IP Mobility Support in Multi-hop Vehicular Communications Networks

by

Sandra Lorena Céspedes Umaña

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Vehicular communications networks are envisioned to be supported by a set of dissimilar wireless access networks and different administrative domains. The heterogeneous infrastructure will serve as the platform for the deployment of safety and infotainment applications, which will help on achieving a safer, efficient, and enjoyable transportation system. Lately, the support of infotainment services —and consequently, of IP-based applications— has drawn substantial attention. Traditional Internet-based applications and driver assistance services, as well as innovative peer-to-peer applications that enable the instant sharing of information between neighboring vehicles, are some of the services that will make traveling a more convenient and pleasant experience. In addition, it is expected that innovative services will incentive a faster adoption of the equipment and the supporting infrastructure required for vehicular communications.

On the other hand, the combination of infrastructure-to-vehicle and vehicle-to-vehicle communications, namely the multi-hop Vehicular Communications Network (VCN) , appears as a promising solution for the ubiquitous access to IP services in vehicular environments. For example, by employing multi-hop communications, the network coverage of the slowly growing infrastructure can be extended. In addition, longer bidirectional connections between road side access routers and vehicles can be established through multi-hop paths. The bidirectionality of links is a strong requirement of most IP applications, and it is difficult to be achieved when asymmetric links appear in the vehicular wireless network.

Although multi-hop communications have been often proposed for disseminating safety and delay-sensitive information, the deployment of seamless infotainment traffic faces unique challenges due to the characteristics of the highly-mobile and multi-hop VCN. Not only the standards for communications in vehicular environments suffer from limitations for the deployment of IP traffic, but also the IP mobility support in VCN has traditionally focused on vehicles using one-hop connections to the infrastructure. Additional complexity is added when urban vehicular scenarios are considered, in which commuters and pedestri-

ans actively access infotainment applications that should be freely transferrable along the heterogeneous VCN.

In this thesis, we address the challenges of multi-hop VCN, and investigate the seamless provision of IP services over such network. Three different schemes are proposed and analyzed. First, we study the limitations of current standards for the provision of IP services, such as 802.11p/WAVE, and propose a framework that enables multi-hop communications and a robust IP mobility mechanism over WAVE. An accurate analytical model is developed to evaluate the throughput performance, and to determine the feasibility of the deployment of IP-based services in 802.11p/WAVE networks. Next, the IP mobility support is extended to asymmetric multi-hop VCN. The proposed IP mobility and routing mechanisms react to the asymmetric links, and also employ geographic location and road traffic information to enable predictive handovers. Moreover, since multi-hop communications suffer from security threats, it ensures that all mobility signalling is authenticated among the participant vehicles. Last, we extend our study to a heterogeneous multi-hop VCN, and propose a hybrid scheme that allows for the on-going IP sessions to be transferred along the heterogeneous communications system. The proposed global IP mobility scheme focuses on urban vehicular scenarios, and enables seamless communications for in-vehicle networks, commuters, and pedestrians.

The overall performance of IP applications over multi-hop VCN are improved substantially by the proposed schemes. This is demonstrated by means of analytical evaluations, as well as extensive simulations that are carried out in realistic highway and urban vehicular scenarios. More importantly, we believe that our dissertation provides useful analytical tools, for evaluating the throughput and delay performance of IP applications in multi-hop vehicular environments. In addition, we provide a set of practical and efficient solutions for the seamless support of IP traffic along the heterogeneous and multi-hop vehicular network, which will help on achieving ubiquitous drive-thru Internet, and infotainment traffic access in both urban and highway scenarios.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Xuemin (Sherman) Shen, for his encouragement, guidance, and support during my graduate studies at the University of Waterloo. Prof. Shen's commitment with high quality research is an inspiration for pursuing excellence and success not only in academic endeavors but in every other aspect of life.

I would like to thank Dr. Liping Fu, Dr. Catherine Rosenberg, Dr. Liang-Liang Xie, and Dr. Patrick Mitran for serving on my Advisory Committee. Their careful reading and comments have significantly improved the quality of this thesis. I am also very thankful to Dr. Ben Liang for serving as the external thesis examiner and for his insightful comments and observations.

My appreciation goes to my colleagues and friends at the BBCR group. It has been a privilege for me to work with so many bright people. In particular, I would like to thank all the VANET/DTN subgroup members, who dedicated long hours of their time to listen and discuss my work. Among them, Mr. Ning Lu, Ms. Sanaa Taha, Dr. Mahdi Asefi, Dr. David (Bong Jun) Choi, Mr. Hassan Omar, Mrs. Khadige H. Abboud, Mr. Hao Liang, Dr. Tom H. Luan, and Mr. Sailesh Bharati. I am also greatful to Dr. Mohammad Towhidul Islam and Dr. Mohamed Mohamed Elsalih for their advice and friendship during the years we shared as office mates.

This dissertation could not have been completed without the company of so many friends who shared happiness and tough times with me for the past four years. I am truly and deeply thankful to Ahmad, a sincere and kind friend who was always generous with time and suggestions for my work, to Kolla, for our long hours of physical exercise and "coffee therapies" that helped us stay healthy, and to N.A., for being the cherished friend that was always there. I also thank Liliana, Jason, Ruth, and the little Jonathan, for being my family here in Canada from the very first moment I arrived. I am very thankful to Monica, I will never forget our endless conversations over skype, also to Luisa, Marcelita,

Milena, Juan, and all my dear friends that were always present from the distance.

My deep appreciation also goes to those friends who made of Waterloo a warm place to live: Rita and Leandro, Alondra, Maricris, Amparo and Nestor, Claudia and Rodrigo, Daniel and Soledad, Tomas and Angeles, Mary and Felipe, Alelí and Pancho, Nathy and German, Xiomara and Fernando, Shahira, and my roommates Sylvia, Monica, and Afsoon. Special thanks go to my dearest Gonzalo, who patiently and lovingly helped me and supported me during the writing of this dissertation.

Grateful acknowledgements are made for financial support from the Colfuturo Scholarship for Graduate Studies, Icesi University, Ontario Research & Development Challenge Fund Bell Scholarship, and numerous assistantship awards from the University of Waterloo. I am indebted to my professors, colleagues, friends, and students from Icesi University who supported me every step of the way.

Finally, I wish to express my deepest gratitude to my dear father, sister, uncles, cousins, grandmother, and my dearly beloved mother. Their unconditional love and support helped me achieve this dream. I would never get this far without you.

Dedication

To the memory of my beloved grandfather León

To my dear mother for her love and endless support

To my dear father for his unconditional love

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3GPP	3rd Generation Partnership Project
AR	Access Router
CBR	Constant Bit Rate
CCH	Control Channel
CN	Correspondent Node
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DAD	Duplicate Address Detection
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
DoS	Denial of Service
EDCA	Enhanced Distributed Channel Access
EPS	Evolved Packet System
ESP	Encapsulating Security Payload transport format
FQDN	Full Qualified Domain Names
GPS	Global Positioning System
GTP	GPRS Tunneling Protocol
HIP	Host Identity Protocol
HIT	Host Identity Tag
HNP	Home Network Prefix
HoA	Home Address

I2V2V	Infrastructure-to-Vehicle-to-Vehicle Communications
I2V	Infrastructure-to-Vehicle Communications
IETF	Internet Engineering Task Force
IP	Internet Protocol
KB	Kilobyte
LFN	Local Fixed Nodes
LMA	Local Mobility Anchor
LRVS	Local Rendezvous Server
LTE	Long Term Evolution
MA-PMIP	Multi-hop Authenticated Proxy Mobile IP
MAC	Medium Access Control
MAG	Mobile Access Gateway
MANEMO	MANET-centric NEMO
MANET	Mobile Ad hoc Network
MIPv6	Mobile IPv6
MITM	Man-In-The-Middle attack
mMAG	Mobile MAG
MN	Mobile Node
MR	Mobile Router
NAT	Network Address Translation
ND	Neighbor Discovery
NEMO BS	Network Mobility Basic Support
NMAG	Next MAG
OBU	On Board Unit
PBA	Proxy Binding Acknowledgement
PBU	Proxy Binding Update
PHY	Physical Layer
PMAG	Previous MAG

PMIP	Proxy Mobile IP
RA	Router Advertisement
RCPI	Received Channel Power Indicator
RN	Relay Node
RS	Router Soliciation
RSU	Road Side Unit
RTT	Round Trip Time
RVS	Rendezvous Server
SA	Security Association
SCH	Service Channel
TCP	Transport Control Protocol
UDP	User Datagram Protocol
V2I	Vehicle-to-Infrastructure Communications
V2V	Vehicle-to-Vehicle Communications
VANET	Vehicular Ad hoc Network
VBR	Variable Bit Rate
VCN	Vehicular Communications Network
VIP-WAVE	Vehicular IP in WAVE
VMN	Visitor Mobile Nodes
WAVE	Wireless Access in Vehicular Environments
WBSS	WAVE Basic Service Set
WLAN	Wireless Local Area Network
WRA	WAVE Router Advertisement
WSA	WAVE Service Advertisement
WSMP	WAVE Short Message Protocol

Chapter 1

Introduction

Vehicular communications networks have been envisioned as a heterogeneous system, in which a variety of applications, communications technologies, and protocols, converge to achieve a safer, efficient, and enjoyable transportation system. Among the applications to be deployed in the vehicular network, interest has been mostly directed to the deployment of safety-oriented applications, such as the notifications of car accidents and monitoring systems. However, the role of infotainment applications has rapidly taken an important place. From traditional Internet-based applications and driver assistance services, such as up-to-the-minute traffic reports and assisted parking, to innovative peer-to-peer applications that enable the instant sharing of information between neighboring vehicles, are some of the services that will make traveling a more convenient and pleasant experience.

In addition to enable access to innovative services designed for vehicular environments, the infotainment applications are likely to incentive a faster adoption of the equipment and the supporting infrastructure required for vehicular communications. In fact, it has been widely accepted that this supporting infrastructure and communications technologies will be heterogeneous in nature [1]. Large coverage access networks, such as 3G/4G cellular and WiMAX networks, will be combined with wireless local area networks (WLAN) , such as 802.11b/g/n. Moreover, they will also integrate WLAN technologies specifically designed

for vehicular environments, such as 802.11p/WAVE .

1.1 Multi-hop Vehicular Communications Networks

In terms of data dissemination, the vehicular network foresees that vehicles communicate via wireless links with other vehicles (V2V communications), in addition to communicate with the infrastructure via roadside access routers (V2I/I2V communications) . According to that, infrastructureless communications in the vehicular network are possible if only V2V communications take place. In addition, I2V and V2V communications can be combined to create an infrastructure-connected vehicular network, which we denominate *the Multi-hop Vehicular Communications Network*. In such scenario, vehicles communicate with the infrastructure using a single hop connection, whenever it is available, but they may also rely on other vehicles for data forwarding when direct connection to the infrastructure is not available.

Besides the scenario in which vehicles relay traffic from other vehicles, another form of multi-hop communications is when devices in the in-vehicle local network are at the same time mobility-enabled nodes or routers. For instance, a passenger traveling on a vehicle is monitored by a body area sensor network. This network selects a gateway sensor node for the forwarding of packets to an external network. Such gateway node is therefore also a mobile router. In such cases, the vehicle's mobile router is the first hop in the chain of hops the body sensor data would have to traverse to reach the destination. Therefore, this scenario also belongs to our definition of *Multi-hop Vehicular Communications Network*.

Although the combination of infrastructure-to-vehicle and V2V communications —also known as I2V2V communications— is promising, it has been often proposed for dissemination of safety and delay-sensitive information [2,3], but little for infotainment applications. In the case of safety applications, the scope is usually of broadcast nature or delimited to a certain geographic area, resulting in a well-defined strategy to be followed if multi-hop

paths become necessary during data dissemination. However, in the case of I2V2V communications for general infotainment applications, such as IP-based services and Internet access, more challenges arise if seamless communications are expected in the multi-hop vehicular network.

1.2 Research Challenges

The special characteristics of multi-hop vehicular communications networks create unique requirements for IP-based services deployment. Some challenges come from the vehicular system itself, such as the high speeds, the dynamic topology, and the spatial-temporal traffic variability. Additional research challenges inherent to the use of multi-hop communications in vehicular environments are described herein:

Limitations for the support of IP applications in current standards

Specialized technologies for the support of vehicular communications networks are being developed and standardized among vendors and standardization bodies. One leading technology example is the IEEE 802.11p [4] combined with the standards for Wireless Access in Vehicular Environments (WAVE) [5, 6]. Such a technology is envisioned to become a fundamental platform for providing real-time access to safety and entertainment information. However, the market penetration is expected to grow slowly, so it is very critical to guarantee seamless, reliable, and ubiquitous communications that provide a satisfactory user experience to the early adopters. In particular, infotainment applications —and consequently IP-based communications— are key to leverage market penetration and deployment costs of the 802.11p/WAVE network.

Previous research evaluate the performance of IP-based applications in I2V vehicular environments [7–9], but they often employ traditional 802.11b/g technologies that do not resemble the intricacies of 802.11p/WAVE for IP communications. Consequently, the op-

eration and performance of IP in 802.11p/WAVE are still unclear, as the WAVE standard guidelines for being IP-compliant are rather minimal. Three limitations for the operation of IP over WAVE have been identified: 1) lack of duplicate address detection; 2) lack of seamless communications for extended services (i.e., services that are consumed along several access networks); and 3) lack of support for multi-hop communications to help extend the coverage of the slowly growing infrastructure. With many open aspects for the operation of IPv6, providing access to IP-based applications, such as assisted parking, route management, and eventually Internet access, becomes a challenging task in 802.11p/WAVE networks.

IP mobility support

Due to the inherent dynamicity of the vehicular network, and the heterogeneity of the supporting infrastructure, it is reasonable to assume that vehicles may transfer their active connections through different IP access networks. Thus, the on-going IP sessions may be affected by the change of IP addresses, and consequently become broken connections. Previously, the research on IP mobility support has focused on vehicles using one-hop connections to the infrastructure [10–13]. The objective is to enhance the performance of existing IP mobility protocols, or to extend the support for the in-vehicle network nodes.

However, there are still not many research efforts devoted to the support of IP mobility when multi-hop connections are employed in the vehicular network. In such a case, the scenario becomes more complex if we consider the variability experienced by the links employed during V2V communications. Given the vehicles high mobility, multi-hop paths are of a short-time duration. WiFi experiments in urban and freeway scenarios, as well as analyses from simulated vehicular networks, indicate a range between 10s and 40s for contact duration between two moving vehicles [14, 15]). Consequently, protocols for IP mobility may take more time in trying to establishing a relayed connection than the time available for actually forwarding packets through the multi-hop path.

Asymmetric links

Asymmetric links in vehicular communications are typically caused by mobility, path losses, and dissimilar transmission powers between road-side Access Routers (ARs) and vehicles; thus, asymmetric links are likely to appear in the vehicular network. However, symmetric links have been a frequently-employed assumption when researching the deployment of IP services in vehicular environments [16].

Although one-way links may not affect some applications, e.g., safety information that requires only a downstream link (i.e., from AR to vehicles), this problem severely affects IP-based applications. On the one hand, if IP mobility is to be supported, the node is expected to indicate a location update to the infrastructure (in the form of a Binding Update message in Mobile IP, a Router Solicitation in Proxy Mobile IP, or a link-layer indicator when the connection is established). However, such location update cannot be delivered to the AR unless a bidirectional link exists. On the other hand, applications will be affected —specially TCP-based that require the packets to be acknowledged—, not only due to the inability to confirm reception of packets, but also due to the impossibility to initiate a client-server application, which requires the node to send a request to the server in order to start a service.

Impact of market penetration

The deployment of services in the vehicular network fundamentally depends on the deployment of in-vehicle and road-side communications equipments. The pace at which the equipments penetrate the market will highly affect the performance of the IP mobility solutions, which employ access routers and anchor points located at the infrastructure side for data dissemination. Therefore, the network-wide connectivity places an important role in the solutions' performance. Moreover, the distribution of equipped vehicles could be highly variable even for a contained geographic area. In the hypothetic case that all new vehicles were fully equipped for vehicular communications, they would be mixed with the existent fleet of vehicles that, in contrast, will follow a slow and gradual adoption pro-

cess [17]. Therefore, IP mobility solutions should handle the different market penetration rates of vehicular communications equipments over the short, medium, and long term.

Lack of motivation for data forwarding

Providing ubiquitous access to infrastructure-based IP services represents a major challenge due to the high cost involved in the installation of the roadside infrastructure. Therefore, the use of multi-hop communications come as a convenient solution for enabling ubiquitous connectivity, by means of using other vehicles as relays in order to reach the nearest road side access router. However, this requires for intermediate vehicles to forward packets that do not belong to the in-vehicle local network. Forwarding external packets consume and compete for resources that are supposed to be fully enjoyed by the local network, so intermediate vehicles may refuse to cooperate, resulting in scarcity of available relays.

Furthermore, if two vehicles decide to cooperate and participate in the relaying of packets, they are arbitrary mobile routers that have not met before. As a result, it becomes difficult to generate a security association between them, and this leads to security threats for both the infrastructure and the vehicles involved in the communications.

Vehicular network heterogeneity

Urban vehicular scenarios involve different patterns of mobility: low-speed mobility of commuters and pedestrians, and high-speed mobility of vehicles combined with complex spatio-temporal traffic conditions [17]. Nowadays, it is common to find people actively using Internet-based applications from high-end mobile devices, e.g., cellular phones and tablets. Such access is available even at vehicular speeds, thanks to 3G/4G cellular networks with large coverage around the urban areas, and a proprietary protocol, called GPRS Tunneling Protocol (GTP) , employed for IP mobility and billing, among other functions. Nevertheless, the on-going IP sessions are normally not transferable when the mobile device switches from the cellular network to a WLAN connection (or viceversa), unless complex

states are maintained by the application developer, in order to avoid resetting the sessions.

Vehicular communications, on the other hand, are envisioned to be supported by dissimilar access networks and independent network operators. Yet, pedestrians and commuters' mobile devices should be also considered as users in the vehicular scenario. If IP sessions are to be transferable along the whole vehicular network, and that means, from bus stations, to in-vehicle wireless networks, to WiFi hostpots, to the cellular network, and more, the IP mobility mechanism should be robust enough to make seamless communications possible.

1.3 Thesis Motivation and Contributions

Multi-hop communications come as a convenient solution for the ubiquitous access to IP services in vehicular communications networks. This research area is actually very important for several reasons. In the case a direct connection between vehicles and infrastructure is not available, the bidirectional links required by IP applications could be established by means of multi-hop communications. In this way, infrastructure networks that are in-process to be deployed, e.g., the recently standardized 802.11p/WAVE network, or that provide limited coverage, such as 802.11b/g/n hot spots, may benefit from an extended coverage thanks to data forwarding mechanisms through V2V communications [18]. Furthermore, when the coverage is not an issue thanks to the presence of a well-deployed infrastructure, such as 3G/LTE networks in urban scenarios, the multi-hop communications may decrease the levels of the energy consumption when signals have to cover shorter distances, as well as to improve the spectral efficiency, and to increase network capacity and throughput [19, 20].

As a result, data forwarding cooperation in VCN has been investigated through theoretical approaches [18, 21], as well as prototype implementations and field testbed evaluations [9, 22], which have demonstrated that multi-hop scenarios in the VCN are technically

sound. In addition, previous incentive schemes for data forwarding have shown that, instead of being a burden, the relay of packets can become a good practice in the benefit of the relay nodes [19, 23]. However, to provide IP services and IP mobility support in the multi-hop vehicular network represents major challenges, as outlined in Section 1.2. Therefore, in this research we aim at addressing those challenges by means of the following contributions:

1. We have studied the 802.11p/WAVE standard and have identified its limitations for the support of infrastructure-based IP communications. This have led us to propose the Vehicular IP in WAVE (VIP-WAVE) framework. VIP-WAVE defines the IP configuration for extended and non-extended IP services, and a mobility management scheme supported by Proxy Mobile IPv6 over WAVE. It also exploits multi-hop communications to improve the network performance along roads with different levels of infrastructure presence. Furthermore, an analytical model considering mobility, handoff delays, collisions, and channel conditions, has been developed for evaluating the performance of IP communications in WAVE. Extensive simulations are employed to demonstrate the accuracy of our analytical model, and the effectiveness of VIP-WAVE in making feasible the deployment of IP applications in 802.11p/WAVE networks.

2. Building upon the I2V2V concept, we have studied the secure provision of infrastructure-based IP services in asymmetric vehicular communications networks (VCN), and have proposed a Multi-hop Authenticated Proxy Mobile IP scheme (MA-PMIP) . MA-PMIP focuses on three different aspects: first, it provides an IP mobility scheme for multi-hop VCN, and employs location and road traffic information in order to predict handovers; second, it considers the asymmetric links in the VCN, and adapts the geo-networking routing mechanism depending on the availability of bidirectional links; and third, it ensures the handover signalling is authenticated when a V2V path is employed to reach the infrastructure, so that possible attacks are mitigated

without affecting the performance of the ongoing sessions. Analytical evaluations and extensive simulations in OMNeT++ are carried out to demonstrate that, by employing MA-PMIP, service availability is improved for supporting seamless access to IP-based applications in the asymmetric VCN.

3. To address the continuity of IP sessions in heterogeneous urban vehicular scenarios, we have proposed a novel interworking scheme between Host Identity Protocol and Proxy Mobile IPv6. The scheme aims at supporting legacy nodes (i.e., mobile devices with no mobility support), mobility-enabled nodes, and in-vehicle mobile networks. Moreover, the scheme does not require any synchronization among the different network operators providing the access. We have provided analytical evaluations that demonstrate the improved performance of our hybrid scheme compared to other global mobility schemes. In addition, a realistic urban vehicular scenario has been simulated to evaluate the performance of our hybrid scheme, when a commuter accesses IP services during a journey that involves both pedestrian and vehicular mobilities.

1.4 Outline of this thesis

This thesis is organized as follows. Chapter 2 reviews the main existent protocols for IP mobility support in mobile networks, as well as their applicability to vehicular scenarios. It also presents a brief literature survey of important research advances for the support of IP mobility and data forwarding cooperation in vehicular environments. Chapter 3 introduces our first contribution for the support of IP services in 802.11p/WAVE networks. In Chapter 4, we present our MA-PMIP scheme for IP services provision in asymmetric vehicular networks. Chapter 5 investigates the multi-hop communications from the in-vehicle network perspective, and proposes the hybrid HIP/PMIP scheme that enables the transferring of on-going IP sessions along dissimilar access networks and different administrative domains. Finally, Chapter 6 gives conclusions of this research and outlines our future work.

Chapter 2

Background and Related Work

2.1 Host-based Mobility

2.1.1 Mobile IPv6 (MIPv6)

Mobility support in IPv6 was initially defined by the IETF in 2004, with an updated version being released in 2011 [24]. In general, MIPv6 is a host-based mobility protocol that enables the mobile node to keep using its home address (i.e., the IP address assigned in its home link), even when the node moves to a visited network. When the mobile node moves to a visited network, it configures an IP address, known as the care-of-address, through conventional IPv6 mechanisms such as stateless or stateful auto-configuration. This care-of-address allows the node to establish communications at the new location, but it also serves for establishing an association, or "binding", between the mobile node's home address and the care-of-address.

The binding is performed when the mobile node is away from the home network. At that point, the mobile node registers the newly acquired care-of-address with a router, known as the home agent, in its home network. Therefore, if a correspondent node sends packets to the mobile node, they are first routed to the mobile node's home network,

where the home agent encapsulates each packet with a new IP header, and redirects them to the visited network (i.e., the extra header indicates the mobile node's care-of-address as the new destination). Then, the mobile node decapsulates and processes the original IP packets.

