speed will have higher mobility than the vehicles which have lower speed. As a result, it is more possible for

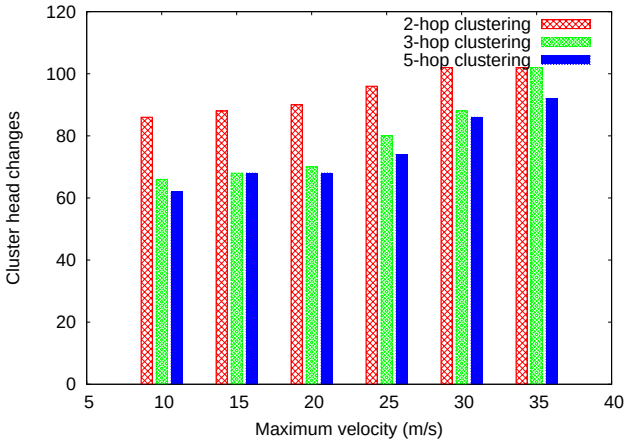


Figure 5.7: Average cluster head changes using Manhattan model

vehicles to change their clusters when the vehicles have high speed. The cluster member duration increases when the maximum hops increases, because it is more possible that the cluster member vehicles maintain their connections with the current cluster head nodes when we use a large maximum hop number.

Figure 5.6 illustrates the average cluster member duration using the freeway mobility model. It shows that the average cluster member duration has the same trend as the one using the Manhattan mobility model. The cluster member duration increases when the maximum velocity of vehicles decreases. In the meantime, when the maximum hop number increases, the cluster member duration time also increases. Moreover, the simulation results show that the performance using the freeway mobility model is better than the performance using the Manhattan mobility model. The reason is that when we use the movement paths which are generated using the freeway mobility model, the vehicles only go forward and there is no turning action. The mobility using the freeway mobility model is lower than the mobility using the Manhattan mobility model. Therefore, the average cluster member duration using the freeway mobility model is higher.

Cluster head change number is the number of the cluster head change during one simulation experiment. The cluster head change number increases when a vehicle becomes the cluster head node. The average cluster head change number of our scheme under two

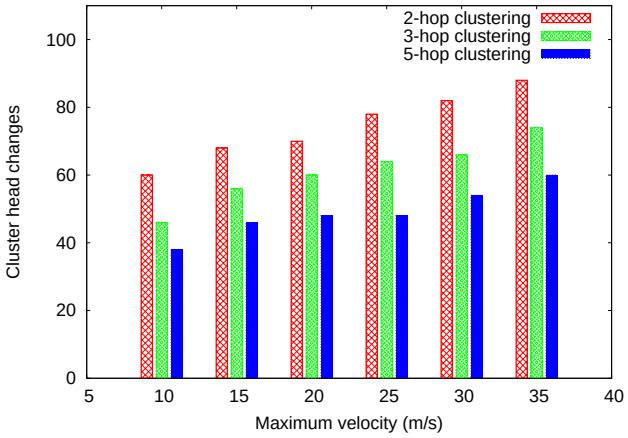


Figure 5.8: Average cluster head changes using the freeway model

mobility scenarios are shown in Figure 5.7 and Figure 5.8 separately. The cluster head change number can also demonstrate the stability of a clustering algorithm. When the vehicles have high mobility, a vehicle node will switch from the cluster member node to the cluster head node or from the cluster head node to cluster member node frequently. Therefore, in Figure 5.7 and Figure 5.8, we can see that the cluster head change number increases when the maximum velocity of vehicles increases, and it decreases when the maximum hop number increases. When the maximum velocity increases, the probability of two cluster head nodes meet each other is becoming higher; therefore, one of the cluster head nodes will become a cluster member node and the cluster will be removed. Some of the cluster member nodes in that group will join other groups, and some of the cluster member nodes will select a new cluster head node and construct a new group. Consequently, the cluster head change number increases. When we set a large number for the maximum hop number, the cluster nodes will have strong connections with each other and the cluster will be stabler. Therefore, the cluster head change number is decreased when the maximum hop number increases. Comparing the results of two mobility models, the average cluster head change number using the freeway mobility model is smaller than the number using the Manhattan model. The reason is that the mobility of vehicles which move according to the freeway mobility model is less than the

mobility of the vehicles using the Manhattan mobility model. Using the freeway model, vehicles only move forward; however, using the Manhattan mobility model, vehicles can take turns at the crosspoints.

5.4.2 Fast Handoff Scheme over VMNs

Performance evaluation of our fast handoff scheme for VMNs is presented in this subsection by illustrating extensive simulation and experimental results. The Network Simulator - ns2 [53] is used to evaluate our fast handoff scheme. To create movement paths for vehicles, the Manhattan mobility model [80] is used. The probability of moving straight is set to 0.5, and the probabilities of turning left and right are set to 0.25 separately. Different simulation scenarios are used to test our handoff scheme using different vehicle speed and background traffic. Table 5.3 illustrates general simulation parameters.