An optimized version to avoid the pass through the home agent has been also defined for MIPv6. If the correspondent node supports mobility, then the mobile node can inform about the new location directly to the correspondent node. As a result, packets coming from the correspondent node are routed directly to the visited network, with no need to be sent first to the home agent. This enhancement also requires a security procedure, called Return Routability, which permits the mobile node to prove that the claimed home address and care-of-address are indeed assigned to it, preventing in this way man-in-the-middle attacks.

2.1.2 NEMO Basic Support (NEMO BS)

NEMO Basic Support [25] is an extension to Mobile IP for the support of mobile networks instead of single hosts, therefore, it has been often considered as the standard IP mobility protocol to be employed in vehicular networks. Nodes in the mobile network are served by a mobile router, and they configure IP addresses from a mobile network prefix advertised by the mobile router. When the mobile router connects to an Access Router (AR) in a visited network, it acquires a care-of-address, and establishes a tunnel with the home agent, similar to MIPv6. In this way, such tunnel is used for communications of the mobile network nodes with any correspondent node. Although the standard does not include an optimized version of NEMO BS, extensive research has been done to improve the performance of the protocol in terms of delay and throughput, because its performance may become sub-optimal in several situations [26].

To analyze the sub-optimality of NEMO BS in a vehicular network context, we look at the connection between the vehicular network and the fixed network from two different

perspectives: A) by using single-hop connections to reach the fixed network, i.e., the vehicle has direct connection to an access point in the infrastructure; and B) by using multi-hop connections to reach the fixed network, i.e., vehicles connect to neighboring vehicles in order to reach the infrastructure.

Single-hop connections between VCN and fixed network

When NEMO BS is employed as the IP mobility protocol for vehicles connected to the infrastructure, packets follow sub-optimal paths to reach the correspondent node, due to the pass through the home agent before reaching the final destination. The use of sub-optimal paths between two peers is a recurrent problem of IP mobility solutions that use intermediary agents. The vehicular scenario is not exempt of that problem either, specially if delay/throughput-sensitive applications are to be deployed. Studies show that, for a NEMO-enabled configuration, the effective throughput of TCP applications is reduced at least in half, compared to the throughput perceived by applications that do not traverse the home agent [27].

In addition, when V2V communications take place between NEMO-enabled vehicles in the same VCN, they start using paths that traverse the fixed network instead of using the direct link between them. Experiments show that this effect could increase a regular Round Trip Time (RTT) between two vehicles using 802.11b technology, from 8ms up to 40ms [28]. In general, sub-optimal paths to the correspondent node result in increased packet overhead, and longer processing and end-to-end delays. Solutions that address the aforementioned issues for single-hop connections have been classified according to the strategy they employ. They are illustrated in Fig. 2.1 and described as follows:

1) Tunnel establishment to correspondent node: This strategy resembles the route optimization technique in MIPv6. It intends to establish a tunnel directly between mobile router and correspondent node. The requirement in this case is for the correspondent node to also support NEMO BS. The approach is especially useful when the in-vehicle network

nodes communicate with only a few correspondent nodes. MIRON [27] is an example of a solution that implements this strategy. This optimization method is offered to those mobile network nodes that have no mobility protocol running on their stack of protocols.

2) Tunnel establishment to Correspondent Router: In this strategy, the access router that is serving the correspondent node caches the binding with the mobile router's information. The duties of the home agent are then shifted to the correspondent router. By assuming that traffic always traverses the correspondent router, the path mobile router–correspondent node is optimized. However, an additional procedure to locate the correspondent router becomes necessary in order to establish the optimized tunnel. ONEMO [29] is a solution based on this strategy. The mobile router discovers the correspondent router by sending a Correspondent Router Discovery Request message, to an anycast address derived from the correspondent node's IP address. Once the optimized tunnel is established, all the mobile network traffic bypasses the home agent. The solution was tested in vehicular scenarios with TCP traffic, and demonstrated improvements in throughput and a reduction of the RTT.

3) Delegation to visiting nodes: In this strategy, every mobility-enabled node (i.e., nodes in the in-vehicle network who also support MIPv6) configures a topologically-valid care-of-address directly from the infrastructure, so that it activates its own MIPv6 route optimization. The mobile router then forwards the packets coming from the mobility-enabled node to the access router, without using the bi-directional tunnel between mobile router and home agent. By surpassing both tunnels: between mobile router and its home agent, and between the mobility-enabled node and its home agent, the path to the correspondent node is optimized. However, an additional prefix delegation mechanism is required for the mobility-enabled node to be able to configure a valid care-of-address from the infrastructure. An alternative mode of operation of MIRON [27] employs this strategy. When the mobile network contains mobility-enabled nodes, they use address delegation with network access authentication to manage their own route optimization procedure in a secure manner.

4) Intra-NEMO optimization: This strategy aims at establishing a direct path between the in-vehicle network nodes and correspondent nodes, when they both are connected to the same access router. By adopting this strategy, packets can be delivered with no use of resources from the fixed network. Direct paths in the ad hoc network are typically established by a MANET routing protocol. Furthermore, there is a family of solutions — the so-called MANEMO— which explores the cooperation of MANET routing and NEMO. Solutions in [28] and [30] exemplify this strategy. Both are designed for vehicular scenarios and use Optimized Link State Routing Protocol (OLSR) to learn routes in the ad hoc network. They use a policy-based routing mechanism at the mobile router to select a NEMO-path or a MANET-path. Criteria such as bandwidth and RTT are used to select the optimal path. The test bed of both solutions involved moving vehicles. Results in [30] showed an improvement in path selection based on available bandwidth for UDP traffic. Accordingly, the experiments in [28] demonstrated a reduction of the total RTT.

Another example of intra-NEMO optimization is provided in VARON [31]. This solution aims to improve the delay and throughput for inter-vehicle communications while providing security. When the route optimization is activated, it establishes a path using the ad hoc routing protocol (ARAN), and performs a secure hop-by-hop binding procedure that employs cryptographically generated addresses. Simulation results in a vehicular environment showed that the TCP throughput of VARON does not improve for sparse scenarios, but outperforms by up to 4 times the one obtained by NEMO BS in dense scenarios.

Multi-hop connections between VCN and fixed network

After NEMO BS was released, extensive research has been conducted to evaluate and improve its performance in nested-NEMO configurations [32]. A nested-NEMO appears when the mobile router employs a multi-hop path to reach the infrastructure, so that it configures the care-of-address from the IP prefix assigned to the mobile router in the upper level. As a result, packets traverse two or more home agents before they can reach the final

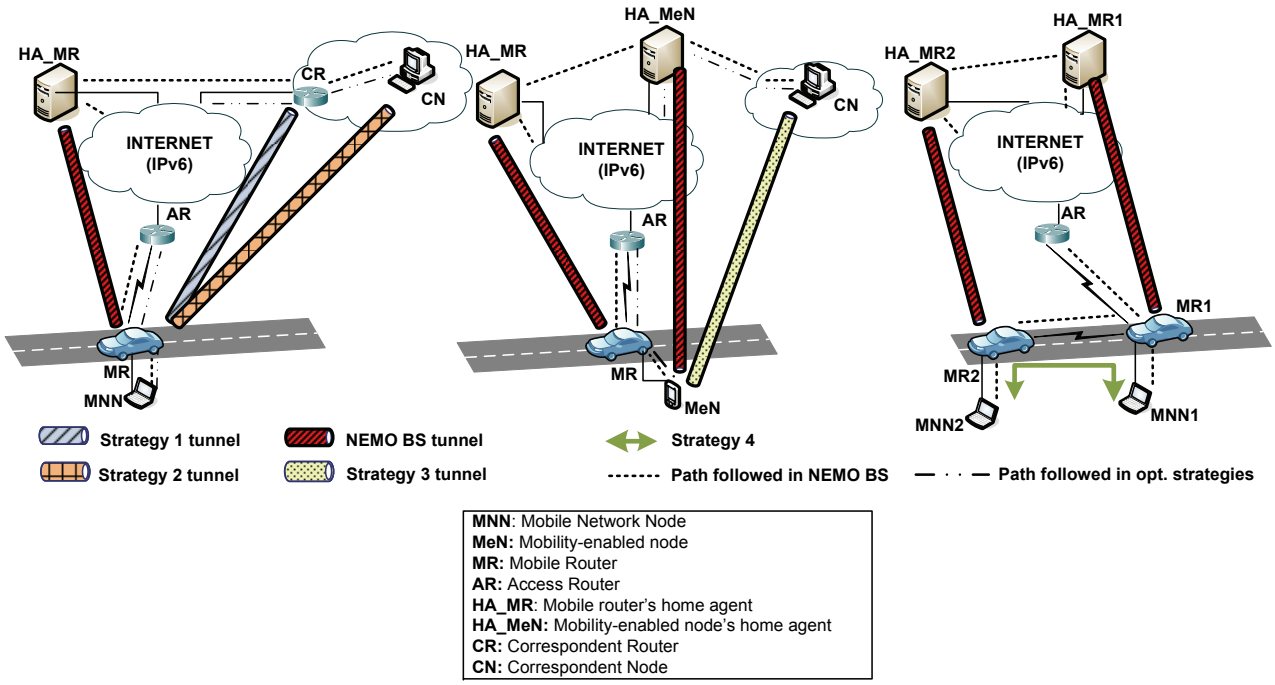


Figure 2.1: Optimization of NEMO BS in single-hop vehicular communications

destination.

However, although this thesis indeed focuses on the multi-hop connections for IP communications, we consider unfeasible the use of nested-NEMO configurations in vehicular scenarios, due to the following reasons: 1) given the short-time duration of V2V communications, the in-motion vehicles would have to constantly reconfigure the IP addresses and update their location to the home agent, every time any of the mobile routers involved in the V2V connection changes; and 2) vehicles may not share the same service provider for accessing the infrastructure, hence, there would be a natural restriction in configuring IP addresses from another vehicle's IP prefix, when they both belong to different administrative domains.

Instead of considering vehicles configuring IP addresses from other vehicles, in multi-hop scenarios we consider more feasible for the in-vehicle mobile router to employ ad-hoc routing in order to obtain IP addresses directly from the infrastructure. Therefore, if

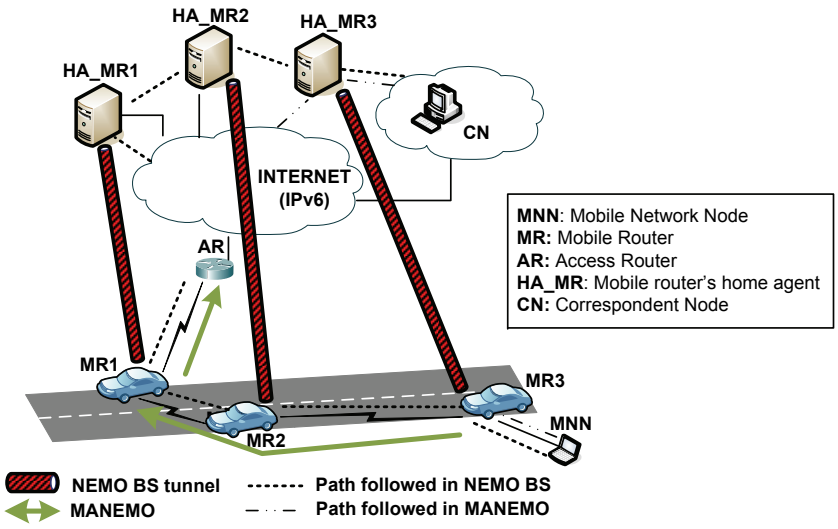


Figure 2.2: Optimization of NEMO BS in multi-hop vehicular communications

NEMO BS is employed in this type of scenario, there should be a way to guarantee that packets coming from a nested mobile router do not suffer from extra encapsulations at intermediate mobile routers. Such a strategy is illustrated in Fig. 2.2 and explained as follows.

MANEMO: If packets come from a nested mobile router and are destined to an external node, an ad hoc sub-IP routing is used to forward IP packets through the multi-hop path. In that way, MANEMO creates a virtual link between the vehicle and the access router. The packets are then forwarded from the access router to the proper home agent, and then delivered to the correspondent node. If packets are destined to nodes in the same ad hoc network, the Intra-NEMO optimization strategy described in Section 2.1.2 is employed. An implementation of this strategy, called MANET-centric solution for NEMO in vehicular scenarios, is presented in [33]. To eliminate the nesting problem, the MANET-centric scheme uses sub-IP geographic routing. Once a nested mobile router encapsulates a packet, the sub-IP layer builds a geo-header pointing to the AR. This geo-header is used to forward the packet until the AR is reached. Consequently, from the IP layer's perspective, the nested configuration is hidden, emulating a direct link between the access router and

the nested mobile router.

2.1.3 Host Identity Protocol (HIP)

HIP has been defined as an experimental standard in RFC 5201 [34]. To solve the locator/identifier problem of IP addresses, HIP uses different addresses for both location and identification. IP addresses are still used for routing protocols to reach the host. However, HIP defines a new host identity for the identification aspect. This identity is a cryptographic address by nature (i.e., a public and private key pair). Due to the introduction of a new host identity, hosts that are HIP-enabled require a new layer in the TCP/IP stack, between the network and the transport layers. Although the modification of the stack of protocols represents a drawback for the deployment of HIP, the IETF has recognized already the separation of identification and location from the IP address namespace as the next important change in the Internet architecture [35].

When two nodes want to communicate using HIP, each peer establishes a pair of Security Associations (SA), which are later used for the encryption/decryption of data packets. The SAs are related to HIP host identities and not to IP addresses. In addition, the identifiers used by application and transport layers are also related to HIP host identities. In order to make the host identity compatible with the IP address format, a hash function is employed to obtain a 128-long bit string. The latter is known as the Host Identity Tag (HIT) , which is compatible with the length of normal IPv6 addresses.

The information about a node's HIT is therefore necessary for establishing incoming communications with such a node. Thus, HIP defines a rendezvous server system (for the query of HITs given the host name), similar to the Domain Name System (for the query of IP addresses given the host name). Two queries must be performed by the initiator to obtain the necessary information before sending the first data packet: one to obtain the HIT, and one more to obtain the IP address of the node. The registration process of each host with an HIP rendezvous server is defined in RFC 5203 [36]. The rendezvous server

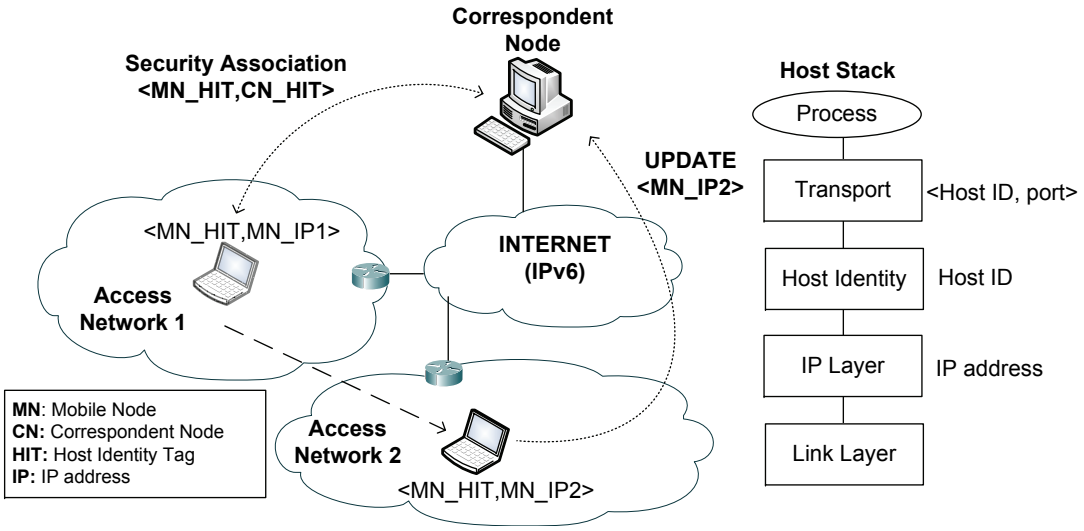


Figure 2.3: HIP Location Update for a mobile node moving from Access Network 1 to Access Network 2

also serves as the supporting mechanism for nodes that are mobile (especially when both ends of the communication are mobile) or multihomed.

In the mobile scenario, if one of the peer nodes changes its IP address, it uses a three-way signalling mechanism, named UPDATE, to inform its peer about the change, so that future packets are routed correctly to the new location of the node. The same mechanism is used for the support of multihoming. Therefore, if the IP address changes in one (or both) side(s) of the communication, HIP allows for the continuation of data packets transmission, because neither the transport layer sessions nor the SAs are related to IP addresses. The location update process of HIP is illustrated in Fig. 2.3. Consequently, HIP is considered as a host-based global mobility management protocol.

HIP has been shown to outperform other global mobility protocols such as MIPv6 [37], especially in terms of signalling overhead. Therefore, research has been conducted mostly to improve the protocol's performance in micro-mobility scenarios (i.e., mobility inside a single administrative domain, or localized mobility) rather than for macro-mobility scenarios [38,39]. In [38], the authors propose microHIP (mHIP), which defines new network entities called mHIP agents. The mHIP agents are in charge of handling signaling messages of

the intra-domain HIP handover, and to re-direct the HIP-based connections to the correct location. In [39], an extension to HIP is proposed to introduce a Local Rendezvous Server (LRVS). The LRVS is located in each administrative domain, and acts in a similar way to a network address translator: it traduces a locally-valid IP address to a globally-routable IP address. Thus, incoming and outgoing packets are intercepted by the LRVS to modify the IP addresses. Additional discussion of other studies devoted to the applicability of HIP in vehicular scenarios are presented in Chapter 5.

2.2 Network-based Mobility

2.2.1 Proxy Mobile IPv6 (PMIP)

The standard PMIP protocol is specified in RFC 5213 [40], and its general operation is mostly based on the signalling messages already defined for MIPv6. However, PMIP is a network-based mobility protocol, which means a mobile node is not required to include any mobility functionalities in its stack of protocols. On the contrary, the network-based mobility concept is that the network performs all the signalling on behalf of the mobile node, for it to maintain the reachability inside a PMIP domain, regardless of the mobile node's change of point of attachment to the network.

The PMIP operation is illustrated in Fig. 2.4. The protocol defines two network entities: a Local Mobility Anchor (LMA) and a Mobile Access Gateway (MAG) . The first one acts as the anchor point for the mobile node inside the PMIP domain. Thus, the LMA is in charge of registering the mobile node's current location, and to assign the same network prefix every time the node joins a different access network. The MAG, on the other hand, detects the connections of mobile nodes, and sends such location updates to the LMA in the form of Proxy Binding Update (PBU) messages. The LMA responds with Proxy Binding Acknowledgements (PBA) that include the IP prefix assignment for mobile nodes (Fig. 2.4-(1)). Upon PBA reception, the MAG is ready to announce the network prefix to

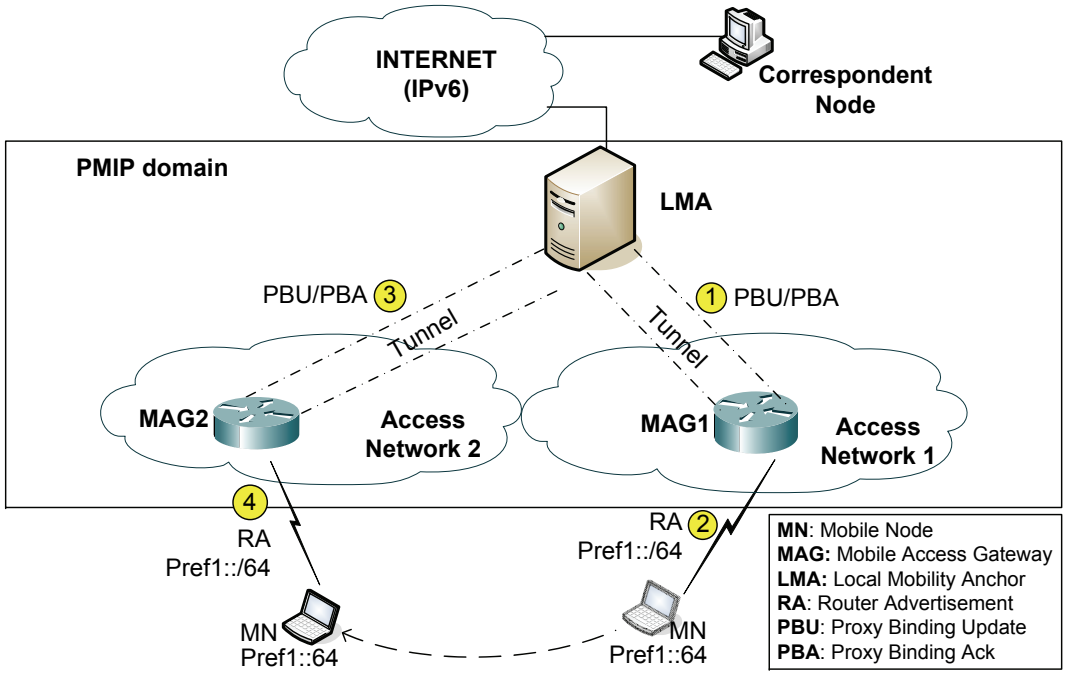


Figure 2.4: PMIP Location Update for a mobile node moving from Access Network 1 to Access Network 2

the mobile node (Fig. 2.4-(2)).

When a mobile node's handover occurs, the new MAG notifies the new connection to the LMA. Then, the LMA recognizes the mobile node by means of a unique identifier, and again assigns the same IP prefix to it (Fig. 2.4-(3)). In this way, every time the mobile node roams inside the PMIP domain, the node does not detect any changes at the network layer, and the mobility is transparent to upper layers in the stack of protocols (Fig. 2.4-(4)).

PMIP has been widely-accepted because it simplifies the stack of protocols required in mobile devices. In addition, it has been shown that, for micro-mobility scenarios, the basic operation of PMIP may outperform the enhanced version of MIPv6, called Fast MIPv6, by showing more robustness against control messages dropping than Fast MIPv6 in reactive mode [41]. In the case of predictive mode, the study in [41] concluded that the basic PMIP does not show any improvements compared to Fast MIPv6. However, PMIP has been enhanced with a recently released standard for Fast Handovers in PMIP [42], which also

enables predictive handovers for this network-based protocol.

On the other hand, the Evolved Packet System (EPS) architecture, which covers the radio access, the core network, and the terminals in the so-called all-IP Long Term Evolution (LTE) network , has indicated that PMIP is the network-based mobility protocol to be employed over non-3GPP and 3GPP accesses (for the latter, EPS also indicates the traditional GTP protocol for network-based mobility) [43]. Since EPS provides the interworking between 3GPP technologies and non-3GPP radio access networks, the support of IP mobility across the IP-based core networks becomes key for the provision of all services.

Due to the aforementioned arguments, and since previous studies of NEMO BS in vehicular environments have confirmed its performance limitations [26,44], PMIP has been also suggested as the mobility protocol for vehicular communications networks. However, the standard PMIP only supports mobility for single hosts, so it requires modifications in order to provide mobility for the in-vehicle mobile network. Lee *et al.* [11] introduce P-NEMO as a way to support session connectivity for hosts traveling attached to a mobile router in Intelligent Transport Systems. In a similar way, Bernardos *et al.* [13] present an extension for PMIP to support mobile networks, by allowing hosts to obtain and maintain connections from both fixed and mobile routers.

In this thesis, we have selected PMIP as the main IP mobility protocol to be employed in our multi-hop vehicular communications networks. In Chapters 3 and 4, we point out which characteristics of the standard PMIP are employed for our network model, and introduce the modifications required for making it usable in multi-hop VCN. Then, in Chapter 5, we extend the mobility support beyond the PMIP domain, in order to enable global IP mobility for the in-vehicle network, as well as for commuters and pedestrians. More details about additional studies of PMIP in vehicular environments are further discussed and compared in the following chapters.

2.3 Data forwarding cooperation in VCN

Vehicular communications networks are envisioned to support a wide variety of infrastructure-based infotainment applications. Accordingly, extensive research has been conducted to improve the performance of these applications through one-hop communications [45]. Nevertheless, considering the race between the always-increasing access demand and the deployment of the supporting infrastructure, applications' availability has been extended through multi-hop connections in the vehicular ad hoc network (VANET).

As a result, data forwarding cooperation in the VCN has been proposed at the PHY/MAC layer [18, 21, 46–48], and at the network layer [33, 49], among others. However, when intermediate nodes are involved in data forwarding of external traffic, it is reasonable to assume that selfish nodes may appear with the goal of simultaneously maximize their benefits and minimize their contribution [50]. As a result, extensive research efforts have been devoted to propose incentive mechanisms and reputation systems that enforce cooperation among nodes. In [50], the authors determine the conditions under which cooperation without incentives can exist, while taking the network topology into account. In [51], Salem et al. propose an incentive mechanism based on a charging/rewarding scheme to make collaboration rational for selfish nodes in multi-hop cellular networks. In [19], the authors provide a payment model for a micropayment system that stimulate nodes' cooperation in multi-hop wireless networks. Similarly, Chen et al. propose a scheme to stimulate message forwarding in VANETs based on coalitional game theory [23].

Although the study of incentive mechanisms and rewarding schemes for data forwarding cooperation in VCN is outside of the scope of this thesis, the non-exhaustive list of works we have provided help envisage the advances made in these topics.

Chapter 3

VIP-WAVE: A Framework for IP Mobility in 802.11p/WAVE Networks

3.1 Preliminaries

In this chapter, we propose the **V**ehicular **IP** in **W**AVE **(VIP-WAVE)** framework. VIP-WAVE is our first contribution for the support of IP communications over multi-hop paths in the VANET domain. The proposed framework is customized to work with the technologies specially designed for vehicular communications, i.e., the IEEE 802.11p/WAVE standards.

Although traditional radio access networks such as cellular (e.g., GSM/GPRS and UMTS) and WiFi may also be employed to enable vehicular communications [7, 8, 52], the strict latency requirement for safety-oriented and emergency communications has resulted in the definition of the IEEE 802.11p/WAVE standards [4–6]. Such standards define a low-latency alternative network for vehicular communications, and their main focus has been the effective, secure, and timely delivery of safety-related information.

However, the deployment of infotainment applications certainly would help to accelerate

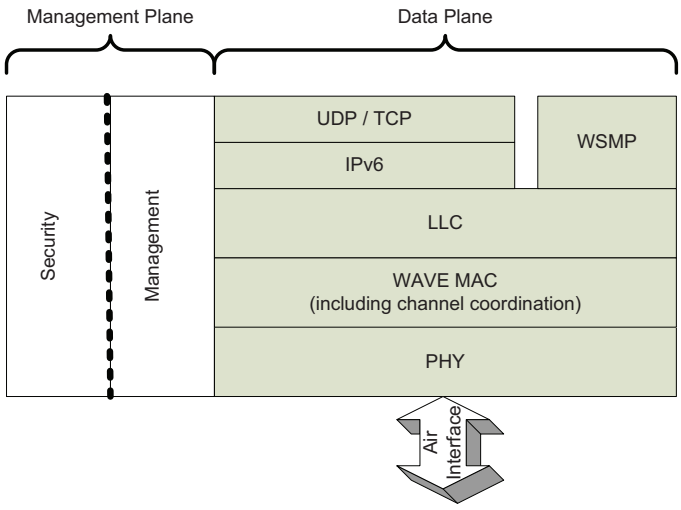


Figure 3.1: WAVE stack of protocols as defined in IEEE 1609.3-2010 [5]

the market penetration and leverage the deployment costs of the infrastructure required by WAVE. Thus, in order to support infotainment traffic, the standards also consider IPv6 data packets transmission, and transport protocols such as TCP and UDP. By supporting IP-based communications, the vehicular network may use well-known IP-based technologies and readily be connected to other IP-based networks.

Fig. 3.1 shows the WAVE stack of protocols. The IEEE 1609.3 standard [5] specifies two network layer data services: WAVE Short Message Protocol (WSMP), which has been optimized for low latency communications, and IPv6. Although the operation of WSMP has been fully specified in [5], it has been found that recommendations for the operation of IPv6 over WAVE are rather minimal [53]. Protocols in which the operation of IPv6 relies for addressing configuration and IP-to-link-layer address translation (e.g., the Neighbor Discovery protocol) are not recommended in the standard.

Additionally, IPv6 works under certain assumptions for the link model that do not necessarily hold in WAVE. For instance, IPv6 assumes symmetry in the connectivity among neighboring interfaces. However, interference and different levels of transmission power may cause unidirectional links to appear in WAVE, which may severely affect IPv6's ef-

fectiveness in its operation. Furthermore, interference and mobility may cause inability to communicate with other WAVE devices unless relayed communications are employed. For example, there are cases in which the Road Side Unit (RSU) (i.e., the point of attachment to the infrastructure) has to deliver configuration information for IPv6 to a vehicle through a multi-hop path. However, the multi-hop support of infrastructure-based IP services is not currently permitted in the IEEE 1609.3 standard.

With many open operational aspects of IPv6, providing access to infrastructure-based IP applications, such as assisted parking, route management, and eventually Internet access, becomes a challenging task in 802.11p/WAVE networks. Previous works evaluate the performance of IP-based applications in I2V vehicular environments, but they often employ traditional 802.11b/g technologies that do not resemble the intricacies of 802.11p/WAVE for IP communications. In [53], the limitations of the operation of IPv6 in 802.11p/WAVE have also been identified, but they can only be used as guidelines regarding the incompatibilities of the two technologies.

Therefore, in this chapter we address the problem of I2V/V2I IP-based communications in 802.11p/WAVE networks by providing the VIP-WAVE framework. Since our general focus in this thesis is the multi-hop IP mobility aspect, in this part of our work we require to step back and first propose solutions for the open issues found in the 802.11p/WAVE standards for the IPv6 support over one-hop connections (i.e., the traditional type of access for V2I). Then, we extend our proposal so that IP-based mobile communications are also supported over two-hop connections for vehicles that employ the WAVE technologies for V2V and V2I communications. Our design goals are the following:

- To design an efficient mechanism for the assignment, maintenance, and duplicate detection of IPv6 global addresses in WAVE devices, which is customized according to the type of user service;

- To support the per-application and on-demand IP mobility for seamless infrastructure-based communications;

- To design a relay detection and routing mechanism for the delivery of IP packets through one-hop and two-hop communications in 802.11p/WAVE networks.

Furthermore, we aim at developing an analytical model for evaluating and comparing the throughput performance of the standard WAVE and the proposed VIP-WAVE. The model integrates the vehicle's mobility, and considers the delays due to handoff, the packet collisions due to MAC layer conditions, and the connectivity probability of vehicles to the infrastructure according to the channel model.

In the following sections, we first discuss the 802.11p/WAVE standards and review the previous works (Section 3.2). Next, we describe our network model, and introduce the VIP-WAVE framework and its extensions for the support of multi-hop communications (Section 3.3). Finally, the proposed analytical model and the performance evaluation of the proposed framework are presented (Sections 3.4 and 3.5, respectively).

3.2 Related Work

In this section, we present the main concepts described in the 802.11p/WAVE standard that are relevant for the transmission of data frames and the operation of IP-based services. We also describe previous works dedicated to the support of IP-based communications in 802.11p/WAVE networks.

3.2.1 The 802.11p/WAVE Standards

At the physical (PHY) and medium access control (MAC) layers, the 802.11p technology for wireless communications while in a vehicular environment has been proposed in [4]. The 802.11p works in the 5.9 GHz frequency band, and employs Orthogonal Frequency Division Multiplexing (OFDM) modulation. It defines CSMA/CA as the fundamental access method to the wireless media. The MAC layer of 802.11p includes the 802.11e

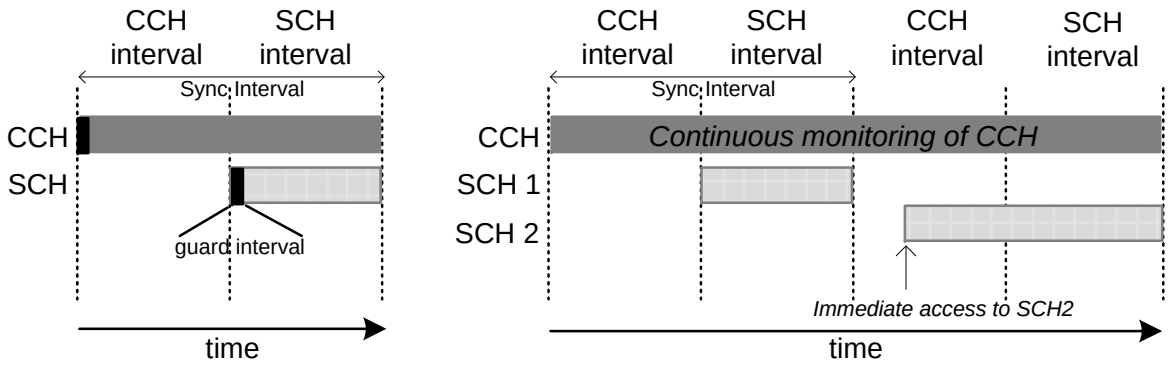
Enhanced Distributed Channel Access (EDCA) function to manage access categories and priorities.

The previous versions of the 802.11p draft allowed for the exchange of data packets between two WAVE entities only if they were inside the context of a WAVE Basic Service Set (WBSS). However, this condition has been removed in the latest version, and a WAVE device is now able to "consume" services from a provider by simply switching the radio to the proper channel where the service is being offered. In this way, the latency associated with establishing a WBSS can be avoided. Out-of-context WBSS is then the recommended transmission mode for data plane services in WAVE [6].

On the other hand, the Wireless Access in Vehicular Environments (WAVE) standards, namely 1609.4-2010 [6] and 1609.3-2010 [5], define the medium-access channel capabilities for multi-channel operation, and the management and data delivery services between WAVE devices. In [6], WAVE frequency spectrum is divided into 1 control channel (CCH) and 6 service channels (SCH), each with 10MHz bandwidth. In addition, each channel has a set of access categories and its own instance of the 802.11p MAC layer.

Among the different types of frames that can be exchanged in WAVE, management frames can be transmitted in both CCH or SCH. Conversely, data frames (i.e., WSMP and IPv6 data frames) should be transmitted in SCH, although WSMP frames are also allowed in the CCH. Furthermore, the 802.11p radios may have a single-physical layer (single-PHY) or multi-physical layer (multi-PHY). The former means the radio is able to exchange information in one single channel at all times; therefore, a single-PHY has to continuously switch between CCH and SCHs every certain time (the default is 50ms). The latter indicates the radio is able to monitor the CCH while at the same time it can exchange data in one or more SCHs. Examples of single-PHY and multi-PHY radios accessing the channels are illustrated in Fig. 3.2

The 1609.3-2010 standard for networking services provides more details regarding the support of IP communications [5]. It specifies as mandatory the support of IPv6 link-local, global, and multicast addresses in WAVE devices. Regarding the IP configuration,



(a) Single-PHY channel access **(b)** Multi-PHY channel access

Figure 3.2: Multi-channel synchronization in WAVE

it indicates that link-local addresses should be derived locally and WAVE devices should accept traffic directed to well-known IPv6 multicast addresses (e.g., all-nodes multicast address). It also states that "WAVE devices may implement any Internet Engineering Task Force (IETF) protocol"; however, it does not specify the operation conditions for the Neighbor Discovery for IPv6 protocol (ND) [54].

According to [5], the announcement of IP services takes place in the Wave Service Advertisement (WSA) management frame. The WAVE device announcing the service takes the role of "provider", whereas the one receiving the WSA and indicating interest in the service takes the role of "user". Each WSA includes 0 to 32 `ServiceInfo` segments, 0 to 32 `ChannelInfo` segments, and up to one `WaveRoutingAdvertisement (WRA)` segment. A `ServiceInfo` includes among others, the definition of the service, the provider information (including its IP address if it is an IP service), the Received Channel Power Indicator (RCPI) level (dBm) recommended to accept the service (also known as the RCPI threshold), and the index for the `ChannelInfo` segment in the WSA that corresponds to the announced service. A `ChannelInfo` includes among others, the service transmission characteristics (e.g., Tx power and data rate), the channel number, and the type of access in the SCH (i.e., continuous access or alternating access between SCH and CCH).

Similarly, if the WSA has at least one `ServiceInfo` segment for an IP Service, it should

also include a `WRA` for global IPv6 addressing configuration and internetwork connectivity. A `WRA` segment includes the IP prefix, prefix length, default gateway, DNS, and router lifetime, among other extension fields relevant for IP configuration at the WAVE user's side. Once the WAVE user receives a WSA with an announced IP service of its interest, it calculates a global IP address by means of stateless configuration, based on the IP prefix received in the `WRA` segment and its own MAC address, after which the WAVE user is ready to start consuming the service. `WRA`s are meant to replace the standard Neighbor Discovery protocol, as a mean to minimize the overhead and latency associated with the latter.

From the described operation of IP services in 802.11p/WAVE networks, one can identify the following limitations:

Lack of duplicate address detection mechanism. Given the broadcast nature of WSA messages for the announcement of services, a WAVE user interested in a specific IP service is allocated with the same IP prefix of all other users subscribing to any other IP service announced in the same WSA. On the one hand, that forces nodes to perform some kind of duplicate address detection (DAD) procedure, to guarantee the uniqueness of IP addresses among all users. The need for DAD comes mainly from the fact that WAVE devices may support readdressing to provide pseudonymity. Therefore, a MAC address may be changed at any moment and be randomly generated, which would increase the chances of collisions for auto-configured IP addresses based on MAC addresses. Nonetheless, as we mentioned before, the ND operation, which includes the standard DAD procedure for IPv6, is not recommended in WAVE.

On the other hand, suppose the infrastructure provides Internet access or route management services. These are examples of extended IP services that are provided through the entire 802.11p/WAVE network, and are continuously announced by all the RSUs. Thus, even if a WAVE device actually performs a DAD and confirms the uniqueness of its IP address among other neighboring users, the DAD will be invalidated as soon as the vehicle moves to the area of coverage of a different RSU, since the set of neighbors will also

change. Furthermore, the DAD will be invalidated when the WAVE user switches to a different SCH to consume another service for which the same `WRA` has been employed.

Lack of seamless communications for extended services. Suppose the DAD problem is alleviated by having each RSU to advertise a unique set of IP prefixes among all the other RSUs. Then, if the DAD is executed among neighboring users, the IP address uniqueness may be guaranteed at the RSU service area level. Although this solution would work for non-extended services, it would cause a breakage for extended services continuity, because when a user moves its connection to a different RSU, it receives a different IP configuration information. Therefore, transport layer sessions have to be reset and service disruption will be experienced as a result of the reconfiguration.

Lack of support for multi-hop communications. The current standard allows for a WAVE user to consume infrastructure-based IP services only if there is a direct connection between RSU (i.e., WAVE provider) and WAVE user. We consider such condition as an undesired limitation of the 802.11p/WAVE standards. Vehicular networks experience highly variable channel conditions due to mobility, obstacles, and interference. Therefore, it is desirable to take advantage of intermediary WAVE devices to relay packets from/to the infrastructure. In this way, access to the IP services could be extended to further than one-hop WAVE users, when there are some WAVE users that do not directly hear the RSU. In addition, service could be provided to users that do hear the RSU but with a signal quality level below the one recommended by the RCPI threshold.

Extensive research has shown that mobile networks may benefit from multi-hop communications, in terms of improving the network capacity and throughput [55]. Also, by serving as relays, nodes may obtain benefits from the network, like earning credits that reward them for their relay services [19]. Moreover, other standards for vehicular communications have already considered the support of IPv6 multi-hop communications by means

of sub-IP geo-routing [56].

3.2.2 Previous Works

IP becomes a natural solution for providing addressing services in WAVE, and for enabling the access to existent IP networks (e.g., the Internet), to legacy applications, and to innovative services. Therefore, the IP addressing configuration in vehicular networks has been further investigated in numerous studies [57–59]. While these studies enable IP configuration in moving vehicles, they are often limited to guarantee uniqueness in a specific area (e.g., around the leading vehicle acting as DHCP server [57], around the service area of RSU [58], or around a specific lane [59]). As a result, they limit the deployment of extended IP services and seamless communications in 802.11p/WAVE. We address this limitation by designing an IP addressing scheme for 802.11p/WAVE that employs a differentiated treatment for location-dependant and extended services, in a way that it does not overload the network, at the same time that it guarantees uniqueness throughout the entire network.

In terms of mobility management, host mobility solutions for vehicular networks, based on the Network Mobility (NEMO) Basic Support Protocol, are proposed and evaluated in [33, 60–63]. Baldessari et al. [33] define a MANET-centric solution that exploits multi-hop communications, so that each vehicle is treated as a NEMO Mobile Router. Prakash et al. [61] propose a vehicle-assisted cross-layer handover scheme for vehicles to help relaying signalling and data packets of a handover vehicle. In [62], on the other hand, vehicular clusters are employed so that cluster-heads are in charge of IP mobility for other vehicles. Different from the aforementioned works, network-based mobility with Proxy Mobile IPv6 has been proposed in [11, 13]. Soto et al. [13] enable mobility for broadband Internet access to be provided in a transparent way in automotive scenarios, whereas Lee et al. [11] propose a set of network mobility support protocols for Intelligent Transport Systems.

In general, those schemes reduce the handover delay and improve the throughput in vehicular networks. However, none of them specifically consider the use of 802.11p for

V2I communications. Instead, they employ a general 802.11 network for connectivity to the infrastructure, or theoretical performance evaluations. In our work, we select network-based mobility since it confines the signalling overhead at the infrastructure side, and it does not require mobility management protocols to be included in the stack of the vehicle's On Board Unit (OBU) (Fig. 3.1). Furthermore, we adapt the signalling and movement detection mechanisms required for mobility management, in a way that the control channel of the 802.11p/WAVE network do not suffer from excessive overhead or congestion. Thus, we propose a customized mobility management mechanism tailored to the characteristics of 802.11p/WAVE networks.

Our premise of extending the network coverage in areas with different levels of infrastructure presence leads to a proposal for multi-hop communications. Employing intermediate nodes to extend the network service area and to improve performance has been previously investigated in the context of vehicular networks [19, 55, 56]. We take the advantage from the findings of these works and further define the relaying services in 802.11p/WAVE, by considering the many service channels of this network, the different levels of the availability of neighboring vehicles as relays, and the restrictions imposed over the control channels to carry data that may interfere with the delivery of emergency and safety information.

On the other hand, a collection of works are devoted to provide measurement studies for evaluating the performance of IP-based applications in V2I vehicular environments [7–9]. However, they employ traditional 802.11b/g technologies and obviate the limitations existent in the current 802.11p/WAVE standard for IP communications. In [64], they do provide an evaluation of UDP/TCP applications in 802.11p/WAVE, but their main focus is to reduce the problem of bandwidth wastage resulting from the switching operation in single-PHY environments. In parallel to those measurement studies, extensive research has been devoted to provide theoretical models for evaluating mobility and spatiotemporal relations, connectivity and access probabilities, MAC layer performance, handovers, and relay strategies in vehicular environments (see [45, 60, 65–68] and references therein). Al-

though we have been inspired by these works, our work is different in that we integrate these many aspects to provide a closed-form expression, from a microscopic point of view, for the throughput evaluation of IP applications in the 802.11p/network.

3.3 The Vehicular IP in WAVE (VIP-WAVE) Framework

3.3.1 Network Model

Consider the infrastructure-based vehicular network shown in Fig. 4.2. The connection to the infrastructure is provided by RSUs located along the road. Vehicles are equipped with On Board Units (OBU) that enable connections to the infrastructure and to other vehicles. Every RSU and OBU is equipped with 802.11p/WAVE radios. It is assumed that RSUs and OBUs are multi-PHY. In this way, we alleviate problems such as bandwidth wastage, longer queuing, and higher end-to-end delay, which have been previously identified as the consequences of the channel switching operation performed by 802.11p single-PHY radios [69] [70].

Two different infrastructure-based IP services are offered in the 802.11p/WAVE network: 1) extended services that are continuously announced by all RSUs in the network, such as mapping applications, route planning, and Internet access; and 2) non-extended services that are location-dependant, such as assisted-parking, and that are provided only by some RSUs.

For a given channel model $\mathcal{C}$, vehicles may establish a direct connection to the RSU. Some other vehicles, however, are located in areas uncovered by the infrastructure (see car A in Fig. 4.2), or with a communication link in deep fade toward the RSU (see car B in Fig. 4.2). Inside such areas, we exploit the use of multi-hop communications, so that at most one intermediate vehicle acts as a relay for another vehicle's communications from/to

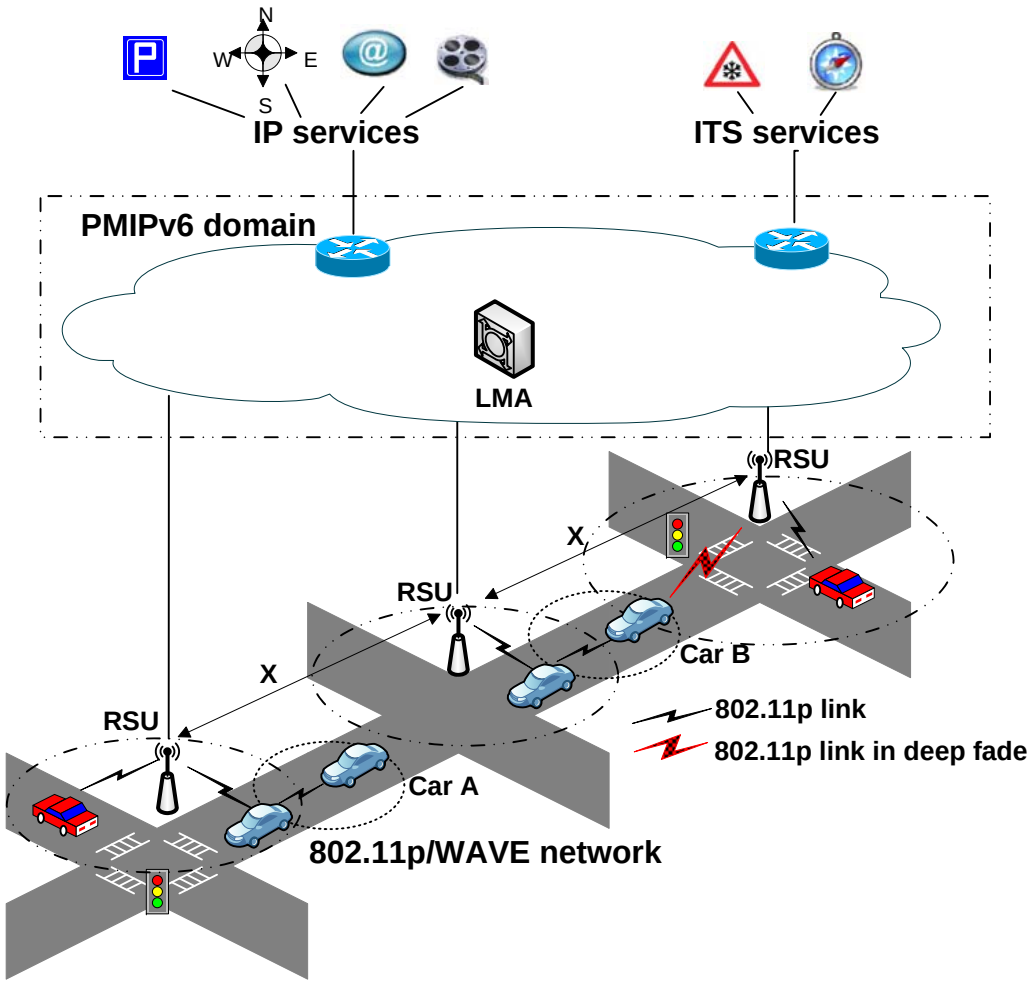


Figure 3.3: IP-enabled 802.11p/WAVE network model

the RSU [66]. Since the transmission power of RSU is higher than the transmission power of OBU, this leads to the RSU radio range, R, to be wider than the OBU radio range, r.