Handoff latency is defined as the transmission interruption time during the handoff process. It is the time interval from when the mobile router disconnects with the old access point to when the home agent of the mobile router registers the new care of address of the mobile router. It is an important parameter to evaluate the efficiency of a handoff scheme. The handoff latency of our handoff scheme under different vehicle speed and background traffic is shown in Figure 5.9. The background traffic is changed from 0 Kbps to 512 Kbps. The handoff latency increases as the background traffic grows. It is reasonable because when the background traffic rises, the probability of packet

Table 5.3: Simulation parameters of the handoff scheme over VMNs

Parameters	Value
Simulation time	600 s
Area range	2000 m * 2000 m
Average velocity	5 - 30 m/s
Number of APs	60
Transmission range	130 m
Transmission rate	54 Mbps
Background traffic	0 - 512 Kbps

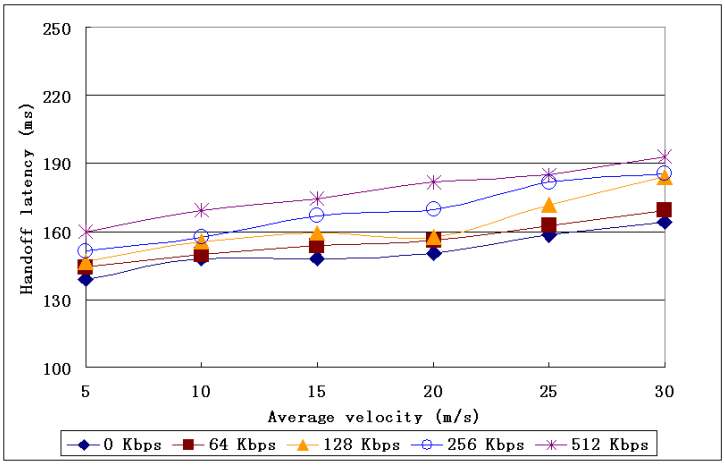


Figure 5.9: Handoff latency under different background traffic

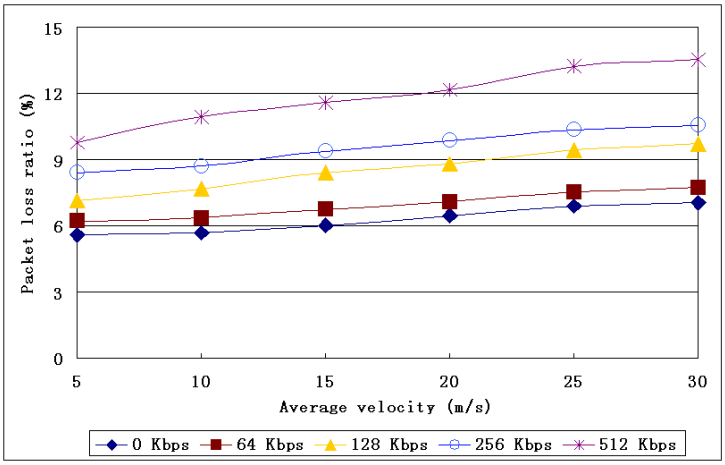


Figure 5.10: Packet loss ratio under different speed and background traffic

collision increases. The heavier background load is, the higher collision probability we get. To send frames, the mobile node has to wait longer time when the background traffic is increased. We also use different vehicle speed to test the handoff latency. The vehicle speed is changed from 5 m/s to 30 m/s. The handoff latency also increases slightly when the speed grows.

The packet loss ratio under different scenarios is illustrated in Figure 5.10. To test the data packet loss ratio, we establish CBR flows between vehicle nodes and wired nodes. To simulate G.711 [1] encoded/decoded VoIP applications, the wired node sends

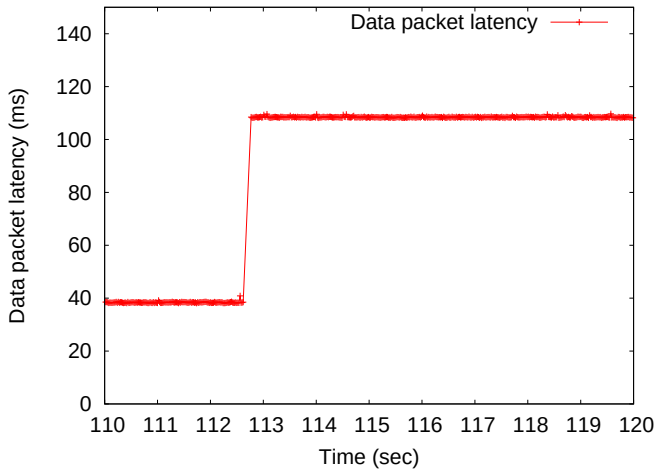


Figure 5.11: Data packet latency over VMNs

UDP packets to the vehicle nodes. The packet size is 160 byte and the packet interval is 20 *ms*. The UDP traffic starts at 100 *sec* and stops at 500 *sec*; and in total, there are 20000 UDP packets for each CBR flow. Similar to the handoff latency, the packet loss ratio also increases when the vehicle speed and the background traffic load grows. This is because when the background load grows, the wireless channel is busier and the probability of packet collision increases. The wireless node has to wait longer time to gain the right to use the wireless channel. Therefore, the packet loss ratio will increase when the background traffic grows.