Furthermore, in the case of extended services, we have selected the standard Proxy Mobile IPv6 (PMIP) protocol, introduced in section 2.2.1, to manage the IP mobility of the OBUs. The general integration of PMIP with the 802.11p/WAVE network has been illustrated in Fig. 4.2. When a MAG detects a new connection, it sends a PBU to the LMA on behalf of the MN. The LMA then assigns an IP prefix and creates a tunnel through which all traffic from/to the mobile node is encapsulated toward the serving MAG. When

the mobile node changes its location, the LMA has to change the tunnel's end-point upon reception of a PBU from the new serving MAG. We also consider the whole 802.11p/WAVE network as a single PMIP domain, and co-locate the MAG functionalities with the RSU.

3.3.2 VIP-WAVE Architecture

As denoted in section 3.2.1, one of the 802.11p/WAVE biggest issues, in terms of IP operation, is the announcement of a per-WSA IP prefix, which forces WAVE users of all IP services announced in a specific WSA (up to 32 services per WSA) to belong to the same IP network. This causes not only a necessity for often having to detect duplicate addresses through out the network and other SCHs, but also contradicts one of the main assumptions IPv6 has for the link-layer model, which says that all nodes belonging to the same IP prefix are able to communicate directly with each other. This assumption does not hold when there are WAVE users that are scattered along different locations or along different service channels.

Additionally, there is a shortage in differentiating extended from non-extended services, and no IP mobility support is indicated to provide seamless communications in the case of extended services. Last but no least, multi-hop communications are not exploited in the 802.11p/WAVE network, although they could boost the network's performance and increase the IP services availability.

The general idea behind our framework is to address those limitations by integrating IP configuration and IP mobility in order to provide differentiated treatment for extended and non-extended services. We intend to enable a per-user IP prefix for the access to extended services an for guaranteeing seamless communications. Moreover, we intend to improve the coverage of IP services by extending the access to OBUs located two hops away from the RSU.

The architecture of VIP-WAVE is illustrated in Fig. 3.4. VIP-WAVE is located in the data plane of the WAVE stack of protocols and it defines three main components

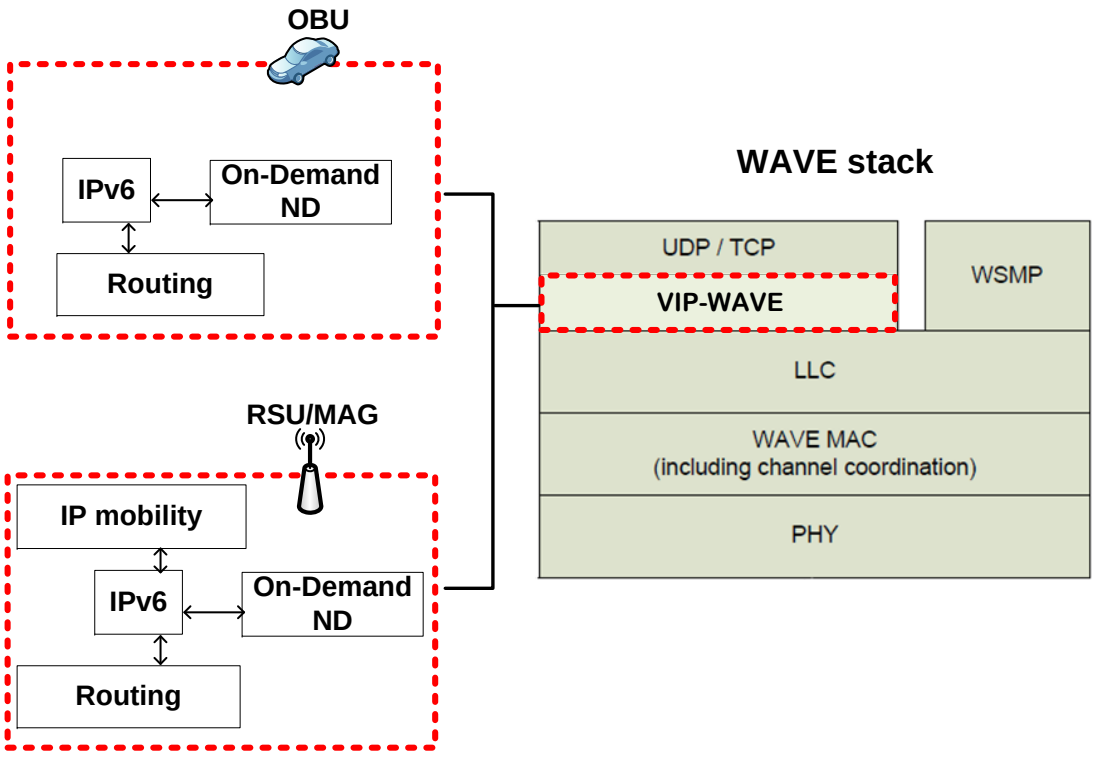


Figure 3.4: Vehicular IP in WAVE (VIP-WAVE) architecture

that interact with the standard IPv6 protocol: 1) the IP addressing and mobility block (only in the RSU), in charge of assigning global IPv6 prefixes to vehicles and guaranteeing IP mobility for extended services throughout the network; 2) the on-demand Neighbor Discovery block, which is a light-weight adaptation of the standard ND; and 3) the routing block, which enables relay selection for multi-hop communications when a user fails to directly consume the IP service from the RSU. Due to our selection of PMIP for network-based mobility, the OBUs do not have to include any component for IP mobility, as depicted in Fig. 3.4.

In the following sections, we describe the interaction of VIP-WAVE's components for the support of IP services to vehicles directly connected (i.e., one-hop away) to the infrastructure, and then we introduce the extensions required for enabling support of two-hop connections in VIP-WAVE.

IP service establishment

The RSU that announces an IP service includes, besides the type of service (i.e., extended or non-extended), the global IP address of the hosting server, its own MAC address that identifies it as the WAVE provider, and the RCPI threshold (i.e., the recommended minimum WSA's received power level). Such information is included in the extension fields of `ServiceInfo`, as specified in [5]. The WSA is transmitted in the CCH. Since OBU has a radio dedicated to monitor the CCH, all one-hop users in the area of service of the RSU can receive the WSA. Upon WSA reception by a potential WAVE user, the user determines if it wants to access the service and checks the type of service to proceed in the following way:

1. **If service is extended:** The OBU tunes a radio to the SCH specified in the `ChannelInfo` segment. At that point, the OBU does not have a global IP address to initiate communications with the hosting server; therefore, the IPv6 module requests the on-demand ND module to trigger a Router Solicitation (RS) message. The RS message is destined to the all-routers multicast address as indicated in [54], and is handed to the routing module for determining the L2 next-hop destination. Since the user is directly connected to the RSU, the routing module selects the WAVE provider's MAC address (i.e., RSU MAC address) as the MAC-layer frame destination; thus, instead of multicast, the RS is delivered as a unicast message. The RSU then exchanges PBU/PBA messages with the LMA for IP prefix assignment, after which the RSU sends a unicast Router Advertisement (RA) message to the OBU.

   The RA message includes all the information required by IPv6 for a proper configuration. Once the global IP address has been calculated, the OBU may start exchanging IP data packets with the hosting server. Note that no DAD mechanism is required after IP address configuration, since the IP address uniqueness is guaranteed by having an IP prefix uniquely assigned to each OBU;

2. **If service is non-extended:** The OBU employs the IP prefix announced in the

WRA to calculate a global IP address. After IP configuration, the OBU tunes to the proper SCH. Since the IP prefix is shared among other users consuming non-extended IP services announced in the same WSA, a DAD procedure has to be executed before the OBU may start transferring IP packets. Hence, our on-demand ND defines a centralized DAD mechanism controlled by the RSU, which is only triggered when the first IP data transmission request appears at the OBU. The RSU keeps a list of the active OBUs and their IP addresses, in order to be able to detect duplicates.

The details of the DAD procedure are depicted in Fig. 3.5. Note that the OBU's IP configuration for non-extended services is only valid inside the area of coverage of the serving RSU; thus, the IP uniqueness only needs to be guaranteed at the serving RSU level, instead of at the entire network level. Once the DAD has been completed, the OBU may start exchanging IP data packets with the hosting server.

Handover of IP services

An OBU transitions through the service area of different RSUs at vehicular speeds. Therefore, we introduce a handover mechanism that allows for seamless communications of extended IP services in the 802.11p/WAVE network. When an OBU is consuming a extended service, it continues monitoring the CCH while is roaming toward a new RSU. Consequently, the reception of a WSA that announces the same extended service but from a different WAVE provider (i.e., the WAVE provider field in `ServiceInfo` now includes a different MAC address), serves as a movement detection hint that is notified by the MAC layer to the VIP-WAVE layer in the OBU. Upon the movement notification, the on-demand ND module triggers the sending of an RS message, which is transmitted over the SCH in which the service is being offered.

The reception of the RS message is then employed by the RSU for connection detection, so that it proceeds to exchange PBU/PBA signalling with the LMA. The RSU sends an RA to the recently detected OBU as a response to the RS message. As a result, the LMA

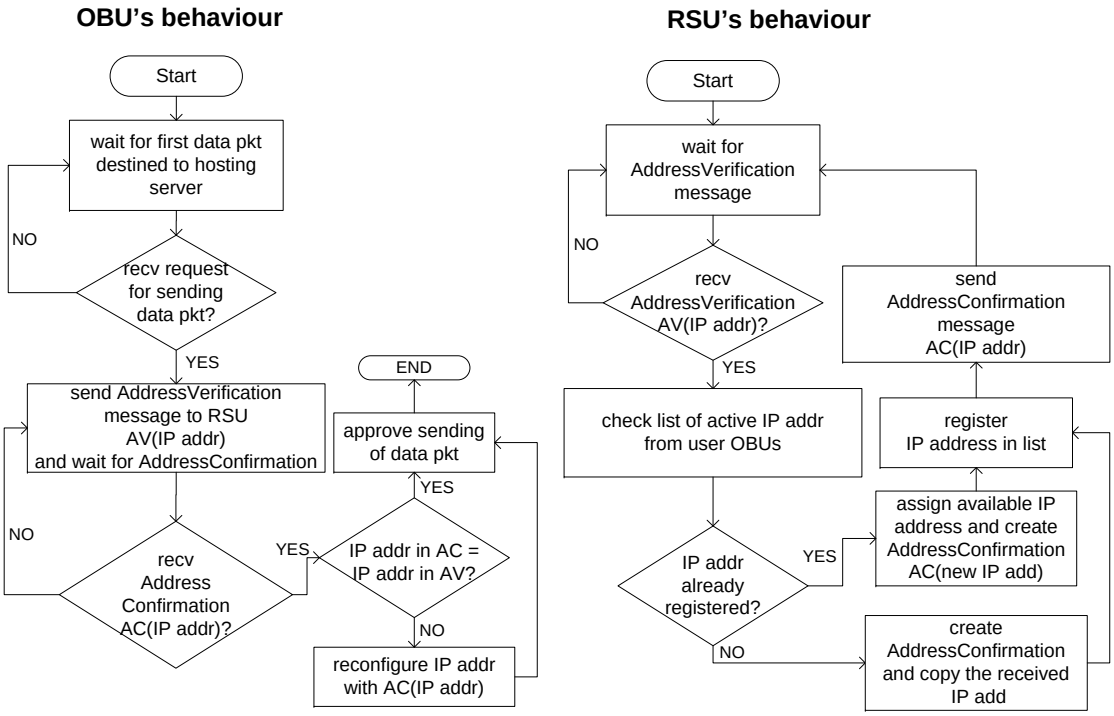


Figure 3.5: DAD mechanism in VIP-WAVE for non-extended services

is able to resume packets forwarding toward the OBU as soon as it sends the PBA to the new RSU. The OBU, on the other hand, is able to resume packets transmission toward the hosting server once it receives the RA.

Note that our on-demand ND does not require the frequent sending of messages. We have replaced the necessity of receiving frequent RA messages by the reception of WSAs that are already defined in the standard. Thus, an IP prefix does not expire, unless announcements for the service that is currently being consumed are no longer received. In this way, the WSA message reception aids the VIP-WAVE layer in two ways: 1) it helps the maintenance of IP addresses by replacing the non-solicited RA messages defined in the standard ND; and 2) it solves the IP-to-link-layer address translation, because the WSA already includes the MAC address of the current WAVE provider. In addition, we alleviate possible congestion in the CCH by having the on-demand ND messages (e.g., RS or RA) being transmitted only over the SCH.

For the non-extended services case, they are no longer available when the OBU moves to a new service area, thus, they do not require the definition of a handover mechanism.

3.3.3 VIP-WAVE Extensions for Two-hop Scenarios

In section 3.2.1, we have introduced the advantages of enabling multi-hop communications in vehicular networks. Therefore, we define the necessary features and services to extend the support of VIP-WAVE in two-hop scenarios. We start by defining two services that are closely related: 1) the *Relay Service*, which is registered in the `ProviderServiceRequestTable` of all OBUs and is announced only when they require another OBU to serve as a relay. A request for relay service may only be sent after the user OBU has started consuming a given service (i.e., after the OBU has acquired its IP configuration from the RSU); and 2) the *Relay Maintenance*, which is announced by the intermediary OBU that has been selected as a relay for IP communications.

Intermediate OBUs may serve as relays for extended and non-extended services. However, only those OBUs with availability to serve as temporary relays will take action when they receive a *Relay Service* request. The procedure for setting up a relay OBU is located in the routing module and described in detail as follows.

Routing through a relay

Table 3.1 presents the algorithm for setting up communications through a relay, and an example of a successful relay establishment is illustrated in Fig. 3.6. Once the procedure has been completed in all the involved parties, the RSU and user OBU have the necessary information for delivering packets through a two-hop path, so that the exchange of IP packets may be resumed.

Procedure at user OBU

Relay detection

1: **if** (WSAs from the RSU are no longer received **or** received WSA signal < RCPI threshold)

2: create a `ServiceInfo` to announce the Relay Service solicitation.

3: include this OBU's ID and IP address in the extension fields of `ServiceInfo`.

4: associate a `ChannelInfo` segment with SCH number of active IP service.

5: send a WSA with the Relay Service announcement in CCH.

6: **else**

7: keep using one-hop connection to RSU.

Relay setup

25: **if** (Relay Service announcement has been sent **and** WSA with Relay Maintenance announcement is received)

26: set the relay OBU's MAC address as next-hop for reaching the RSU.

Procedure at relay OBU

Relay provision

8: **if** (reception of WSA with Relay Service solicitation **and** availability to serve as relay)

9: tune to the SCH of the service as indicated in `ChannelInfo`.

10: create a Relay Notification message.

11: include the user OBU's information in the Relay Notification message.

12: set Relay Notification's destination address to `ALL_ROUTERS`.

13: send Relay Notification message through SCH.

Relay setup

21: **if** (reception of Relay Confirmation)

22: create a `ServiceInfo` to announce the Relay Maintenance to the user OBU.

23: send a WSA with the Relay Maintenance announcement in CCH.

24: set the forwarding route for packets from/to the user OBU.

Procedure at RSU

Relay setup

14: **if** (reception of Relay Notification **and** user OBU's information corresponds to an active user OBU)

15: create a Relay Confirmation message.

16: set relay OBU's MAC address in Relay Confirmation's MAC frame destination.

17: send Relay Confirmation message to relay OBU in SCH.

18: set the relay OBU's MAC address as next-hop for reaching user OBU.

19: **else if** Relay Confirmation has been already sent

20: discard Relay Notification.

Table 3.1: Relay setup procedure in VIP-WAVE

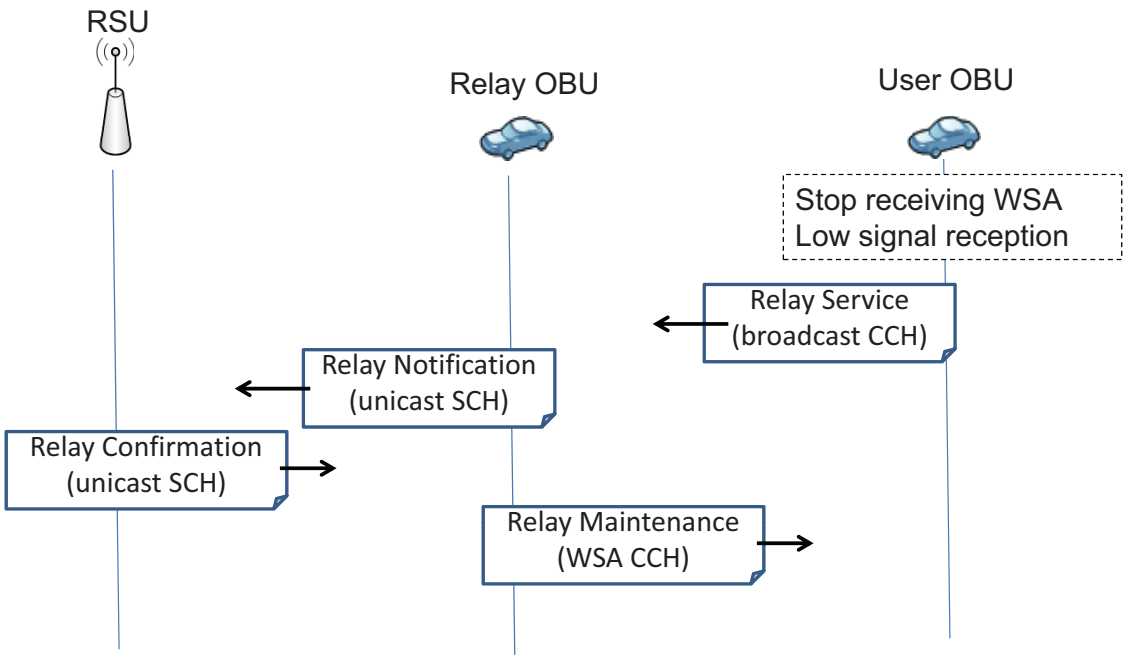


Figure 3.6: Example of successful relay establishment according to Algorithm 3.1

Depending on the direction of traffic, the routing protocol works in the following way for multi-hop communications:

1. **Traffic from hosting server to user OBU:** Once the packet arrives at the RSU, the IPv6 protocol queries the routing module about the next-hop to reach the user OBU. The routing module selects the relay OBU MAC address as the MAC layer frame destination, as per configured by the relay setup procedure. The packet is then forwarded to relay OBU.

2. **Traffic from user OBU to hosting server:** Once the data packet is generated at the user OBU, the IP layer determines that hosting server belongs to an external network; thus, it decides the packet should be sent toward the default gateway, which in this case is the RSU. The IPv6 module then queries the routing module about the next-hop to reach the RSU. As configured by the relay setup procedure, the route to reach the RSU indicates the relay OBU as the next-hop; therefore, the relay OBU MAC address is selected as the MAC layer frame destination. The packet is then forwarded to relay OBU.

If at any moment during the two-hop communications, the user OBU detects again the reception of WSA directly from the RSU in the CCH, and with a signal level above the RCPI threshold, then the user OBU sends a Router Solicitation to re-establish direct communications with the RSU. The Router Advertisement response message sent by the RSU is overheard and employed by the relay OBU for terminating the relay service.

Handover in two-hop scenarios

When the vehicle is in motion, it may experience handovers of communications in different scenarios: 1) it may move the connection to a relay OBU, where both relay and user OBUs remain in the service area of the same RSU; and 2) it may move the connection to a relay OBU, where the relay OBU is connected to an RSU different from the user OBU's serving RSU. The first case holds for extended and non-extended services, whereas the second case only holds for extended services. Note that the handover procedure when the vehicle maintains a direct connection to the RSU has been already defined in section 3.3.2.

1. **Handover to a relay in the same service area:** In this scenario, the signalling required to maintain seamless communications is no different from that described in Table 3.1. Since both relay OBU and user OBU remain in the service area of the same RSU, when the RSU receives the Relay Notification message (step 14), it finds the information about the user OBU registered in its list of active IP users. Therefore, it does not require to trigger any signalling for IP mobility. Moreover, the procedure is the same regardless of whether the service is extended or non-extended.

2. **Handover to a relay in a different service area:** The procedure of two-hop handover to a different service area is illustrated in Fig. 3.7. In this scenario, the handover may be triggered by the conditions described in Table 3.1 (step 1), so the relay detection procedure is started. However, given that the relay is connected to a different service area, when the RSU receives the Relay Notification message (step 14), it does not have an active tunnel configured for the user OBU. Therefore, the RSU uses the Relay Notification message as a hint for connection detection, and triggers the PBU/PBA signalling toward the LMA. Once the PMIP signalling is completed, the RSU continues with the sending of Relay Confirmation (step 15) to the relay OBU. This message serves for triggering the *Relay Maintenance* announcements from relay OBU to user OBU (step 23), after which bi-directional communications are resumed.

3.4 Analytical Model

We derive an analytical model to evaluate the performance of the proposed VIP-WAVE framework compared with the standard network layer in WAVE. The analysis focuses on modelling the OBU's mobility, and calculating the handover delay and packet collision probability. Based on those aspects, we examine a randomly tagged vehicle and calculate

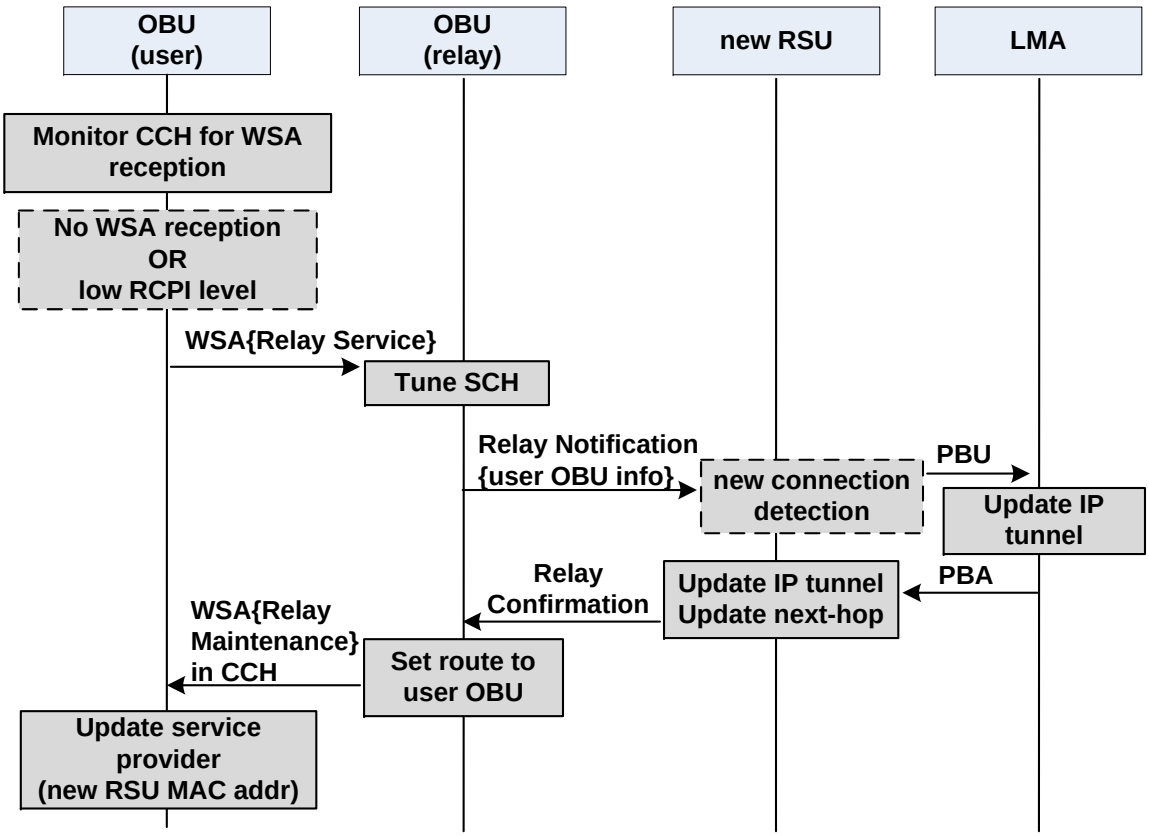


Figure 3.7: Handover of extended IP services through a relay in VIP-WAVE

its nodal downstream throughput experienced when it has an active extended IP service in the 802.11p/WAVE network.

3.4.1 Mobility Model

Consider the network model illustrated in Fig. 4.2. To make our analysis tractable, assume the RSUs are uniformly deployed along the roads and separated by a distance X. Similar to [66], we analyze the subnetwork placed in the range $[0, X]$ and bounded by two consecutive RSUs. We further divide such subnetwork in smaller segments $\mathbb{S} = \{1, 2, 3, ..N\}$, where each $s \in \mathbb{S}$ is of length d_s. A vehicle that moves along the 802.11p/WAVE network, iteratively transits through the segments while traversing the different subnetworks.

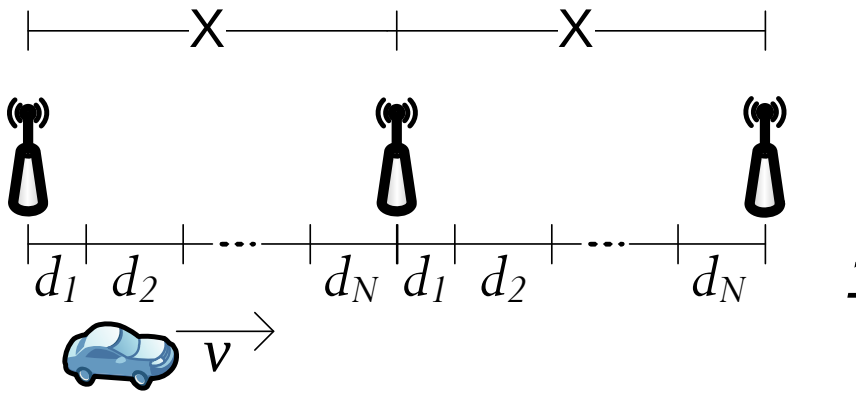


Figure 3.8: Spacial division of the 802.11p/WAVE network to model a vehicle's mobility

The relation between the spacial division of the 802.11p/WAVE network and the vehicle's mobility are shown in Fig. 3.8. The residence times in each segment are considered to be geometrically distributed with mean t_s. The mean residence time in each segment is determined by $t_s = d_s/v$, where v is the average velocity of the vehicle.

On the other hand, for a user OBU subscribed to an IP service, its connection to the RSU may be of three different types: direct connection (i.e., one-hop), connection through a relay (i.e., two-hop), and no connection at all. In [66], for a vehicle located at x in $[0, X]$, $p_1(x)$ denotes the probability of the vehicle to be directly connected either to RSU in 0 or RSU in X, and $p_2(x)$ denotes the probability of the vehicle to be connected to at least one relay (where a relay is any other vehicle with direct connection to the infrastructure). These access probabilities are defined as:

$$p_1(x) = 1 - (1 - g_b^{\mathcal{C}}(x))(1 - g_b^{\mathcal{C}}(X - x)) \tag{3.1}$$

$$p_2(x) = 1 - e^{-\int_0^X g_v^{\mathcal{C}}(\|x-y\|)\rho p_1(y)dy} \tag{3.2}$$

where $g_b^{\mathcal{C}}(x)$ and $g_b^{\mathcal{C}}(X-x)$ are the V2I connectivity probabilities for a given channel model $\mathcal{C}$ and a given location with respect to both RSUs (i.e., at x for RSU in 0 and at $X-x$ for RSU in X). $g_v^{\mathcal{C}}(\|x-y\|)$ is the V2V connectivity probability between two vehicles located at x and y respectively, and ρ represents the density in vehicles per meter (vpm).

The number of vehicles in $[0, X]$ is assumed to be Poisson distributed with mean ρX. Despite of the fact that our model relies on the assumption of a Poisson distributed population of vehicles, it has been previously demonstrated, by means of validation with real world traffic traces and synthetic mobility models [71, 72], that it is a reasonable assumption that does not detract the adequacy of our model, instead, it helps to make our analysis tractable. Although a Poisson distribution is commonly employed for sparse vehicular ad hoc networks, the results in [72] show that, for all traffic densities, the exponential distribution accurately estimates the inter-vehicle spacing distribution, especially for a spacing larger than 50m.

Accordingly, we represent the connection type of a vehicle in segment $s \in \mathbb{S}$ as $G_s = \{1, 2, 0\}$ for one-hop, two-hop, and no connection, respectively. Thus, for a vehicle located in segment s, the probability distribution of G_s can be calculated as:

$$P\{G_s = a\} = \begin{cases} p_1(\omega_s), & \text{if } a = 1, \\ (1 - p_1(\omega_s))p_2(\omega_s), & \text{if } a = 2, \\ (1 - p_1(\omega_s))(1 - p_2(\omega_s)), & \text{otherwise.} \end{cases} \tag{3.3}$$

For the simplicity of the analysis, we use the middle point of the segment to represent the location of vehicles in that segment. Thus, we denote by ω_s the location of the middle point of segment s. The connection type of the user OBU is therefore integrated with our mobility model, so that $P\{G_s = a\}$ represents the probability of the user OBU of having connection type a to the RSU while the vehicle is in segment $s \in \mathbb{S}$.

3.4.2 Handover Delay

Definition 1. *The handover delay H_G is the time duration between the breakage of the user OBU's connection to the infrastructure (i.e., through direct or relayed connection) and the resumption of data packets transmission from the infrastructure to the user OBU. The handover delay varies according to the type of connection G acquired by the user OBU in the new location.*

Handover delay in standard WAVE

In the standard WAVE, the vehicle may experience only two states: directly connected to RSU or disconnected. We define two possible configurations for the standard WAVE. In scenario A, we consider the current standard as-is with no mobility management scheme. Then, it is reasonable to assume that each RSU includes a different IP prefix in its `WRA`, as mentioned in section 3.2.1. In such case, the vehicle should reset its connection for an extended IP service every time it enters the service area of a new RSU. The handover delay of this scenario (WV-A) is calculated as follows:

$$H_{G=1}^{\text{WV-A}} = R_{\text{WSA}} + RESET \tag{3.4}$$

where R_{WSA} indicates the time delay for the user OBU to receive a WSA from the new RSU, and $RESET$ corresponds to the time for a user OBU's transmission of a connection reset toward the server, and its corresponding re-configuration time above the network layer (e.g., the 3-way TCP handshake).

In scenario B, we consider the standard 802.11p/WAVE network to be PMIP-enabled, so that network-based mobility is offered to maintain the IP prefix assignment of the OBUs along the domain. Although this configuration is not mentioned in the standard, by considering this scenario we account for basic IP mobility management employed in

the standard WAVE, at the same time that we provide a fair comparison to our proposed VIP-WAVE framework. Note that this scenario would require a basic Neighbor Discovery signalling in order to re-establish the flow of IP traffic at the new location. The handover delay of this scenario (WV-B) is derived as follows:

$$H_{G=1}^{\text{WV-B}} = R_{\text{WSA}} + T_{\text{RS}} + RTT_{\text{PMIP}} + R_{\text{RA}} \tag{3.5}$$

where T_{RS} indicates the transmission time for RS message, RTT_{PMIP} indicates the round trip time for exchanging PBU/PBA messages between MAG and LMA, and R_{RA} indicates the time delay for the user OBU to receive the RA message from the infrastructure.

Handover delay in VIP-WAVE

In VIP-WAVE, a roaming vehicle may experience different types of connection breakages. When in a one-hop connection, the vehicle may lose signal reception due to distance, blocking of line of sight, or poor signal quality reception. On the other hand, besides the aforementioned causes of connection breakage, in a two-hop connection the OBU may also terminate its current two-hop connection when it again detects a one-hop connection with better link quality conditions.

Among all those possibilities, we analyze the worst-case scenario, in which every time the vehicle experiences a change of connection (i.e., to one-hop or two-hop), it involves also a change of RSU, hence it triggers PMIP signalling at the infrastructure side. Although this may not be the case for real deployments, because the OBU may change its type of connection and still be connected to the same RSU, the assumption allows us to give an upper bound estimation of the handover delay induced by the proposed VIP-WAVE framework.

The handover delay in VIP-WAVE is calculated as follows:

$$H_{G=1}^{\text{VIP}} = R_{\text{WSA}} + T_{\text{RS}} + RTT_{\text{PMIP}} + R_{\text{RA}} \tag{3.6}$$

$$H_{G=2}^{\text{VIP}} = T_{\text{R.SOL}} + T_{\text{R.NOT}} + RTT_{\text{PMIP}} + T_{\text{R.CONF}} + R_{\text{R.MAIN}} \tag{3.7}$$

Note that $H_{G=2}$ does not require waiting for WSA reception, as the relay selection and configuration process starts as soon as the user OBU stops receiving WSAs from the RSU (or when the RCPI threshold is no longer met). The calculation involves the transmission and reception delays for R.SOL, R.NOT, R.CONF and R.MAIN (i.e., the messages defined in Table 3.1 for selecting and setting the relayed connection).

3.4.3 Packet Collision Probability

Definition 2. *The packet collision probability p_{col} is the probability of packet losses due to collisions occurring between two or more nodes transmitting at the same time when they are all tuned to the same SCH.*

Packet collision probability in standard WAVE

Let M_s denote the mean population of vehicles in segment s, $s \in \mathbb{S}$. Then, M_s can be expressed by:

$$M_s = \rho d_s \tag{3.8}$$

where ρ is the density of vehicles (vpm) and d_s is length of segment s (m). Let us consider P_α as the probability that an OBU subscribed to service α is active (i.e., the OBU is tuned to the SCH where service α is being provided and is transmitting/receiving data packets). Then, the conditional transmission probability $\tau_1(s)$ given that a vehicle is located in segment s is given by:

$$\tau_1(s) = P\{G_s = 1\}P_\alpha \tag{3.9}$$

where $P\{G_s = 1\}$ is the one-hop connectivity of vehicles in segment s.

For the standard WAVE, we denote by $p_{col}^{\mathrm{WV}}(s)$ the conditional collision probability of a tagged node in segment s, given that the tagged node is active. Thus,

$$p_{col}^{\mathrm{WV}}(s) = 1 - (1 - \tau_1(s))^{M_s - 1} \prod_{s' \in S_r(s), s' \neq s} (1 - \tau_1(s'))^{M_{s'}} \tag{3.10}$$

where S_r denotes the set of segments that fall into the radio range of the tagged vehicle. For the simplicity of the analysis, if the middle point of the segment falls into the radio range of the tagged vehicle, that segment is considered in S_r. Therefore, we have,

$$S_r(s) = \{s' | \omega_s - r < \omega_{s'} < \omega_s + r\} \tag{3.11}$$

Packet collision probability in VIP-WAVE

In VIP-WAVE, a vehicle communicates with the RSU either directly or through two-hop relaying. Then, the conditional transmission probability $\tau_2(s)$ given that a vehicle is located in segment s is given by:

$$\tau_2(s) = (P\{G_s = 1\} + P\{G_s = 2\})P_\alpha. \tag{3.12}$$

Recall that $P\{G_s = 2\}$ is the two-hop connectivity of vehicles in segment s. For VIP-WAVE, we denote by $p_{col}^{\mathrm{VIP}}(s)$ the conditional collision probability of a tagged node in segment s, given that the tagged node is active. Thus,

$$p_{col}^{\text{VIP}}(s) = \begin{cases} 1-(1-\tau_2(s))^{M_s-1} \prod\limits_{s'\in S_r(s), s'\neq s} (1-\tau_2(s'))^{M_{s'}}, & \text{if } G_s = 1, \\ 1-(1-\tau_2(s))^{M_s-1} \prod\limits_{s'\in S'_r(s), s'\neq s} (1-\tau_2(s'))^{M_{s'}}, & \text{if } G_s = 2 \end{cases} \tag{3.13}$$

where $S_r(s)$ is given by (3.11) and $S'_r(s)$ is given by

$$S'_r(s) = \{s' | \omega_s - 2r < \omega_{s'} < \omega_s + 2r\}. \tag{3.14}$$

S'_r indicates that for guaranteeing the transmission of the tagged vehicle, vehicles within the two-hop range of the tagged vehicle should be inactive.

3.4.4 Nodal Downstream Throughput

Definition 3. *The nodal downstream throughput T is the average rate of packets received at the user OBU when traversing the subnetwork in $[0, X]$. It is expressed in bits per seconds.*

Let B denote the total number of bits received by an individual OBU when traversing the subnetwork in $[0, X]$. According to our mobility model, d_s/v is the average time the vehicle spends in each segment s. Consequently, the expected number of bits received while in segment s, $E[B_s]$, and the total number of bits B received in $[0, X]$, are computed as follows:

$$E[B_s] = \sum_{a=0}^{2} B_s P\{G_s = a\} \tag{3.15}$$

$$B = \sum_{s=1}^{N} E[B_s]. \tag{3.16}$$

The average nodal downstream throughput T experienced by the tagged vehicle is then expressed as:

$$T = \frac{B}{(\sum_{s=1}^{N} d_s)/v}. \tag{3.17}$$

Nodal downstream throughput in standard WAVE

According to the previous definition of B_s, we express the number of bits received in state s, B_s^{WV} as follows:

$$B_s^{\mathrm{WV}} = \begin{cases} \lambda_d(1 - p_{\mathrm{col}}^{\mathrm{WV}}(s))(d_s/v - H_{G_s}^{\mathrm{WV}}), & \text{if } G_s = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.18}$$

where λ_d is the downstream data rate (in bits per second) from the IP server to the OBU, and H_G^{WV} is given by either (3.4) or (3.5). Overall, the expression computes the total number of bits received during the available transmission time (i.e., after deducting the handover delay), while the OBU is in segment s. Note that an OBU operating under the standard WAVE does not receive data packets when $G_s = 2$ or $G_s = 0$.

Nodal downstream throughput in VIP-WAVE

The number of bits B_s^{VIP} received in state s while the OBU operates under VIP-WAVE is defined as:

$$B_s^{\mathrm{VIP}} = \begin{cases} \lambda_d(1 - p_{\mathrm{col}}^{\mathrm{VIP}}(s))(\frac{d_s}{v} - H_{G_s}^{\mathrm{VIP}}), & \text{if } G_s = 1, \\ \lambda_d(1 - p_{\mathrm{col}}^{\mathrm{VIP}}(s))(\frac{d_s}{2v} - H_{G_s}^{\mathrm{VIP}}), & \text{if } G_s = 2, \\ 0, & \text{if } G_s = 0. \end{cases} \tag{3.19}$$

Note that for $G_s = 2$ the effective time available for transmission is considered to be roughly $\frac{d_s}{2v} - H_{G_s}^{\text{VIP}}$, since the packets go through an intermediary node before they can be forwarded to the user OBU.

3.5 Performance Evaluation

For evaluation purposes, we compare our VIP-WAVE framework with the standard WAVE with no mobility management (WAVE-A), and with the standard WAVE with IP mobility provided by Proxy Mobile IPv6 (WAVE-B). In WAVE-A, each RSU announces a different prefix in the `WRA` segment, which forces vehicles to reset communications for extended IP services every time they roam to a different RSU. In WAVE-B, we assume the network supports Proxy Mobile IPV6, as previously mentioned in Section 3.4.2. The comparisons evaluate the nodal downstream throughput for variable network characteristics, and the delay due to handovers and during data packets delivery.

3.5.1 Model Validation

We obtain the numerical results for our analytical model in Matlab. The average nodal downstream throughput in standard WAVE is obtained by replacing (3.18) in (3.15), and by calculating B^{WV} and T^{WV} according to (3.16) and (3.17), respectively. The average nodal downstream throughput in VIP-WAVE is obtained by replacing (3.19) in (3.15), and by calculating B^{VIP} and T^{VIP} according to (3.16) and (3.17), respectively.

The settings for such evaluation are provided in Table 3.2. In order to obtain $P\{G_s\}$, we calculate $p_1(\omega_s)$ by assuming a unit disk model $\mathcal{U}$, so that connectivity is determined mainly by the distance between vehicle and RSU. However, we also integrate the RCPI threshold (see Section 3.2.1) in determining connectivity, because a received power level below the RCPI threshold results in a disconnection from the vehicle to the provider RSU. Thus, we calculate the V2I connectivity probability as:

$$\text{(unidirectional)} \quad g_b^{\mathcal{U}}(\omega_s) = \begin{cases} 1, & \text{if } (\omega_s \leq R) \text{ and } (rxPw \geq RCPI), \\ 0, & \text{otherwise} \end{cases} \tag{3.20}$$

where $rxPw$ is the OBU's reception power level calculated as $rxPw = 10log_{10}(\text{Tx Power RSU}) - PL$, which is a reduction of the log-normal shadowing model to the unit disk model when the path loss component, PL, has no fading.

We also consider a more restrictive bidirectional connectivity probability. This is to account for the asymmetry existent in the transmission power of RSUs and OBUs, in which case a distance $\omega_s \leq R$ only guarantees connection from RSU to OBU, but not from OBU to RSU. Thus, to guarantee bidirectionality we have:

$$\text{(bidirectional)} \quad g_b^{\mathcal{U}}(\omega_s) = \begin{cases} 1 & \text{if} \quad (\omega_s \leq r) \text{ and } (rxPw \geq RCPI)] \\ 0 & \text{otherwise.} \end{cases} \tag{3.21}$$

In other words, the unidirectional connectivity probability given by (3.20) allows for one-way reception of traffic from RSU to OBU, but it does not necessarily guarantees reception from OBU to RSU. Examples of such IP-based applications that require one-way reception of traffic are audio and video streaming. In the case of bidirectional connectivity probability, as calculated by 3.21, two-way reception of traffic is enabled between RSU and OBU when they meet the connectivity conditions. Examples of IP-based applications that require two-way reception of traffic are IP telephony and general TCP-based applications.

For the calculation of $p_2(\omega_s)$, we modify the integral limits in (3.2) to calculate the average number of nodes in $[\omega_s - r, \omega_s + r]$, and consider only a percentage of that number, given by the parameter p_r, as available to serve as relays (i.e., OBUs that process *Relay Service* requests and are available at the time of reception of a request).

Table 3.2: VIP-WAVE Performance Evaluation Parameters

Parameter	Value
Tx Power RSU	50mw (500m radio range)
Tx Power OBU	11mW (250m radio range)
Frequency band	5.9GHz
Link data rate	6Mbps
PHY/MAC Layer	Inetmanet 802.11p / CSMA-CA
RCPI threshold	-85dBm
Download data rate (λ_d)	100Kbps (default) $\sim$ 3Mbps
Available relays (p_r)	40%, 70%, 100%
Average speed (v)	35Km/h (default) $\sim$ 100Km/h
Density (ρ)	1/25vpm
RSUs inter-distance (X)	500m $\sim$ 2000m
Segment length (d_s)	50m
R_{WSA}	25ms
$RESET$	150ms
$T_{\text{RS}}, R_{\text{RA}}$	5ms
$T_{\text{R.SOL}}, T_{\text{R.NOT}}, T_{\text{R.CONF}}, R_{\text{R.MAIN}}$	5ms
RTT_{PMIP}	10ms
P_α	1% $\sim$ 12%
Session time	600s

In order to validate our model, the numerical results are compared with the simulation results in the next section.

3.5.2 Simulation Results

Extensive simulation results have been obtained based on the discrete event simulator OMNeT++ [73] and the Inetmanet framework [74]. The simulation parameters are presented in Table 3.2, and a simulated sample topology is depicted in Fig. 3.9. RSUs and OBUs are equipped with two wireless interfaces transmitting in different channels. In this way, we emulate the multi-PHY capabilities with simultaneous transmissions over CCH and SCH. Each radio implements the Inetmanet 802.11p PHY and MAC model, and parameters are set according to the recommended values in [4]. Connectivity among nodes is initially determined by a unit disk model. However, signals are attenuated following a Log-normal propagation model with path-loss exponent of 2.4. We have also modified the Inetmanet package so that it delivers the OBU's received power to the network layer; in this way,

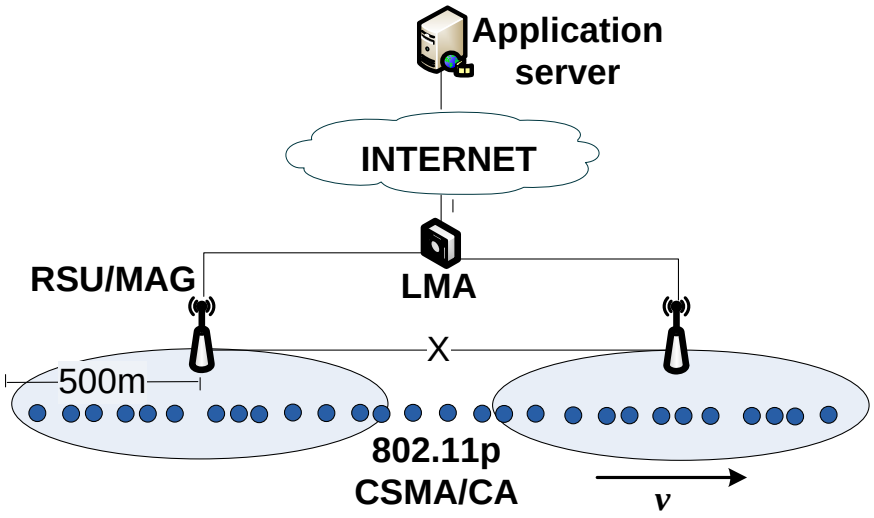


Figure 3.9: Simulation setup in Omnet++

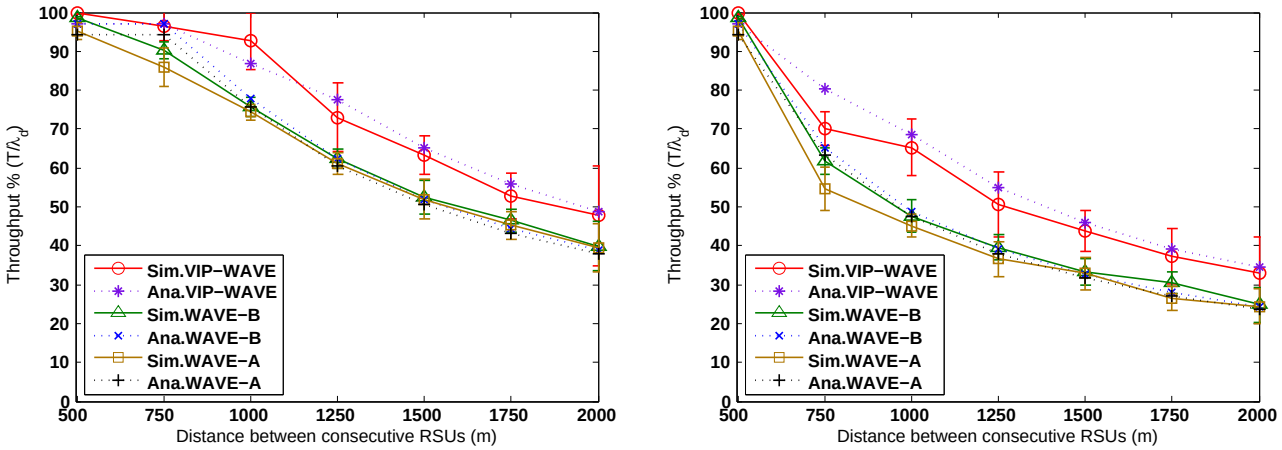
we employ the RCPI threshold to determine connectivity between OBU and RSU. An Internet-located application server for the downloading of data traffic is connected to the 802.11p/WAVE network with an RTT of 40ms.