The data packet latency and the inter-frame delay are the two other metrics to evaluate a handoff scheme. Like testing the packet loss ratio, the same CBR flow is used to test the data packet latency and the inter-frame delay. The data packet latency is defined as the time interval from when the packet is sent by the sender to when the mobile node receives the data packet. Figure 5.11 illustrates the data packet latency from 110 second to 120 second. According to Figure 5.11, data packet latency changes around 113 second because the mobile node completes a handoff process around that time. The data packet latency increases from around 40 *ms* to 110 *ms* after switching its access point, because the mobile node moves from its home agent's domain to a foreign domain. All of the

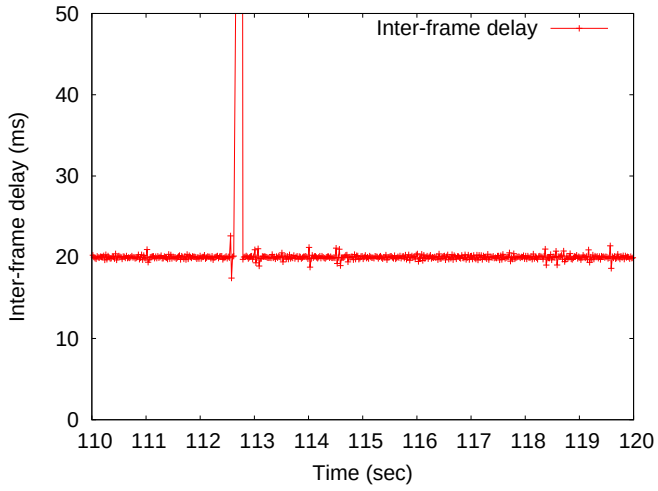


Figure 5.12: Inter-frame delay over VMNs

data packets sent to the mobile node will be sent to its home agent first. The home agent then sends these data packets to the mobile node's foreign agent. Therefore, the data packet latency increases.

Figure 5.12 illustrates the inter-frame delay from 110 second to 120 second accordingly. Inter-frame delay is the time interval between the receiver receives two consecutive data packets. Therefore, the best case is that the inter-frame delay is equal to the data packet interval, which should be 20 *ms* in our simulation. However, because of the inherent characteristic of the wireless communication, the value of inter-frame delay will change around the data packet interval. The inter-frame delay reflects the jitter level of the communication. According to Figure 5.12, around 113 second, the inter-frame delay increases to more than 50 *ms* when the handoff occurs. This is because during the handoff process, the mobile node disconnects with the old access point, and the communication between the mobile node and its correspondent node is interrupted. Some data packets are lost during this time period; consequently, the inter-frame delay increases.

In summary, the proposed handoff scheme can reduce handoff latency for VMNs significantly. In the meantime, the packet latency and the inter-frame delay are controlled in an acceptable level.

5.5 Summary

The vehicular mesh network is an innovative infrastructure to support new applications in urban environment. However, it also encounters the same problem of the traditional wireless networks caused by the mobility. Due to the limited transmission range of antennas, mesh clients in vehicular networks have to complete handoffs to keep wireless connections. Compared with traditional wireless networks, the vehicular mesh network has its own characteristic, such as high speed, dynamic topology, etc. These characteristics lead to more frequent handoffs in vehicular networks. Therefore, an efficient handoff scheme should be implemented in such networks. In vehicle mesh networks, constructing clusters can improve the handoff performance using NEMO scheme. The question of how to construct stable clusters for reducing cluster head changes and increasing cluster member duration is very important in VMNs. A lot of solutions have been presented in the literature; however, to the best of our knowledge, there is no solution to construct multi-hop clusters for VANETs. In this chapter, a new mobility metric to represent N-hop mobility for vehicular networks is proposed. Based on this mobility metric, a multi-hop clustering scheme is presented. Simulation results demonstrate that our scheme can cluster vehicle nodes efficiently. Using the proposed multi-hop clustering scheme, we divide vehicle nodes into different clusters; and each cluster works as a mobile network. In the cluster, we assign some nodes as assistance nodes, and these nodes can help the mobile router nodes scan channels and get new care of addresses before the handoff occurs. Simulation results illustrate that the handoff latency can be reduced using our scheme, and the packet interval can be maintained.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this thesis, our research work on handoff management is presented. An efficient handoff management scheme for traditional WMNs is proposed in this thesis. It includes two parts: the MAC layer handoff and the network layer handoff. Moreover, to support smooth roaming in high speed scenarios, a fast handoff scheme is introduced to reduce handoff latency for VMNs.