RSUs are uniformly deployed along the road segment with a distance X. We employ one-way lane where vehicles are moving at a constant average velocity v. Each topology employed during the experiments has an average number of vehicles of ρX per subnetwork in $[0, X]$, with vehicles randomly located along the road and following an exponentially distributed inter-distance. Only application layer packets, sent from the application server and received at the user OBU, are considered for the throughput calculation in each simulation run. The results are plotted with a 95% confidence interval.

Level of presence of infrastructure

Fig. 3.10 shows the throughput obtained when X increases from 500m to 2000m. The analytical results are verified by the simulation results in both one-way and two-way application traffic scenarios.

Furthermore, as shown in Fig. 3.10a, the performance of VIP-WAVE outperforms the



(a) One-way traffic **(b)** Two-way traffic

Figure 3.10: Nodal downstream throughput for different levels of presence of infrastructure, average speed $v = 35$Km/h, constant density $\rho = 1/25$ vpm, and $p_r = 0.4$

standard one even when the same IP mobility protocol is employed. It is also observed how the effective throughput drops for all, as soon as $X > 2R$. This is due to the existence of uncovered areas between consecutive RSUs; in the case of VIP-WAVE, the greater X is, the more the vehicle depends on the density ρ for being able to find a two-hop connection toward an RSU, as it is shown later in Fig. 3.13. Furthermore, it is observed that is more probable for vehicles to find a two-hop connection to the RSU when $X < 2R+r$. However, this condition only benefits the VIP-WAVE scheme, as neither WAVE-A nor WAVE-B support multi-hop communications. On the other hand, in Fig. 3.10b can be seen how the reduced coverage observed by two-way traffic applications results in a steeper decrease in throughput. Due to a shorter connectivity range, the effective throughput starts decreasing as soon as $X > 2r$.

Impact of velocity and available relays

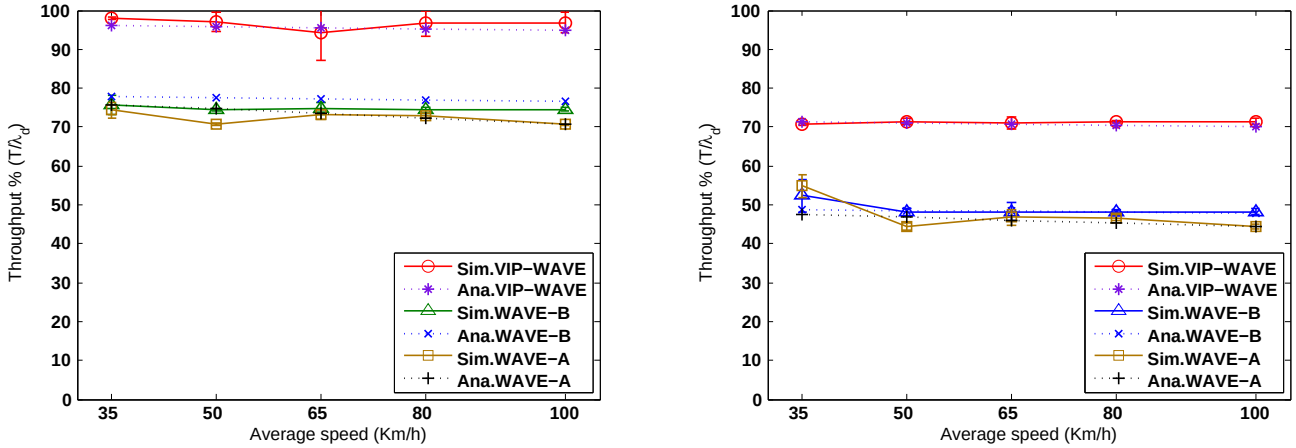
The impact on throughput performance given different values of v is illustrated in Fig. 3.11. Once more, the numerical results are shown to be accurate when compared to simulation

results. It is observed that both VIP-WAVE and standard WAVE are stable for different average speeds. This can be explained as the result of the reduced handover signalling overhead, thanks to the on-demand Neighbor Discovery coupled with the use of WSA messages for movement detection and relay establishment. Thus, despite of the handovers occurring at a higher rate with the increase of velocity, it does not reflect on a higher number of packet losses at higher speeds.

With regard to the type of traffic, in Fig. 3.11b, we observe nearly a 30% reduction of successful reception of packets when the IP application requires bidirectional connection. However, the extended area of coverage provided by the relay-aided communications in VIP-WAVE demonstrate its benefit: it improves the effective throughput by nearly 20% compared to the standard WAVE. Consequently, we also evaluate the impact of the available number of OBUs, p_r, willing to serve as relays in VIP-WAVE. The results of these experiments are depicted in Fig. 3.12. They indicate that even for a low availability of 40%, the difference in the effective throughput is minimum, i.e., VIP-WAVE only requires one neighboring OBU to be available (and connected to the RSU) to take advantage of two-hop connections in uncovered areas.

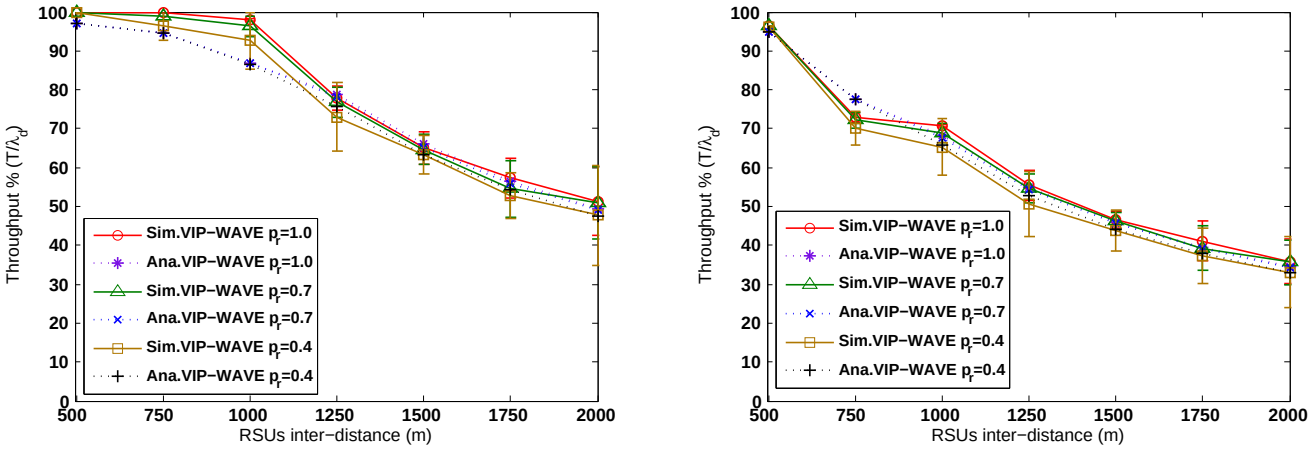
Impact of vehicle density

Fig. 3.13 depicts the analytical throughput given different densities in a low-level presence of infrastructure (i.e., X=1500m). It can be observed the trends of throughput in terms of the vehicle density when the percentage of available relays decreases from 100% to 70% and 40%. For both WAVE-A and WAVE-B, the throughput decreases almost linearly when the vehicle density increases, regardless of the values of p_r. This is the result of an increase in congestion when there are more nodes in the vehicular network. Instead, since VIP-WAVE supports multi-hop communications, a greater p_r value directly translates into an increase of throughput, and a better performance than that obtained by the standard WAVE in all three cases. However, it can also be observed that VIP-WAVE's throughput increases up to a maximal value, but thereafter it starts decreasing with the increase of



(a) One-way traffic **(b)** Two-way traffic

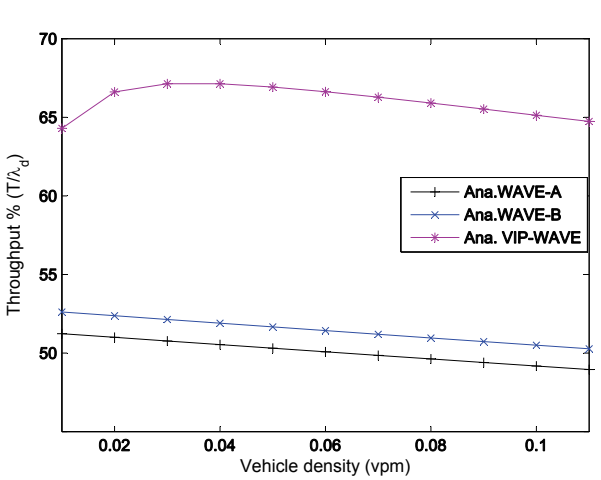
Figure 3.11: Nodal downstream throughput for different average speeds, RSUs inter-distance $X = 1000$m, and constant density $\rho = 1/25$vpm



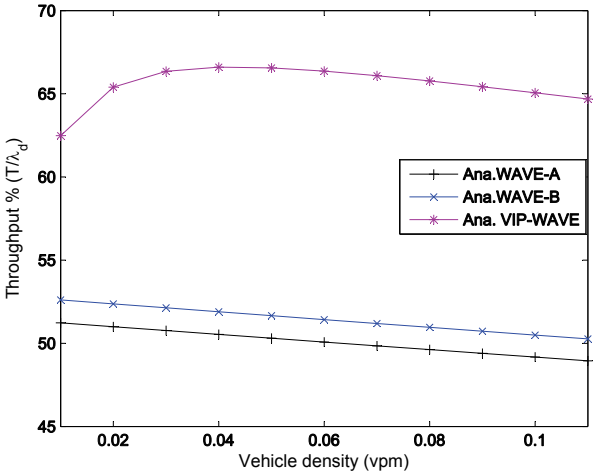
(a) One-way traffic **(b)** Two-way traffic

Figure 3.12: Nodal downstream throughput for different relays availability and RUSs inter-distance, RSUs inter-distance $X = 1000$m, average speed $v = 35$Km/h, and constant density $\rho = 1/25$ vpm

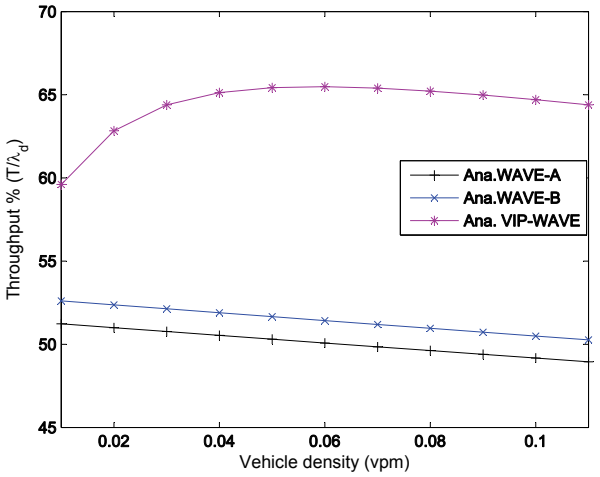
the vehicle density. The reason of the throughput increase before the maximum point is due to a greater number of available relays when the vehicle density increases. After the maximum point, the throughput goes down because as there are more vehicles on the road, the congestion of communications is dominant over the benefit from the increase of available relays. Figures 3.13a, 3.13b, and 3.13c exemplify how the maximum point varies



(a) $p_r = 1.0$



(b) $p_r = 0.7$



(c) $p_r = 0.4$

Figure 3.13: Nodal downstream throughput for different vehicle densities, RSUs inter-distance $X = 1500$m, and average speed $v = 35$Km/h

according to the different values of p_r.

Impact of downloading data rates

An evaluation of how data rate demanding IP applications (i.e., λ_d >1Mpbs) affect the overall performance of the nodal throughput is illustrated in Fig. 3.14. In the experiment, we calculate the throughput of VIP-WAVE and WAVE standards under saturated conditions, for a vehicular network with low-level presence of infrastructure. In all three cases, simulation and analytical results are configured to allow for a 60% of the nodes around the tagged vehicle to be actively transmitting in the same service channel. Since every active vehicle intends to transmit at a larger data rate, the congestion of communications become more and more severe, and thus, the performance of throughput degrades when the data rate increases. At the same time, a larger amount of data packets are lost when the OBU is experiencing a handover.

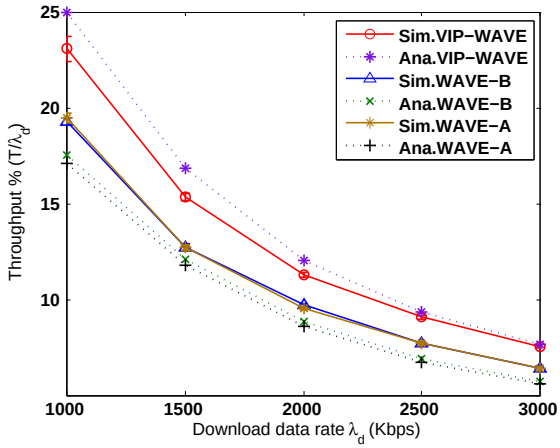


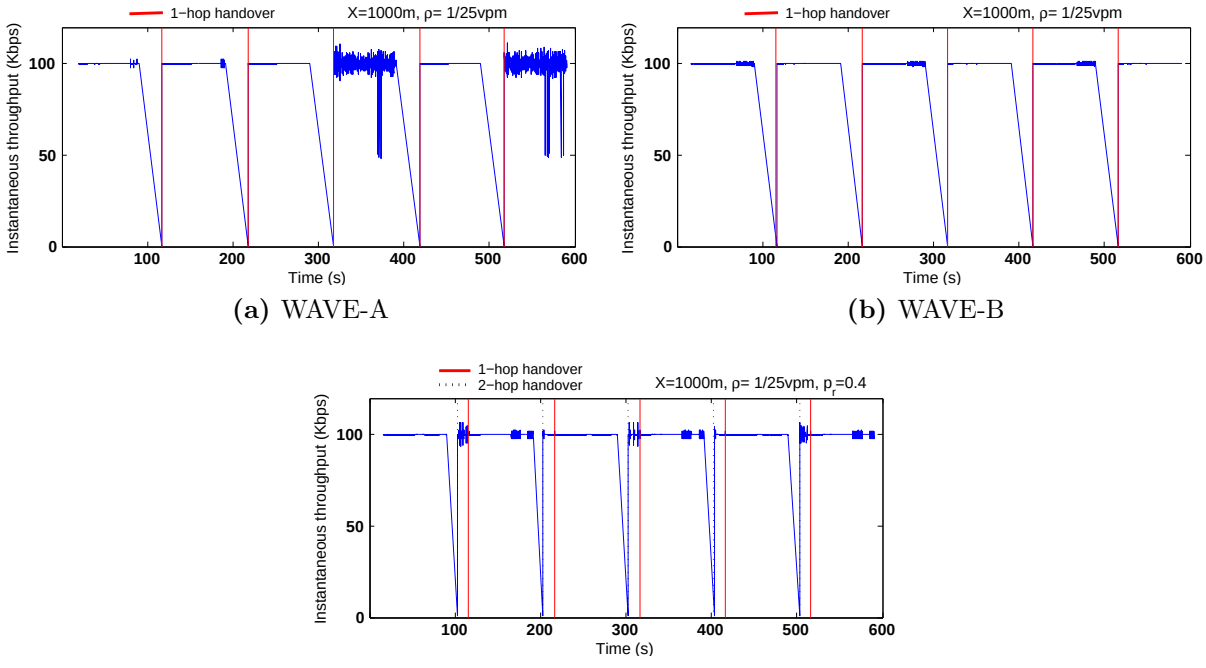
Figure 3.14: Nodal downstream throughput under saturated conditions for highly demanding IP applications, RSUs inter-distance $X = 1500$m, average speed $v = 35$Km/h, and constant density $\rho = 1/25$ vpm

We can also observe that the improvement obtained by VIP-WAVE compared to the standard WAVE tends to be reduced due to the congestion of communications becoming dominant for larger data rates. However, these throughput measurements may actually be better in real life scenarios, since the MAC layer in 802.11p/WAVE allows for prioritization of traffic by means of the EDCA mechanism (for simplicity, our simulation employs a single

access category queue). Furthermore, access control and quality of service policies could be imposed in order to guarantee the minimum level of the quality to the OBUs that are consuming the IP service [75].

Instantaneous throughput and delay

In order to evaluate the throughput behavior during a given session time, we show in Fig. 3.15 the instantaneous throughput for the three different schemes. In all three schemes, 60% of the nodes around the tagged vehicle are subscribed to the same service, which means there are other nodes that are actively transmitting in the same service channel. In the case of VIP-WAVE, this condition translates to having a 40% probability of finding an available relay among the neighboring vehicles.



(a) WAVE-A **(b)** WAVE-B

(c) VIP-WAVE

Figure 3.15: Instantaneous throughput and handover delay for different WAVE schemes

The figures illustrate the times at which every handover occurs. Given the constant

average speed and the fixed distance between RSUs, it is expected for the handovers to occur every fixed number of seconds. Nonetheless, the results help on understanding the nature of handovers in each scheme. It is observed how the presence of an IP mobility management scheme makes smoother the transition during handovers when comparing WAVE-A (Fig. 3.15a) with WAVE-B (Fig. 3.15b). Moreover, although the region between RSUs is fully covered when X=1000m, the handover delay in WAVE-A and WAVE-B is longer than the one experienced in VIP-WAVE. This is because the OBU needs to re-establish the connection with the new RSU, and given that $r < R$, it takes some time until when the RSU is able to receive the location update in the form of a Router Solicitation or a RESET message from the OBU, which is only possible when $x < R$ or $x > X - R$ (where x is the OBU's location). This phenomenon has a smaller impact in VIP-WAVE, since the framework allows for two-hop communications toward the RSU when the OBU is unable to communicate directly. Thus, the total handover delay in VIP-WAVE is reduced, and a smaller number of packet losses is perceived by the IP application.

Additionally, since we only consider the received application-layer IP packets in the calculation of the instantaneous throughput, in Fig. 3.15c can be observed that the overhead incurred in establishing the relayed connection plays a minor impact in the overall performance of the end-to-end communications, as the throughput remains fairly stable, while at the same time the relaying helps on reducing the total packet losses, as we mentioned before.

Furthermore, as many IP applications are delay-sensitive, we evaluate the effect of two-hop communications in the data packets end-to-end delay. Fig. 3.16 depicts the latency experienced by individual packets received at the OBU during a session time. For those packets being transmitted through a two-hop connection in the 802.11p network, they perceive a slightly higher latency than those using a one-hop connection. However, the total delay, which is less than 37ms in all cases, fits well into the delay requirements for the main multimedia applications, such as 150ms for real time audio, and 250ms for video conferencing and video streaming. The variations observed in the delay of packets using

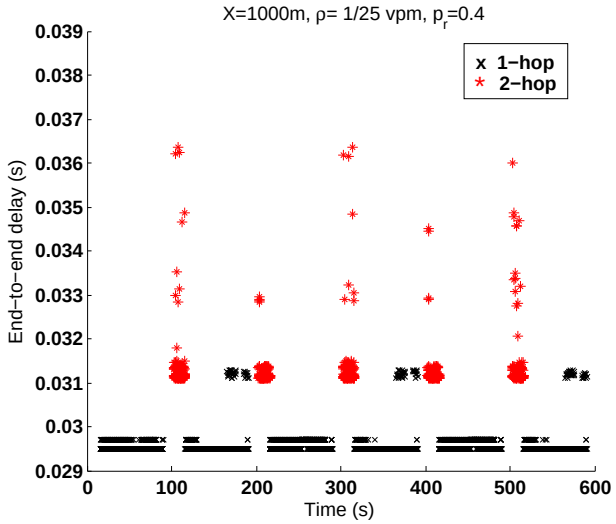


Figure 3.16: Data packet end-to-end delay in VIP-WAVE

the same number of hops, come from the MAC layer retransmissions that are caused when there are colliding transmissions in the wireless domain.

3.6 Summary

This chapter has presented VIP-WAVE, a novel framework for the support of IP communications in 802.11p/WAVE vehicular networks. In particular, we have studied the limitations in the IEEE 802.11p/WAVE standards for the operation and differentiation of IP applications, and have proposed the VIP-WAVE framework to address those limitations. The key advantages of VIP-WAVE can be summarized as follows:

- It has been demonstrated to notably improve the performance of IP applications even when a low presence of infrastructure results in large gaps between areas of coverage.
- The protocols and mechanisms proposed in VIP-WAVE for IP addressing, mobility management, and multi-hop communications, have been all designed according to the intricacies and special characteristics of 802.11p/WAVE networks. Thus, they

re-use existent signalling messages defined in WAVE and do not attempt to stress the control channel employed for safety communications.

- It provides an accurate analytical model that allows for the integration of aspects from different layers, such as mobility and channel conditions, probability of connectivity to the infrastructure, handover delays, and packet collision probabilities, in order to estimate the nodal downstream throughput perceived by a WAVE user that is consuming an IP service from the infrastructure. The model has been validated through extensive simulations.

Furthermore, we have reinforced our observation that the individual downloading data rate perceived by an OBU is highly dependant on the road density and the inter-distance of the RSUs. Our results suggest that it is beneficial for 802.11p/WAVE networks to put in place multi-hop communications that may extend the area of coverage and may help to make smoother the transitions during handovers.

Although we have exploited the use of multi-hop paths in this chapter, during the simulations stage we have also observed that establishing and fixing the forwarding of data packets through a selected relay may become a costly task in cases when the topology is highly variable. Therefore, in the following chapter, we employ this observation and exploit other features of the vehicular network, such as the availability of location information, not only to improve the relay selection, so that decisions are made on a per-packet basis, but also to accelerate the handover process and to reduce the handover delay.

Chapter 4

MA-PMIP: A Multi-hop Authenticated Proxy Mobile IP scheme for Asymmetric VCN

4.1 Preliminaries

In this chapter, we propose a **M**ulti-hop **A**uthenticated **P**roxy **M**obile **IP** (**MA-PMIP**) scheme for asymmetric vehicular communications networks (VCN). Continuing with the provision of IP services in multi-hop vehicular environments, this part of our work focuses on vehicular communications networks powered by WLAN technologies (802.11a/b/g/n).

In this context, drive-thru Internet and IP-based infotainment applications are supported by road-side Access Routers (ARs) that connect the VCN to external IP networks. However, connections between ARs and vehicles suffer from asymmetric links due to variable transmission ranges caused by mobility, obstacles, and dissimilar transmission powers, which makes them difficult to maintain the bidirectional connections, and to provide the IP mobility required by most IP applications. Moreover, vehicular mobility results in short-lived connections to the AR, affecting the availability of IP services in the VCN.

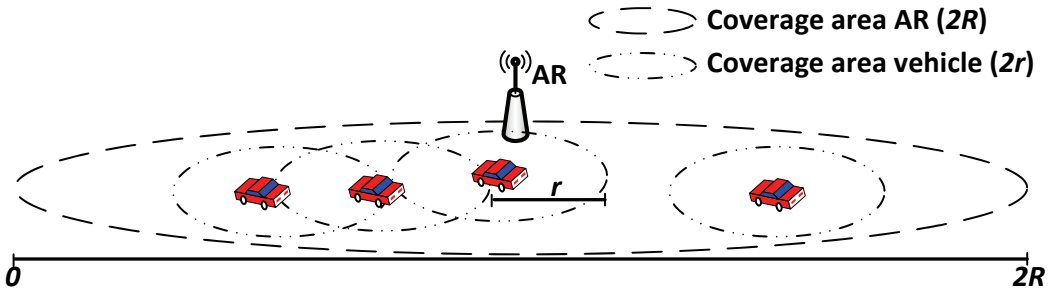


Figure 4.1: Asymmetric links in VCN

Building upon the concept of Infrastructure-to-Vehicle-to-Vehicle (I2V2V) communications, introduced in Section 1.1, and different from previous works that assume symmetric links among all the devices in the vehicular network, in this chapter we demonstrate that multi-hop communications are key for avoiding service breakage in asymmetric VCN. An asymmetric VCN is illustrated in Fig. 4.1. Such a network suffers from asymmetric transmission ranges due to mobility, path losses in the presence of obstacles, and dissimilar transmission powers among the VCN devices. Although one-way links may not affect some applications that require only the link AR→vehicle (e.g., safety-related information), this problem severely affects IP-based applications. In particular, TCP applications require the packets to be acknowledged, but one-way links make it impossible to confirm reception of packets. In fact, the vehicle will not be able to initiate any client-server application unless it establishes a bidirectional link with the AR. Note that the client-server architecture is the most common architecture deployed by Internet communications.

As a result, symmetric links have been a frequent assumption for investigating the deployment of IP services. However, when the asymmetric links are discounted by routing protocols in ad hoc networks, it can result in low data transmission rates and low network connectivity [76]. Thus, the presence of asymmetric links has been studied from the point of view of data dissemination in vehicular ad hoc networks (VANET) [77] and its implications in geographic routing [78], but the impact on the provision of IP services in the VANET is yet to be studied. Since previous works have shown that the inclusion of asymmetric links in the routing decisions may result in a better network performance [76], we consider the

asymmetric VCN as the foundation for our network model.

With MA-PMIP, we aim at better selecting relays on a per-packet basis and depending on the directionality of links. Furthermore, we intend to employ the geo-networking features of the vehicular network, and take advantage of information such as geographical location and road traffic conditions, in order to enable predictive handovers. The detailed design goals of our scheme are explained as follows:

- The provision of an IP mobility scheme for multi-hop VCN that integrates location and road traffic information in order to predict handovers;
- The consideration of asymmetric links in the VCN, in order adapt the geo-networking routing mechanism depending on the availability of bidirectional links;
- The secure handover signalling when a V2V path is employed to reach the infrastructure, so that possible attacks are mitigated without affecting the performance of the ongoing sessions.

The aforementioned design goals are, to the best of our knowledge, the first to combine a predictive IP mobility scheme designed for multi-hop asymmetric VCN, with the security issues of employing I2V2V communications.

In the following sections, we first review the related work and describe our reference system model (Sections 4.2 and 4.3, respectively). Next, we introduce the MA-PMIP scheme (Section 4.4), followed by an analytical evaluation (Section 4.5). After that, we present extensive simulations results that corroborate the findings from our analytical evaluation (Section 4.6).

4.2 Related work

IP addressing and mobility solutions for vehicular environments have been studied from different perspectives. In the case of multi-hop VCN, studies have been proposed to enable

network mobility (i.e., for providing mobility to all the users in the in-vehicle local network) based on the Network Mobility Basic Support (NEMO BS) protocol [25], or based on PMIP. The NEMO-based solutions in [33] and [79], employ a geographic routing protocol to obtain IP addresses directly from the infrastructure. Geographic routing has been shown to be effective, to the point that it has been standardized for communications in Intelligent Transport Systems [56]. Nevertheless, although NEMO BS minimizes the binding update signalling, it also brings a costly tunneling overhead. Thus, there have been proposals to balance the tradeoff between these two factors in one-hop scenarios [44]. However, that is yet to be explored when NEMO BS is extended through multi-hop communications.

On the other hand, since the standard PMIP only supports mobility for a single node, the solutions in [11, 13, 80] adapt the protocol to reduce the signalling when a local network is to be served by the in-vehicle mobile router. Lee *et al.* [11] propose P-NEMO to maintain the Internet connectivity at the vehicle, and provides a make-before-break mechanism when vehicles switch to a new access network. In [80] and [13], the authors propose to forward solicitations from local users, so that nodes in the local network may obtain addresses directly from the PMIP domain; the first solution proposes to use a proxy router to forward such solicitations, whereas the second extends some functionalities for the mobile router to serve as a mobile MAG, so that it exchanges mobility signalling with the LMA. However, these works do not address the mobility problem when other vehicles are connected through multi-hop paths, which is the main concern addressed in this paper.

Although PMIP has a good acceptance for its applicability in vehicular scenarios, it has an important restriction for its deployment in I2V2V communications. The protocol, by definition, requires the MN to have a direct connection to the MAG for two reasons. Firstly, the MAG is expected to detect new connections and disconnections based on one-hop communications. Secondly, the network-based mobility service should be offered only after authenticating and authorizing the mobile node for that service; however, those tasks are assumed to happen over the MAG–MN link, but not in the presence of intermediate routers [40]. Therefore, it is still necessary to devise a solution in which the multi-hop links

in the VCN are considered.

Moreover, none of the aforementioned studies explore the problem of security. In the case of NEMO-based solutions, they let the routing protocol to be responsible for securing the communications, whereas the PMIP-based solutions rely on the assumption that the intermediate node —in this case, the proxy mobile router— is by some means a secure entity in the PMIP domain.

Continuing with the problem of security and authentication schemes for multi-hop networks, previous works are mainly focused on two different approaches: 1) end-to-end authentication, which employs a relay node (RN) to only forward the authentication credentials between mobile node and the infrastructure; and 2) hop-by-hop authentication, which implements authentication algorithms between every two hops. Following the first approach, in [81] the MN uses its public key certificate to authenticate itself to the foreign gateway. On the other hand, the scheme in [82] uses both symmetric key for authenticating an MN to its home network, and public key for mutual authentication between home network and foreign network. However, the expensive computation involved with public key operations tends to increase the end-to-end delay.

Conversely, a symmetric key-based authentication scheme for multi-hop Mobile IP is proposed in [83]. In that work, an MN authenticates itself to its home authentication server, which derives a group of keys to be used by the MN. Despite the low computation and communication overheads, the symmetric key-based schemes cannot achieve as strong levels of authentication as those achieved by public key-based schemes. This is because the sharing of the secret key between the two peers increases the chances for adversaries to identify the shared key. Instead, public key-based schemes create a unique secret key for each user; hence, it is more difficult for adversaries to identify the keys.

Following the second approach, a mutual authentication that depends on both secret splitting and self-certified schemes is proposed in [84]. However, both schemes are prone to DoS attacks. Another scheme for hop-by-hop authentication, called Alpha, is presented in [85]. In Alpha, the MN signs the messages using a hash chain element as the key for

signing, and then delays the key disclosure until receiving an acknowledgement from the intermediate node. Although Alpha protects the network from insider attacks, it suffers from a high end-to-end delay. A hybrid approach, the adaptive message authentication scheme (AMA), is proposed in [86]. It adapts the strength of the security checks depending on the security conditions of the network at the moment of packet forwarding.

Different from the aforementioned authentication schemes, in MA-PMIP we propose a light-weight mutual authentication scheme to be employed between the source node and the relay, which mitigates the high delay that is introduced by previous hop-by-hop schemes. Therefore, the proposed scheme can be used with seamless handover operations in our multi-hop VCN during the I2V2V communications.

4.3 Reference System

4.3.1 Network Model

A vehicular communications network such as the one shown in Fig. 4.2 is considered. Connections to the infrastructure are enabled by means of road-side Access Routers (ARs), each one in charge of a different wireless access network. Vehicles are equipped with wireless interfaces, as well as GPS systems that feed a location service from which the location of vehicles is obtained. Beacon messages are employed by vehicles to inform about their location, direction, speed, acceleration, and traffic events to their neighbors.

We consider the presence of asymmetric links in the VCN (Fig. 4.1). The delivery of packets is assisted by a geographic routing protocol. To serve this protocol, a location server stores the location of vehicles, and is available for providing updated responses to queries made by nodes participating in the routing of packets. In order to forward packets within the multi-hop VCN, a virtual link between AR and vehicle is created [87]. That means that a geo-routing header is appended to each packet, where the location and geo-identifier of the recipient are indicated. In this way, the geo-routing layer is in charge of

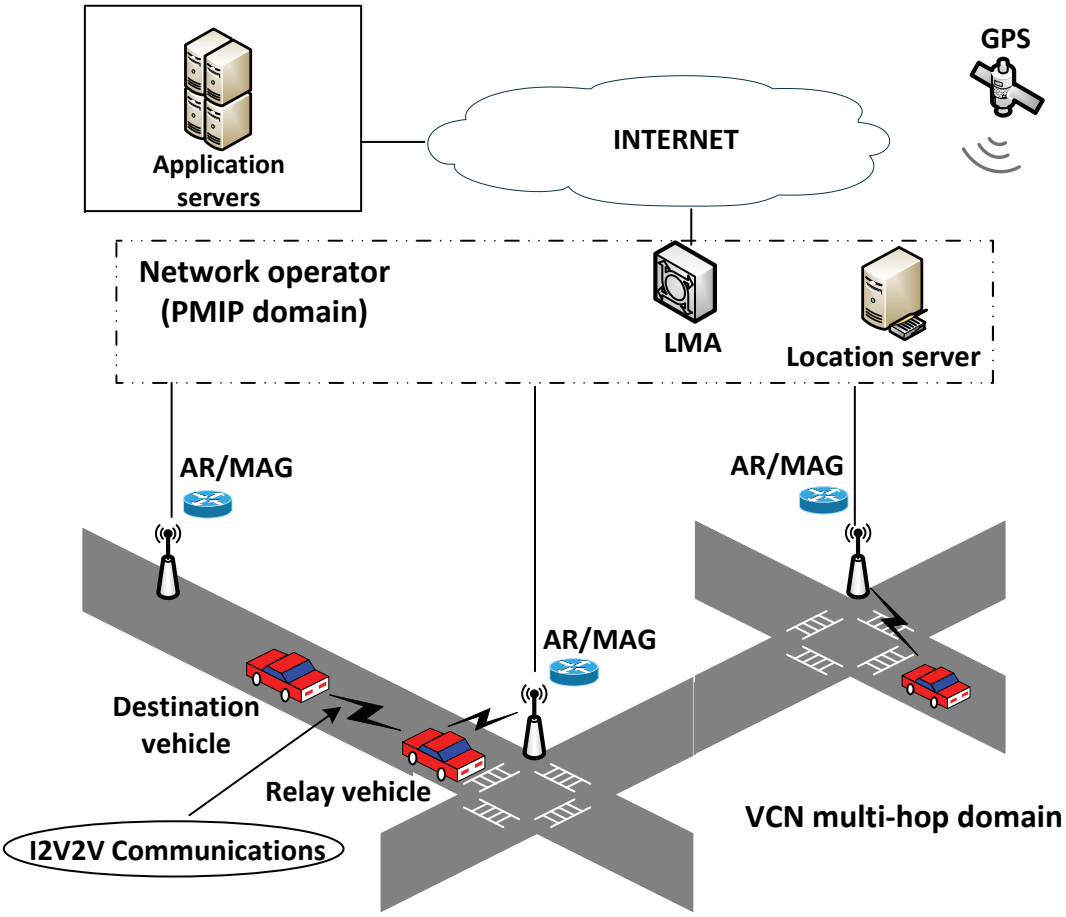


Figure 4.2: Network Topology

the hop-by-hop forwarding through multi-hop paths, with no need of processing the IP headers at the intermediate vehicles.

The ARs service areas are well-defined by the network operator. A well-defined area means that messages from ARs to the VCN are only forwarded within a certain geographic area [58]. Each AR announces its services in geocast beacon messages with the flag `AccessRouter` activated. The beacons are forwarded through multi-hop paths as long as the hops are located inside the coordinates indicated by the geocast packet header. In this way, vehicles in the connected VCN can extract information from the geocast header, such as AR's location, AR's geo-identifier, and the service area limiting coordinates. We

assume the infrastructure is a planned network with non-overlapping and consecutive service areas. Note that, although service areas are consecutive, some locations within them are not reachable through one-hop connections. This may be caused by weak channel conditions, and by the asymmetric links between ARs and vehicles.

On the other hand, to ensure the proper operation of the geo-routing protocol and MA-PMIP, it is required to maintain state information at the entities exchanging IP packets. The following are the required data structures:

Neighbors Table: stores information about the neighboring nodes. The table indicates a link type —unidirectional or bidirectional— for each neighbor. A node detects the bidirectional links in the following way: incoming links are verified when beacon messages are received from neighbors (i.e., this node can hear its neighbors); outward links are verified by checking the neighbors' locations and the node's transmission power, in order to calculate if such neighbors are inside the radio range (i.e., the neighbors can hear this node). The table is stored by vehicles and ARs.

Default gateway table: stores information about the AR in the current service area. It contains the AR's geo-identifier and the service area coordinates. If the destination of a packet is an external node, the geographic routing forwards the packet toward the default gateway indicated in this table. Then, the AR routes the packet to its final destination. The table is stored by vehicles.

We only consider IP-based applications accessed from the VCN. Such applications are hosted in external networks that may be private (for dedicated content), or public, such as the Internet. Since we have selected PMIP for handling the IP mobility in the network, all the ARs are assumed to belong to a single PMIP domain. The AR and MAG are co-located in our model. Therefore, the terms AR and MAG are used interchangeably in the following sections.

Different from [58], in our scheme the AR does not send Router Advertisement (RA) messages announcing the IP prefix to vehicles in the service area. Instead, when a vehicle

joins the network for the first time, individual IP prefixes are allocated through PMIP. It is required by MA-PMIP to obtain this initial IP configuration only when a one-hop connection exists between vehicle and MAG, so that authentication material is securely exchanged for future handovers of the vehicle over multi-hop paths. Note that a one-hop connection is only established when a bidirectional link exists between the two entities.

4.3.2 Threat and Trust Models

We consider both internal and external adversaries to be present during I2V2V communications. Internal adversaries are legitimate users who exploit their legitimacy to harm other users. Two types of internal adversaries are defined: impersonation and colluder. The former impersonates another MN's identity and sends neighbor discovery messages such as Router Solicitation through the relay node. The latter colludes with other domain users in order to identify the shared secret key between two legitimate users.

External adversaries are unauthorized users who aim at identifying the secret key and breaking the authentication scheme. We consider replay, man-in-the-middle (MITM), and denial of service (DoS) attacks as external adversaries. The goal of the MITM and replay attacks is to identify a shared key between two legitimate users, while the goal of DoS attack is to exhaust the system resources. In our model, we assume the LMA and MAG entities to be trusted nodes.

4.4 Multi-hop Authenticated Proxy Mobile IP Scheme (MA-PMIP)

In this section, we introduce the basic and predictive operation of MA-PMIP, the handling of asymmetric links, and the multi-hop authentication mechanism that allows for secure signalling during handovers.

4.4.1 Basic Operation

The signalling of MA-PMIP for initial IP configuration follows the one defined by the standard PMIP. Once the vehicle joins the domain for the first time, it sends Router Solicitation(RS) messages, which are employed by the MAG as a hint for detecting the new connection. Once the PMIP signalling has been completed, the MAG announces the IP prefix in a unicast RA message delivered to the vehicle over the one-hop connection. In order to enable communications from the in-vehicle local network, the MR may obtain additional prefixes by means of prefix delegation [88] or prefix division [89], as it is currently proposed at the IETF for network mobility support with PMIP.

Fig. 4.3 shows the basic MA-PMIP signalling employed when a vehicle experiences a handover through a relay. The movement detection could be triggered by any of the following events: 1) the vehicle has started receiving AR geocast messages with a geo-identifier different from the one registered in the *default gateway table*; or 2) the vehicle has detected its current location falls outside the service area of the registered AR. If the vehicle losses one-hop connection toward the MAG, but it is still inside the registered service area, then no IP mobility signalling is required and packets are forwarded by means of the geo-routing protocol.

After movement detection, the RS message is an indicator for others (i.e., relay vehicle and MAG) of the vehicle's intention to re-establish a connection in the PMIP domain. Thus, an authentication is required to ensure that both nodes source and relay are legitimate and are not performing any of the attacks described in Section 4.3.2. Details of the authentication procedure are later explained in section 4.4.4. Once the nodes are authenticated, the RS packet is forwarded until it reaches the MAG, and the PMIP signalling is completed in order to maintain the IP assignment at the vehicle's new location.

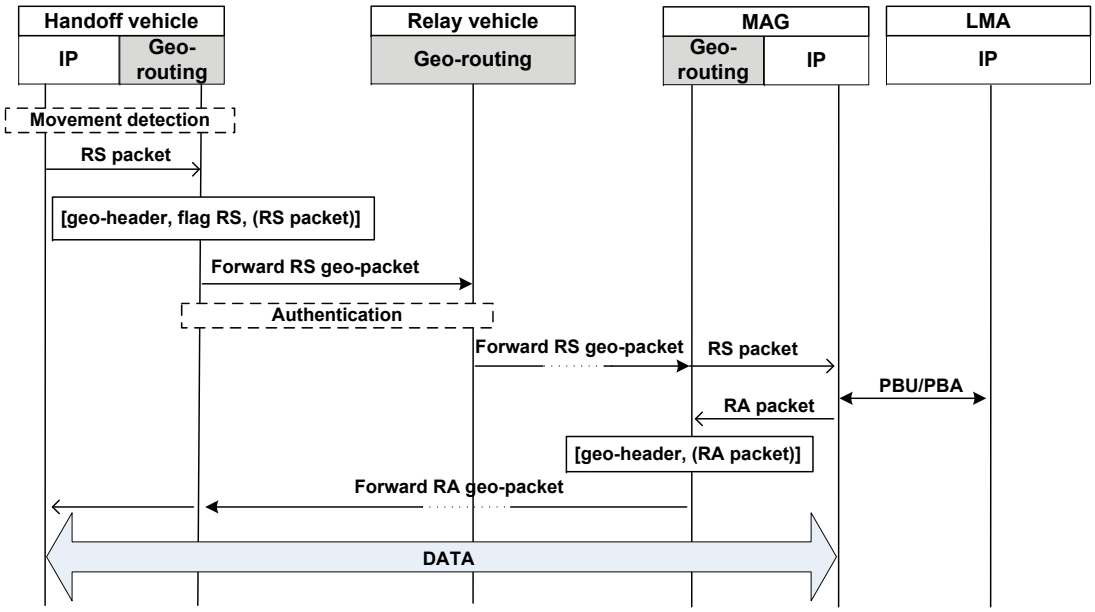


Figure 4.3: Handover through I2V2V communications in MA-PMIP

4.4.2 Predictive handovers

We propose a prediction mechanism that enables a faster handover procedure, which takes advantage of the location information available in the VCN. It consists of an estimation of the time at which the vehicle will move to a new service area, and is coupled with the recently standardized Fast handovers for Proxy Mobile IPv6 [FPMIP] (RFC 5949 [42]). FPMIP in predictive mode defines the signalling between previous MAG (PMAG) and new MAG (NMAG) for pre-establishing a tunnel and forwarding the data packets to the new access network. This aims at minimizing packet losses when the mobile node loses connectivity in both previous and new access networks. Once the node is detected in the new access network, the NMAG forwards the buffered packets to the node, and signals the LMA so that the MAG-to-MAG tunnel can be deactivated.

We do not introduce any changes to the standard FPMIP. Instead, the extensions necessary at the PMAG for estimating the time at which the handover will occur are introduced. In this way, the MAG-to-MAG tunnel can be timely established. Furthermore, the pro-

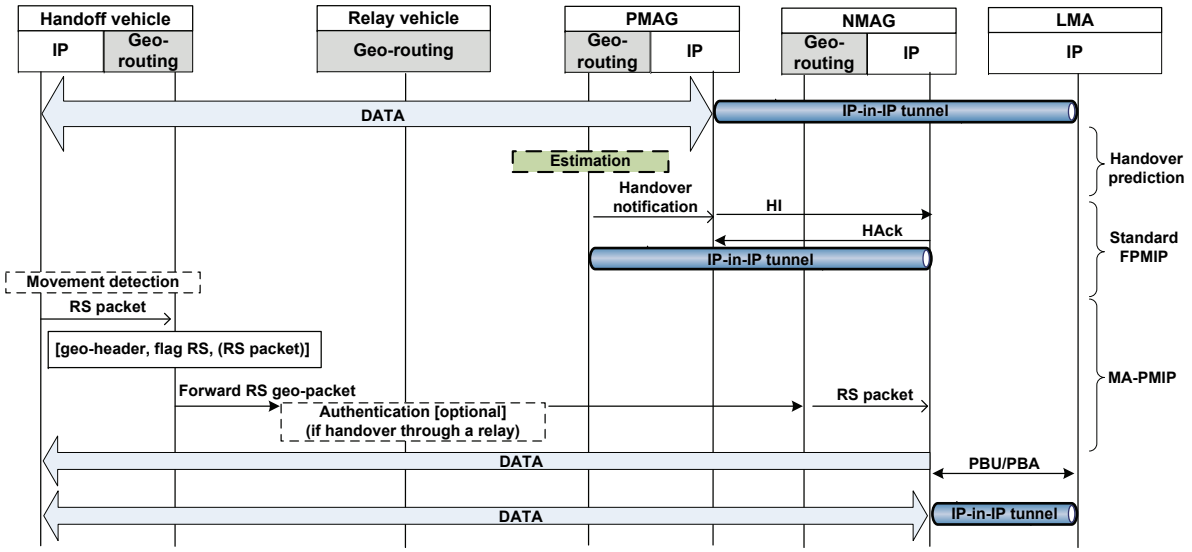


Figure 4.4: Prediction Mechanism for Fast Handovers in MA-PMIP

posed predictive mechanism works for both one-hop and multi-hop connected vehicles in the VCN, and is meant to be enabled only for those vehicles that have active communications. For inactive vehicles that handover across the PMIP domain, they may follow the basic MA-PMIP signalling described in Section 4.4.1.

The prediction process is depicted in Fig. 4.4 and explained in detail as follows. The PMAG queries the location of a vehicle for which a packet has to be delivered. That information is retrieved from the location server, together with the destination vehicle's velocity and traffic density (i.e., vehicles per meter). The latter is calculated by the location server based on the information received about vehicles in that particular service area. In order to estimate the time at which the handover will occur, we construct a weighted average that considers two aspects: the current driving characteristics at the destination vehicle (i.e., current or last reported velocity v_r), and the average flow velocity v_{avg} determined from traffic conditions. According to the Greenshields model, the average flow velocity v_{avg} can be related to traffic conditions as follows [45]:

$$v_{avg} = \left(1 - \frac{k}{k_{jam}}\right) v_f, \tag{4.1}$$

where k is the traffic density, k_{jam} is the density associated with a completely stopped traffic flow, and v_f corresponds to the free-flow speed, i.e., the road speed limit. Therefore, we calculate the estimated vehicle's velocity as follows:

$$v_{est} = (1 - \kappa) \times v_r + \kappa \times v_{avg}. \tag{4.2}$$

If there exists a bidirectional link to the destination vehicle, the PMAG obtains the current velocity v_r from the *Neighbors Table*. Otherwise, it uses the last reported velocity that is retrieved from the location server. The value for κ could be adjusted at the service area level, according to different priorities. For example, a value of $\kappa = 0.875$ could be employed for a service area in which drive-thru traffic is dominant (i.e., not an area where vehicles typically park), so that velocity is mostly determined by the road density. It is important to note that since v_r is close to a "real-time" report of the current velocity, it encloses not only the velocity due to past traffic conditions, but also the isolated driver's behavior.

Once v_{est} is estimated, the time to reach the edge of the service area level is easily calculated as $t_{est} = d_{est}/v_{est}$, where d_{est} corresponds to the Euclidean distance from the current location of the vehicle to the edge of the service area. We then form a heuristic to make the following decision: if the time to reach the edge is less than the one determined by a threshold value, the PMAG initiates the signalling for FPMIP depicted in Fig. 4.4. After the vehicle moves to the new service area, it sends the RS message as a result of the movement detection, and the tunnel between LMA and NMAG is set for the normal routing of traffic to the vehicle's new location.

4.4.3 Handling of asymmetric links

The asymmetric links in MA-PMIP are detected and handled in two different layers: 1) at the network layer, by means of the Neighbor Discovery protocol and the Neighbor

Unreachability Detection mechanism [54]; and 2) at the geo-networking layer, which follows the procedure described in Section 4.3.1 for the link type identification.