To reduce the MAC layer handoff latency for IEEE 802.11-based wireless mesh networks, a self-configured handoff scheme is proposed in this thesis. The standard handoff scheme defined in IEEE 802.11 [31] lasts more than 100 milliseconds, which is not suitable for real-time applications [47]. To provide smooth roaming for real-time applications, the maximal interruption time during the MAC layer handoff should be less than 50 ms [36]. Until now, most of the approaches are based on the fixed value of *MinChannelTime* and *MaxChannelTime*, which is not suitable for WMNs. Therefore, dynamic adaptation is used in our handoff scheme. Before the mesh client starts the probe process, it configures the value of *MinChannelTime* for each channel; this is done according to the scan result in the last handoff or the neighboring ARs' information received from the served AR. If a channel is empty in the last scan, we decrease *MinChannelTime*; otherwise we increase

MinChannelTime. During the probe process, *MaxChannelTime* is adapted based on the signal strength of the frames captured before the *ProbeTimer* reaches *MinChannelTime*. In addition, the mesh client can terminate the scan when it senses that the signal quality of a candidate AR is good enough to waive scans for other ARs. As a result, our scheme can minimize handoff latency by reducing the waiting time for each channel and the number of scanned channels simultaneously.

In the meantime, an efficient MAC layer authentication scheme is proposed to improve the handoff performance. The tunnel technique is introduced to reduce the handoff latency and provide secure communications. During the MAC layer handoff, when the mesh client selects the new access router which has the best signal quality, it starts a fast authentication process. The ID of the new access router, which can be the MAC address of the new access router, is sent to the old access router. The old access router uses the MAC addresses of the new access router and the mesh client to generate the temporary tunnel key. The temporary tunnel key is then sent to the mesh client and the new access router. After the mesh client receives the temporary tunnel key, it triggers a general open system authentication with the new access router. If the open system authentication is passed, the mesh client is permitted to associate with the new access router conditionally. The condition is that all of the data packets sent from the mesh client to the new access router are encrypted by the temporary tunnel key, and these packets are then forwarded to the old access router. The old access router will finally send these packets to their destinations. In the meantime, the new access router triggers the Extensible Authentication Protocol-Transport Layer Security (EAP-TLS) [3] process with the authentication server to generate the new PMK and PTK. After receiving the new PMK and PTK, the temporary tunnel key becomes obsolete, and the mesh client can communicate with the new access router as usual. As a result, the mesh client can communicate with its correspondent node via the old access router and implement the authentication simultaneously.

For the network layer handoff, a hybrid routing based handoff management solution

is proposed. The hybrid routing algorithm involves both the link layer (layer 2) routing and the network layer (layer 3) routing. The key idea is that the mesh client reply to the ARP message with its access router's MAC address. Therefore, packets forwarded among mesh routers can use the link layer routing to decrease the encapsulation and decapsulation delay of IP datagram; and, packets communicated between the access routers and the mesh clients can use the network layer routing. Based on the hybrid routing, during the intra-domain handoff, the mesh client sends a gratuitous ARP message to the correspondent node or the gateway. In the gratuitous ARP message, the new AR's MAC address is encapsulated. Upon receiving the gratuitous ARP message, the correspondent node, or the gateway, maps the new AR's MAC address to the client's IP address. Therefore, after the intra-domain handoff, the data packets will be forwarded to the client correctly according to the hybrid routing algorithm. Using the proposed intra-domain handoff scheme, location updates in the centralized location server are avoided, and re-routing after the handoff is not required. When the mobile client moves among different domains, the inter-domain handoff is activated. Similar to the intra-domain handoff, the mesh client uses the gratuitous ARP message to update its new location. Compared with intra-domain roaming, the main difference is that the IP address of the mesh client changes during the inter-domain roaming. When the mesh client joins in the new domain, it will get a new private IP address assigned by the new server. Moreover, redundant tunnels which are used to forward packets between old access routers are eliminated periodically to save resources and reduce forwarding latency.

To support smooth roaming in VMNs, a NEMO-based handoff management scheme is proposed in this thesis. Using this scheme, vehicles are first divided into different multi-hop clusters. To represent N-hop mobility for VMNs, a new mobility metric is proposed; and based on this mobility metric, a multi-hop clustering scheme is presented. To the best of our knowledge, this is the first multi-hop clustering scheme for VMNs. Then, within each cluster, the network mobility solution is used to reduce the total number of handoffs. In each cluster, one vehicle is selected as the cluster head node and works

as the mobile router in the mobile network. Cluster nodes in the front of the cluster, concerned with the moving direction, are selected as the assistant nodes, they will detect the signal strength of access routers in their neighborhood. When one assistant node finds that the signal quality of the current access router with which the mobile router associates is below the threshold, it starts to scan channels to find the new access point, and gets the new care of address for the mobile router. After that, it sends a message which contains its current position, the information of the new access router and the new care of address, to the mobile router. Consequently, before the mobile routers start the actual handoff process, they can receive their new care of addresses through assistant nodes in the same cluster. Thus, the handoff latency can be significantly reduced by the proposed scheme.

6.2 Future work

For the future work, there are many possible directions that can develop from this thesis. An important direction is the implementation of vertical handoff to support smooth roaming in heterogeneous wireless mesh networks. In this thesis work, we assume that only one physical standard is used to construct the network, and the mesh clients only need to implement horizontal handoff. However, in a heterogeneous WMNs, different physical technologies can be used to provide wireless connections for mesh clients, and mesh clients can move among these different networks freely. Therefore, when a mesh client triggers the handoff, it would connect to another network which uses different technologies. For example, in a mesh network which includes an IEEE 802.11 network and an IEEE 802.16 network, when the mesh client starts the scan process, it has to scan channels for both the IEEE 802.11 network and the IEEE 802.16 network. Therefore, the scan latency in this vertical handoff would be larger than the latency in horizontal handoff. To support smooth roaming, the scan latency should be minimized. Moreover, the criteria of selecting a new access router should also be changed. Considering only

the signal quality is not adequate in the case of vertical handoff: the cost of wireless connection and the transmission range of the new access router should also be considered. Another possible direction is to implement our schemes in the real environment in order to analyze the performance. Although extensive simulation experiments are run to test the performance of our schemes, we would like to deploy the schemes in the testbed and to find if there is any problem with the implementation of the schemes in the real world. To implement our handoff schemes in the real world, there are two possible solutions: modifying existing drivers or implementing a middleware to provide additional handoff functions. Moreover, mathematical modeling and analytical studies could be given to demonstrate the benefits on the handoff performance using our schemes. We hope this research work will trigger more innovations to make wireless roaming seamlessly.

Bibliography

[1] ITU-T recommendation G.711. http://www.itu.int/rec/T-REC-G.711-198811-I/en, last accessed in 20-Jan-2012, 1989. G.711.

[2] B. Aboba, L. Blunk, J. Vollbrecht, J. Carlson, and H. Levkowetz. Extensible authentication protocol (EAP). http://www.ietf.org/rfc/rfc3748.txt, last accessed in 20-Jan-2012, June 2004. RFC-3748.

[3] B. Aboba and D. Simon. PPP EAP TLS authentication protocol. http://www.ietf.org/rfc/rfc2716.txt, last accessed in 20-Jan-2012, October 1999. RFC-2716.

[4] I. F. Akyildiz, J. Mcnair, J. S. M. Ho, H. Uzunalioglu, and W. Wang. Mobility management in next-generation wireless systems. *Proceedings of the IEEE*, 87(8):1347–1384, August 1999.

[5] I. F. Akyildiz, X. Wang, and W. Wang. Wireless mesh networks: A survey. *Computer Networks*, 47(4):445–487, March 2005.

[6] I. F. Akyildiz, J. Xie, and S. Mohanty. A survey of mobility management in next-generation all-IP-based wireless systems. *IEEE Wireless Communications*, 11(4):16–28, August 2004.

[7] Y. Amir and C. Danilov. Reliable communication in overlay networks. In *Proceedings of the IEEE Dependable Systems and Networks*, pages 511–520, June 2003.

[8] Y. Amir, C. Danilov, M. Hilsdale, R. Musaloiu-Elefteri, and N. Rivera. Fast handoff for seamless wireless mesh networks. In *Proceedings of the 4th International Conference on Mobile Systems, Applications and Services*, pages 83–95, June 2006.

[9] Y. Amir, C. Danilov, R. Musaloiu-Elefteri, and N. Rivera. An inter-domain routing protocol for multi-homed wireless mesh networks. In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, pages 1–10, June 2007.

[10] R. Baldessari, A. Festag, and J. Abeille. NEMO meets VANET: A deployability analysis of network mobility in vehicular communication. In *7th International Conference on ITS Telecommunications (ITST07)*, pages 1–6, 2007.

[11] S. Basagni. Distributed clustering for ad hoc networks. In *Parallel Architectures, Algorithms, and Networks, 1999. (I-SPAN '99) Proceedings. Fourth International-Symposium on*, pages 310–315, 1999.

[12] P. Basu, N. Khan, and T. D. C. Little. A mobility based metric for clustering in mobile ad hoc networks. In *International Workshop on Wireless Networks and Mobile Computing (WNMC2001)*, pages 413–418, April 2001.

[13] A. Boukerche. Algorithms and protocols for wireless and mobile ad hoc networks, Wiley & Sons, 2005.

[14] A. Boukerche. Handbook of algorithms for wireless networking and mobile computing, Chapman & Hall/CRC, 2005.

[15] A. Boukerche. Algorithms and protocols for wireless sensor networks, Wiley & Sons, 2008.

[16] A. Boukerche, H. A.B.F. Oliveira, E. F. Nakamura, and A. A.F. Loureiro. Vehicular ad hoc networks: A new challenge for localization-based systems. *Computer Communications*, 31(12):2838–2849, July 2008.

[17] A. T. Campbell, J. Gomez, S. Kim, A. G. Valko, C.-Y. Wan, and Z. R. Turanyi. Design, implementation, and evaluation of cellular IP. *IEEE Personal Communications*, 7(4):42–49, August 2000.

[18] A. T. Campbell, J. Gomez, S. Kim, C.-Y. Wan, Z. R. Turanyi, and A. G. Valko. Comparison of IP micromobility protocols. *IEEE Wireless Communications*, 9(1):72–82, February 2002.

[19] Y.-S. Chen, C.-H. Cheng, C.-S. Hsu, and G.-M. Chiu. Network mobility protocol for vehicular ad hoc networks. In *WCNC 2009*, pages 1–6, 2009.

[20] V. M. Chintala and Q. Zeng. Novel mac layer handoff schemes for ieee 802.11 wireless lans. In *IEEE Wireless Communications and Networking Conference*, pages 4435–4440, March 2007.

[21] T. Clausen and P. Jacquet. Optimized link state routing protocol (OLSR). http://www.ietf.org/rfc/rfc3626.txt, last accessed in 20-Jan-2012, October 2003. RFC-3626.

[22] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thubert. Network mobility (NEMO) basic support protocol. http://www.ietf.org/rfc/rfc3963.txt, last accessed in 20-Jan-2012, January 2005. RFC-3963.

[23] K. Egevang and P. Francis. The ip network address translator (NAT). http://www.ietf.org/rfc/rfc1631.txt, last accessed in 20-Jan-2012, May 1994. RFC-1631.

[24] S. R. Fluhrer, I. Mantin, and A. Shamir. Weaknesses in the key scheduling algorithm of RC4. In *the 8th Annual International Workshop on Selected Areas in Cryptography*, pages 1–24, August 2001.

[25] M. Gerla and J. T.-C. Tsai. Multicluster, mobile, multimedia radio network. *Wireless Networks*, 1:255–265, 1995.

[26] D. Gupta, J. LeBrun, P. Mohapatra, and C.-N. Chuah. Wds-based layer 2 routing for wireless mesh networks. In *Proceedings of the 1st International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization*, pages 99–100, September 2006.

[27] E. Gustafsson, A. Jonsson, and C. E. Perkins. Mobile IPv4 regional registration. http://www.ietf.org/rfc/rfc4857.txt, last accessed in 20-Jan-2012, June 2007. RFC-4857.

[28] J. Harri, C. Bonnet, and F. Filali. Kinetic mobility management applied to vehicular ad hoc network protocols. *Computer Communications*, 31(12):2907–2924, July 2008.

[29] C.-M. Huang, M.-S. Chiang, and T.-H. Hsu. PFC: A packet forwarding control scheme for vehicle handover over the its networks. *Computer Communications*, 31(12):2815–2826, July 2008.

[30] R. Huang, C. Zhang, and Y. Fang. A mobility management scheme for wireless mesh networks. In *IEEE Global Telecommunications Conference*, pages 5092–5096, November 2007.

[31] 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. http://standards.ieee.org/getieee802/802.11.html, last accessed in 20-Jan-2012, 2007.

[32] Channel deployment issues for 2.4-ghz 802.11 wlans. http://www.cisco.com/en/US/docs/wireless/technology/channel/deployment/guide/Channel.html, last accessed in 20-Jan-2012, 2007.

[33] Draft amendment to standard IEEE 802.11: Ess mesh networking. http://standards.ieee.org/getieee802/download/802.11g-2003.pdf, last accessed in 20-Jan-2012, 2006.

[34] IEEE 802.15 WPAN task group 1a (tg1a). http://www.ieee802.org/15/pub/TG1.html, last accessed in 20-Jan-2012, 2009.

[35] IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems amendment for physical and medium access control layers for combined fixed and mobile operation in licensed bands. http://standards.ieee.org/getieee802/802.16.html, last accessed in 20-Jan-2012, 2005.

[36] International Telecommunication Union. General characteristics of international telephone connections and international telephone circuits, 1988. ITU-TG.114.

[37] B. Jackson. History of voip. http://www.utdallas.edu/ bjackson/history.html, last accessed in 19-May-2008, November 2007.

[38] K. Karakayali, J. H. Kang, M. Kodialam, and K. Balachandran. Joint resource allocation and routing for ofdma-based broadband wireless mesh networks. In *IEEE International Conference on Communications, 2007*, pages 5088–5092, June 2007.

[39] M. Kassab, A. Belghith, J.-M. Bonnin, and S. Sassi. Fast pre-authentication based on proactive key distribution for 802.11 infrastructure networks. In *Proceedings of the 1st ACM Workshop on Wireless Multimedia Networking and Performance Modeling*, pages 46–53, October 2005.

[40] R. Koodli. Fast handovers for mobile IPv6. http://www.ietf.org/rfc/rfc4068.txt, last accessed in 20-Jan-2012, July 2005. RFC-4068.

[41] R. Koodli and C. E. Perkins. Mobile ipv4 fast handovers. http://www.ietf.org/rfc/rfc4988.txt, last accessed in 20-Jan-2012, October 2007. RFC-4988.

[42] G. Kousalya, P. Narayanasamy, J. H. Park, and T. Kim. Predictive handoff mechanism with real-time mobility tracking in a campus wide wireless network considering its. *Computer Communications*, 31(12):2781–2789, July 2008.

[43] S. Kuklinski and G. Wolny. Density based clustering algorithm for vanets. In *Testbeds and Research Infrastructures for the Development of Networks Communities and Workshops, 2009. TridentCom 2009. 5th International Conference on*, pages 1–6, April 2009.

[44] T. J. Kwon and M. Gerla. Efficient flooding with passive clustering (pc) in ad hoc networks. *Computer Communication Review*, 32(1):44–56, 2002.

[45] Y. Liao and L. Cao. Practical schemes for smooth MAC layer handoff in 802.11 wireless networks. In *International Symposium on a World of Wireless, Mobile and Multimedia Networks*, pages 1–10, June 2006.

[46] N. Maslekar, M. Boussedjra, J. Mouzna, and L. Houda. Direction based clustering algorithm for data dissemination in vehicular networks. In *Vehicular Networking Conference (VNC), 2009 IEEE*, pages 1–6, October 2009.

[47] A. Mishra, M. Shin, and W. Arbaugh. An empirical analysis of the IEEE 802.11 MAC layer handoff process. *ACM SIGCOMM Computer Communication Review*, 33(2):93–102, April 2003.

[48] A. Mishra, Min Ho Shin, N.L. Jr. Petroni, T.C. Clancy, and W.A. Arbaugh. Proactive key distribution using neighbor graphs. *IEEE Wireless Communications*, 11(1):26–36, February 2004.

[49] A. Misra, S. Das, A. Dutta, A. McAuley, and S. K. Das. IDMP-based fast handoffs and paging in IP-based 4G mobile networks. *IEEE Communications Magazine*, 40(3):138–145, March 2002.

[50] N. Mustafa, W. Mahmood, A. A. Chaudhry, and M. Ibrahim. Pre-scanning and dynamic caching for fast handoff at mac layer in ieee 802.11 wireless lans. In *IEEE International Conference on Mobile Adhoc and Sensor Systems Conference*, pages 1–8, November 2005.

[51] V. Navda, A. Kashyap, and S. R. Das. Design and evaluation of iMesh: an infrastructure-mode wireless mesh network. In *Sixth IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, pages 164–170, June 2005.

[52] M. Nekoui, A. Ghiamatyoun, S. N. Esfahani, and M. Soltan. Iterative cross layer schemes for throughput maximization in multi-channel wireless mesh networks. In *Proceedings of 16th International Conference on Computer Communications and Networks, 2007.*, pages 1088–1092, August 2007.

[53] The network simulator - ns-2. http://www.isi.edu/nsnam/ns/, last accessed in 22-Jul-2011.

[54] S. Pack and Y. Choi. Pre-authenticated fast handoff in a public wireless LAN based on IEEE 802.1x model. In *Proceedings of the IFIP TC6/WG6.8 Working Conference on Personal Wireless Communications*, pages 175–182, October 2002.

[55] C. E. Perkins. IP mobility support for IPv4. http://www.ietf.org/rfc/rfc3220.txt, last accessed in 20-Jan-2012, January 2002. RFC-3220.

[56] D. C. Plummer. An ethernet address resolution protocol or converting network protocol addresses to 48.bit ethernet address for transmission on ethernet hardware.

http://www.ietf.org/rfc/rfc826.txt, last accessed in 20-Jan-2012, November 1982. RFC-826.

[57] J. Postel. User datagram protocol. http://tools.ietf.org/html/rfc768, last accessed in 20-Jan-2012, August 1980. RFC-768.

[58] K. Ramachandran, S. Rangarajan, and J. C. Lin. Make-before-break mac layer handoff in 802.11 wireless networks. In *2006 IEEE International Conference on Communications*, volume 10, pages 4818–4823, June 2006.

[59] R. Ramjee, T. L. Porta, S. R. Thuel, K. Varadhan, and S.-Y. Wang. HAWAII: A domain-based approach for supporting mobility in wide-area wireless networks. In *Seventh International Conference on Network Protocols*, pages 283–292, November 1999.

[60] A. Raniwala and T. Chiueh. Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network. In *Proceeding of IEEE INFOCOM 2005*, volume 3, pages 2223–2234, March 2005.

[61] M. Ren, C. Liu, H. Zhao, T. Zhao, and W. Yan. MEMO: An applied wireless mesh network with client support and mobility management. In *IEEE Global Telecommunications Conference*, pages 5075–5079, November 2007.

[62] C. Rigney, W. Willats, and P. Calhoun. RADIUS extensions. http://www.ietf.org/rfc/rfc2869.txt, last accessed in 20-Jan-2012, June 2000. RFC-2869.

[63] C. Rigney, S. Willens, A. Rubens, and W. Simpson. Remote authentication dial in user service (RADIUS). http://www.ietf.org/rfc/rfc2865.txt, last accessed in 20-Jan-2012, June 2000. RFC-2865.

[64] M. Sabeur, G. A. Sukkar, B. Jouaber, D. Zeghlache, and H. Afifi. Mobile party: A mobility management solution for wireless mesh network. In *Third IEEE Interna-*

tional Conference on Wireless and Mobile Computing, Networking and Communications, pages 45–53, October 2007.

[65] C. Shea, B. Hassanabadi, and S. Valaee. Mobility-based clustering in vanets using affinity propagation. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6, December 2009.

[66] M. Shin, A. Mishra, and W. A. Arbaugh. Improving the latency of 802.11 handoffs using neighbor graphs. In *Proceedings of the 2nd International Conference on Mobile Systems, Applications, and Services*, pages 19–26, June 2004.

[67] S. Shin, A. G. Forte, A. S. Rawat, and H. Schulzrinne. Reducing mac layer handoff latency in ieee 802.11 wireless lans. In *Proceedings of the Second International Workshop on Mobility Management and Wireless Access Protocols*, pages 70–83, October 2004.

[68] skype. http://www.skype.com/intl/en-us/home/, last accessed in 15-Oct-2011, 2011.

[69] H. Soliman, C. Castelluccia, K. El Malki, and L.Bellier. Hierarchical mobile IPv6 mobility management (HMIPv6). http://www.ietf.org/rfc/rfc4140.txt, last accessed in 20-Jan-2012, August 2005. RFC-4140.

[70] S. Speicher. OLSR-FastSync: Fast post-handoff route discovery in wireless mesh networks. In *IEEE 64th Vehicular Technology Conference*, pages 1–5, September 2006.

[71] S. Speicher and C. H. Cap. Fast layer 3 handoffs in AODV-based IEEE 802.11 wireless mesh networks. In *3rd International Symposium on Wireless Communication Systems*, pages 233–237, September 2006.

[72] The spines overlay network. http://www.spines.org, last accessed in 20-Jan-2012.

[73] H. Velayos and G. Karlsson. Techniques to reduce the ieee 802.11b handoff time. In *2004 IEEE International Conference on Communications*, volume 7, pages 3844–3848, June 2004.

[74] H. Wang, Q. Huang, Y. Xia, Y. Wu, and Y. Yuan. A network-based local mobility management scheme for wireless mesh networks. In *IEEE Wireless Communications and Networking Conference*, pages 3792–3797, March 2007.

[75] S.-S. Wang and Y.-S. Lin. Performance evaluation of passive clustering based techniques for inter-vehicle communications. In *Wireless and Optical Communications Conference (WOCC), 2010 19th Annual*, pages 1–5, May 2010.

[76] H.-Y. Wei, S. Kim, S. Ganguly, and R. Izmailov. Seamless handoff support in wireless mesh networks. In *1st Workshop on Operator-Assisted(Wireless Mesh) Community Networks*, pages 1–8, September 2006.

[77] G. Wolny. Modified dmac clustering algorithm for vanets. In *Systems and Networks Communications, 2008. ICSNC '08. 3rd International Conference on*, pages 268–273, October 2008.

[78] World of warcraft. http://us.battle.net/wow/en/, last accessed in 15-Oct-2011, 2007.

[79] J. Y. Yu and P. H. J. Chong. A survey of clustering schemes for mobile ad hoc networks. *Communications Surveys Tutorials, IEEE*, 7(1):32–48, May 2005.

[80] B. Zhou, K. Xu, and M. Gerla. Group and swarm mobility models for ad hoc network scenarios using virtual tracks. In *IEEE Military Communications Conference, 2004. MILCOM 2004.*, pages 289–294, 2004.

[81] H. Zimmermann. OSI reference model - the ISO model of architecture for open systems interconnection. *IEEE Transactions on Communications*, 28(4):425–432, April 1980.