MA-PMIP employs the two mechanisms in order to react to the directionality of links during the delivery of IP packets. For instance, consider an IP application that requires a bidirectional link for its proper operation. We then assume IP packets are marked by the application server to indicate the required application's directionality. Such a marking could be set in the Flow Label field of the IPv6 header. In case the server does not employ/support flow labeling, the LMA may still set the mark by checking the transport protocol in the IP packet header. In either case, the LMA codes the directionality in the Flow Label field of the outer header of packets sent in the tunnel LMA $\rightarrow$ MAG. In this way, the MAG has the necessary information for routing the packets accordingly in the VCN.

Before packets are forwarded, the MAG checks the directionality requirement for each packet and proceeds as follows:

Bidirectional flow

- If Neighbor Discovery detects the destination vehicle is disconnected from the MAG, the packet is discarded unless the prediction mechanism has been activated.
- Else, if the vehicle is still connected, the packet is delivered to the geo-networking layer to continue with routing.

Unidirectional flow

- If the prediction mechanism has been activated, then forward the packet accordingly.
- Else, the packet is delivered to the geo-networking layer to continue with routing.

Once the geo-networking layer receives a packet, it employs the information in *Neighbors Table* to select relays that are close to the destination. If the flow of packets requires

bidirectionality, the selected relays are additionally filtered depending if they are set as bidirectional in the *Neighbors Table*. This combined routing metric distance/type-of-link is employed in both directions: from MAG to destination vehicle, and from destination vehicle to MAG.

4.4.4 Authentication

As depicted in Fig. 4.3 and Fig. 4.4, an authentication scheme should be employed to mutually authenticate the roaming vehicle (also known as Mobile Router [MR]) and the relay vehicle (RN). In this way, the signalling related to the handover process is protected against the threats identified in Section 4.3.2. Therefore, MA-PMIP integrates the Efficient Mutual Multi-hop Mobile Authentication (EM^{3}A) scheme for PMIP Networks we have introduced in [90]. The authentication mechanism employs the concept of symmetric polynomials, which has been used in decentralized key generation schemes for arbitrary nodes in a heterogeneous network [91,92]. However, the proposed authentication increases the secrecy level achieved by the previous symmetric polynomial schemes. In MA-PMIP, the secrecy increases from t to $t \times 2^n$, where n is the number of MAGs in the domain, and t is the degree of the network polynomials.

EM^{3}A consists of three main phases: key establishment phase, for establishing and distributing keys; registration phase, for obtaining the secure material from the PMIP domain; and authentication phase, for mutually authenticating the MR and the RN. During the first phase, the LMA generates a domain symmetric polynomial $F(w, x, y, z)$, which is evaluated for each MAG's identity and sent to every MAG. The polynomial function received at each MAG, $F(ID_{\text{MAG}_i}, x, y, z), \quad i = 1, 2, ..., n$, is later used for keys generation at the MNs in a decentralized manner.

Then, at the registration phase, a vehicle that connects for the first time in the PMIP domain, receives the polynomial $F(ID_{\text{FMAG}}, ID_{MR}, y, z)$ from the MAG (which has been evaluated for the identity of the First MAG and the MR's identity), along with the list of

valid MAGs in the domain. In this way, during the authentication phase, the MR and RN generate a shared key and authenticate each other through a challenge-response scheme, which consists of a three-way hand shake [90]:

1. The MR sends its own credentials attached to the Router Solicitation message to the RN.

2. The RN verifies the MR's identity, creates a challenge message encrypted with the generated shared key, and sends it to the MR.

3. The MR verifies the RN's identity and responses to MR in an encrypted message employing the shared key.

Once the authentication phase is completed, the RN forwards the MR's router solicitation message to the MAG, which allows for MA-PMIP to continue its operation and to maintain seamless communication. In this way, MA-PMIP thwarts both the internal and external adversaries identified in Section 4.3.2.

4.5 Analytical evaluation of MA-PMIP

In this section, we evaluate the performance of MA-PMIP with respect to the following metrics:

- *Location update signalling cost*: Traffic load necessary to update the current location of the vehicle (e.g., PBU/PBA, BU/BA messages).

- *Packet delivery cost*: Overhead traffic necessary to deliver a data packet to its final destination (e.g., an IP tunnel header).

- *Handover delay*: The elapsed time from the moment the vehicle looses connection at the old location, to the moment it is able to resume the transmission/reception of data packets at the new location.

To calculate these metrics, we describe the vehicle's mobility based on a fluid flow model, and calculate the probability $\alpha(i)$ of the vehicle crossing i service areas during an IP session [11]. We have chosen the MANET-centric NEMO (MANEMO) scheme [33], introduced in Section 4.2, for comparison purposes. Both MANEMO and MA-PMIP enable IP mobility in multi-hop VCN scenarios, and consider communications from the in-vehicle local network. MA-PMIP by default employs authentication and predictive handovers. However, we also analyze the MA-PMIP Basic operation.

Further security analysis, as well as the computation and communication overheads introduced by the authentication mechanism described in Section 4.4.4 can be found in [90].

4.5.1 Location update and packet delivery cost

Assume the infrastructure is composed of N service areas and each area is served by one AR. Each well-defined area is square-shaped, with perimeter D and area A. The service area residence time (i.e., the time a vehicle spends inside a service area) is assumed to have a general distribution $f_{\mathrm{SA}}(t)$ with mean $1/\mu$. According to the fluid flow model, the service area crossing rate μ can be calculated as $\mu = vD/(\pi A)$, where v indicates the average velocity, and π indicates the vehicle's direction.

Since the IP services are mainly for the downloading of information from servers in the infrastructure, only incoming data sessions are considered for simplicity of the analysis. Sessions have an average length of L (packets), with exponentially distributed inter-session arrival times, and arriving at an average rate λ_I. Each vehicle has independent and identically distributed session arrival rates.

The inter-session arrival time is defined as the elapsed time between the arrival of the first data packet of a session and the arrival of the next session's first data packet. During an inter-session arrival time, the probability of crossing i service areas, $\alpha(i)$, is expressed by:

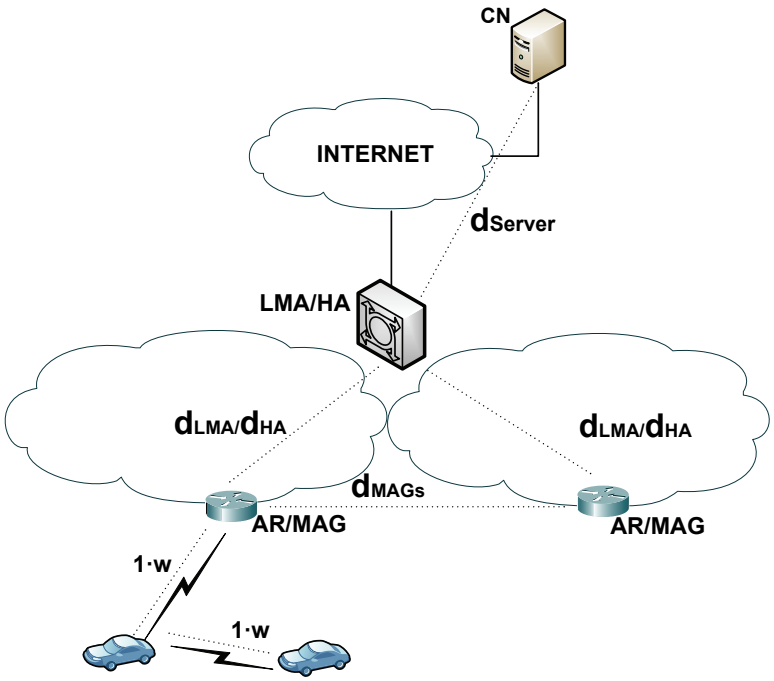


Figure 4.5: MA-PMIP Performance Analysis

$$\alpha(i) = \begin{cases} 1 - \frac{1}{\rho_s}[1 - f^*_{\mathrm{SA}}(\lambda_I)] & \text{if } i = 0, \\ \frac{1}{\rho_s}[1 - f^*_{\mathrm{SA}}(\lambda_I)]^2[f^*_{\mathrm{SA}}(\lambda_I)]^{i-1} & \text{if } i > 0, \end{cases} \tag{4.3}$$

where $\rho_s = \lambda_I/\mu$ indicates the session to mobility ratio, and $f^*_{\mathrm{SA}}(\lambda_I)$ is the Laplace transform of the service area residence time distribution [44].

The distances between network elements (i.e., number of intermediate hops) are shown in Fig. 4.5 by $d_{\mathrm{LMA}}, d_{\mathrm{HA}}$, and d_{MAGs}. The latter, d_{MAGs}, is the number of hops between previous MAG and next MAG, whereas d_{HA} and d_{LMA} are the number of hops for the AR to reach the anchor point. On the other hand, n indicates the number of links of the multi-hop path traversed by signalling messages, and ω is the relative weight of transmitting packets over a wireless link. Although wireless links only have one-hop distance, the weight factor indicates their high cost compared to links in the wired network.

The location update signalling cost per handover, BU, is obtained according to the number of hops the signalling messages have to cross to reach the anchor point (i.e., LMA in MA-PMIP, and Home Agent in MANEMO). It is calculated as follows:

$$BU^{\text{MA-PMIP}} = d_{\text{MAGs}} \times P + d_{\text{LMA}} \times U^{\text{MA-PMIP}}, \tag{4.4}$$

$$BU^{\text{MA-PMIP(Basic)}} = d_{\text{LMA}} \times U^{\text{MA-PMIP}}, \tag{4.5}$$

$$BU^{\text{MANEMO}} = (n \times \omega + d_{\text{HA}})U^{\text{MANEMO}}, \tag{4.6}$$

where P is the size (bytes) of the Handover Indicator/Handover Ack messages in the MA-PMIP with predictive handover, and U is the size (bytes) of Binding Update/Binding Ack (BU/BA) and PBU/PBA messages. Note that the PBU/PBA messages in (4.4) and (4.5) are only transmitted at the infrastructure side, as defined by the PMIP standard. In the contrary, MANEMO's signalling is transmitted in the wireless domain.

The total location update signalling cost, C_{BU} (bytes*hops), incurred by a vehicle moving across several service areas, is calculated as follows:

$$C_{\text{BU}} = \sum_{i} i \times BU \times \alpha(i), \tag{4.7}$$

where BU is replaced by (4.4), (4.5), and (4.6), accordingly. The crossing probability, $\alpha(i)$, is calculated from (4.3).

The delivery overhead cost per packet, PD, accounts for extra information and extra links traversed when delivering a data packet from a server to the vehicle. It is computed as follows:

$$PD^{\text{MA-PMIP}} = d_{\text{server}} + \beta(H(d_{\text{LMA}} + d_{\text{MAGs}}) + (n \times \omega)) + (1-\beta)(H \times d_{\text{LMA}} + (n \times \omega)), \quad (4.8)$$

$$PD^{\text{MA-PMIP(Basic)}} = d_{\text{server}} + H \times d_{\text{LMA}} + (n \times \omega), \quad (4.9)$$

$$PD^{\text{MANEMO}} = d_{\text{server}} + H(d_{\text{HA}} + n \times \omega), \quad (4.10)$$

where d_{server} is the distance from the application server to the anchor point, and H is the size of the tunnelling IP header.

In (4.8), β represents the portion of packets that traverse the extra PMAG-to-NMAG tunnel, before the vehicle is fully detected at the new location during predictive handovers (Fig. 4.4). Although MANEMO and MA-PMIP (Basic) require data packets to traverse the same number of hops, i.e., if LMA and Home Agent are equally distanced from the AR, the packets in (4.10) are encapsulated up to the destination vehicle. Instead, MA-PMIP (Basic) employs the tunnel only between LMA and serving MAG.

The total packet delivery cost, C_{PD} (bytes*hops), considers the number of active hosts m in the in-vehicle network, and the average session length L (packets). L depends on the downloading data rate γ, the packet size S, and the inter-session arrival rate λ_I. Thus, C_{PD} is calculated as follows:

$$C_{\text{PD}} = m \times PD \times L, \quad (4.11)$$

where PD is replaced by (4.8), (4.9), and (4.10), accordingly.

The total cost C_{T} is obtained by adding the total location update and total packet delivery cost of each scheme. Therefore, C_{T} is expressed by:

$$C_{\mathrm{T}} = C_{\mathrm{BU}} + C_{\mathrm{PD}}. \tag{4.12}$$

4.5.2 Handover delay

We quantify the delay D_{HD} incurred during a handover event as $D_{HD} = t_{L2}+t_{MD}+t_{BU}+a$. The layer 2 connection delay is represented by t_{L2}, t_{MD} is the movement detection delay, t_{BU} is the location update delay, and a is the anchor point's processing time. Suppose t_{L2} and a are equivalent in MANEMO and MA-PMIP, so they can be neglected for the comparison. The movement detection is completed when an RS message is received by the AR at the new location. Thus, when employing MANEMO, a vehicle first exchanges RS/RA messages, and then sends the location update signalling to the Home Agent. We calculate t_{MD} and t_{BU} of MANEMO as follows:

$$t_{MD}^{\mathrm{MANEMO}} = 2n\tau, \tag{4.13}$$

$$t_{BU}^{\mathrm{MANEMO}} = 2n\tau + RTT_{\mathrm{AR\text{-}HA}}, \tag{4.14}$$

where τ corresponds to the delay between transmission and reception of a data packet in the wireless domain. τ depends on the propagation delay δ, the link speed C, and the access delay due to contention T_w. The round-trip-time between AR and Home Agent, $RTT_{\mathrm{AR\text{-}HA}}$, considers the time it takes to exchange BU/BA messages.

Conversely, when MA-PMIP (Basic) is employed, the MAG triggers a location update as soon as the RS is received. Nonetheless, we have to consider the extra delay imposed by the authentication mechanism between source and relay vehicles. Thus, the delays are expressed by:

$$t_{MD}^{\text{MA-PMIP (Basic)}} = n\tau + AUTH, \tag{4.15}$$

$$t_{BU}^{\text{MA-PMIP (Basic)}} = RTT_{\text{MAG-LMA}} + n\tau, \tag{4.16}$$

where $AUTH = 2\tau + 2(T_k + T_e)$. $AUTH$ considers the delays for key generation, T_k, and for encryption/decryption, T_e. Moreover, when MA-PMIP with predictive handovers is employed, packets have been redirected to the new location during the handover. Hence, the reception of packets is resumed immediately after the movement detection is completed. Consequently, the delay calculations are derived as follows:

$$t_{MD}^{\text{MA-PMIP}} = n\tau + AUTH, \tag{4.17}$$

$$t_{BU}^{\text{MA-PMIP}} = 0. \tag{4.18}$$

4.5.3 Numerical Results

Numerical results are obtained based on the values presented in Table 4.1. The service area residence times are assumed to follow an exponential distribution [11]. Fig. 4.6 shows that MA-PMIP and MA-PMIP (Basic) achieve less location update cost compared with MANEMO. In particular, Fig. 4.6a shows that the difference among the schemes becomes larger for increasing values of ρ, i.e., for longer residence times compared with the session length.

However, a different behavior is observed when ρ becomes extremely large. In such a case, the longer session lengths dominate compared with mobility (Fig. 4.6b), and the three schemes tend to reduce the location update cost. It is also observed that the reduced packet losses in the predictive MA-PMIP, come at the cost of a nearly 30% increase of

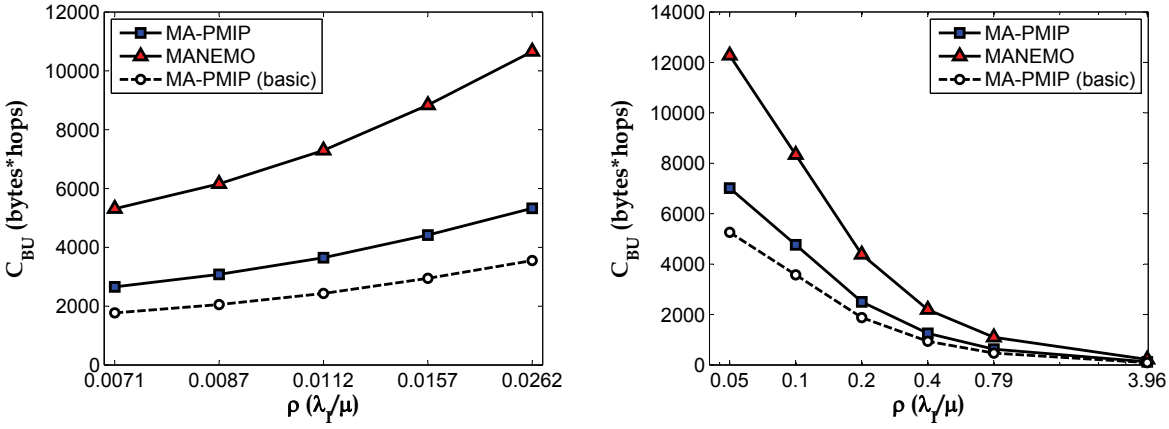
Table 4.1: MA-PMIP Cost and Handover Delay Parameters

Parameter	Value	Parameter	Value
D	700m	A	490Km^2
ω	2	n	2
$1/\lambda_I$	10s–800s	N	50
d_{HA}	3hops	d_{LMA}	3hops
d_{server}	8hops	d_{MAGs}	1hop
U^{MANEMO}	124bytes	$U^{\text{MA-PMIP}}$	124bytes
P	124bytes	v	30Km/h–110 Km/h
H	40bytes	β	5%
m	5hosts	γ	150Kbps–1Mpbs
S	1024bytes	$\delta + S/C$	2.5ms
T_w	0ms–5ms	$RTT_{\text{AR-HA}}$	10ms
$RTT_{\text{MAG-LMA}}$	10ms	$RTT_{\text{PMAG-NMAG}}$	10ms
T_k	3μs	T_e	2μs

location signalling cost when compared with MA-PMIP (Basic).

We study the impact of different session lengths (packets) in the packet delivery cost. Different downloading data rates and sessions arrival rates are considered for this study. Fig. 4.7 shows how the packet delivery cost naturally increases for longer data sessions. However, MA-PMIP still outperforms MANEMO with a reduced cost. Based on the same figure, it is observed that the packet overhead introduced by the prediction mechanism is almost equivalent to the basic MA-PMIP. This is because only a percentage of packets are affected by the double encapsulation when the MAG-to-MAG tunnel is employed.

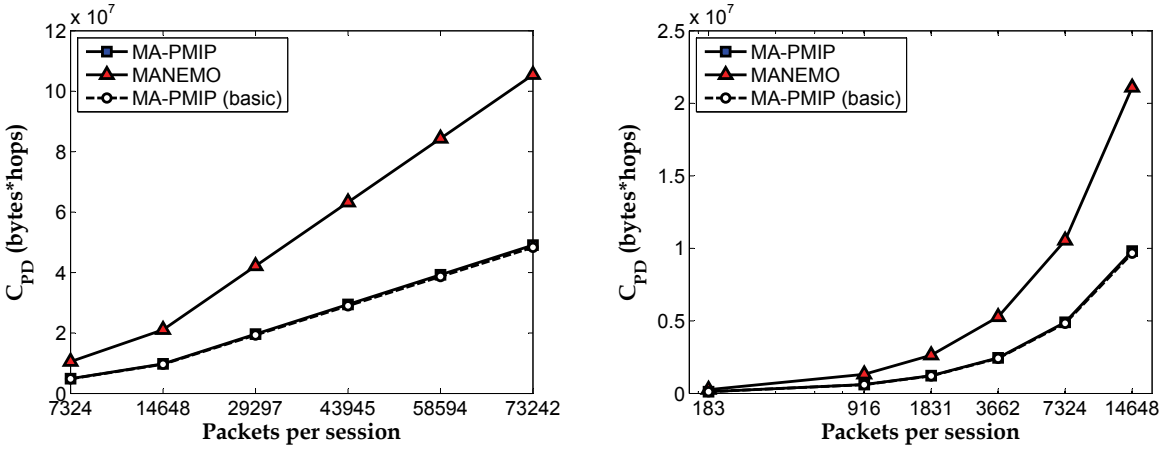
Furthermore, we calculate the total cost gain as $C_T(\text{MANEMO})/C_T(\text{MA-PMIP})$. Since the results obtained in Fig. 4.6b shows a decreasing difference among the different location update costs, we employ the cost gain to demonstrate that even when the three schemes behave similar for large values of ρ, the total reduction in cost is still dominated by the



(a) Different velocities $\frac{1}{\lambda_I}$=600s and v=110Km/h–30Km/h

(b) Different session lengths v=50Km/h and $\frac{1}{\lambda_I}$=800s–10s

Figure 4.6: MA-PMIP Location Update Comparison



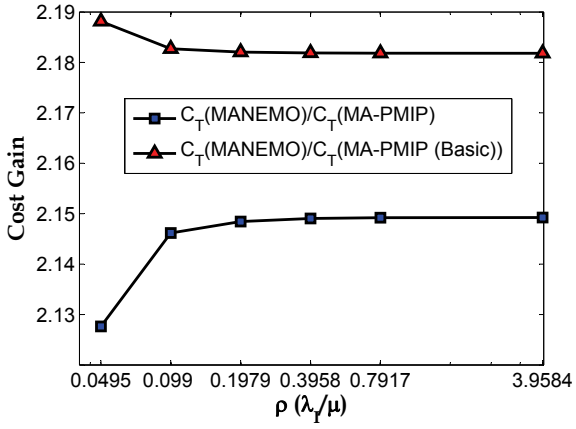
(a) Different rates γ=200Kbps–1Mbps, S=300B, and $\frac{1}{\lambda_I}$=600s

(b) Different session lengths γ=150Kbps $\frac{1}{\lambda_I}$=800s–10s, and S=300B

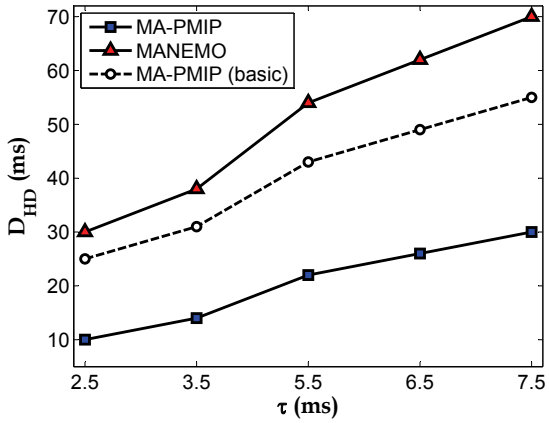
Figure 4.7: MA-PMIP Packet Delivery Comparison

reduced packet delivery overhead. Fig. 4.8a illustrates that the gain becomes stable when ρ becomes large. Both MA-PMIP and MA-PMIP (Basic) achieve less than half of the total cost of MANEMO.

Although we introduce additional signalling for authenticating the handovers through



(a) Cost gain comparison, $\frac{1}{\lambda_I}$=800s–10s, v=50Km/h, S=300B, and γ=150Kbps



(b) Handover Delay, T_w=0ms–5ms and $\delta + S/C$=2.5ms

Figure 4.8: MA-PMIP Cost gain and Handover delay

I2V2V communications, the MA-PMIP handover delay remains lower than the one in MANEMO, even though the latter does not consider any authentication mechanism. This behavior is illustrated in Fig. 4.8b. It is observed that the additional signalling employed by MA-PMIP with predictive handovers (Fig. 4.6) significantly reduces the handover delay (Fig. 4.8b). Thus, the reception of data packets is resumed near 2 times faster than in MA-PMIP (Basic), and 2.3–3 times faster than in MANEMO.

4.6 Experimental evaluation

We have performed OMNeT++ [73] simulations to corroborate the analytical evaluation presented in Section 4.5. The MiXiM [93] and INET [94] packages are used for simulating wireless communications and the TCP/IP stack, respectively. We have implemented the MA-PMIP and MA-PMIP (Basic) schemes, which are compared with the implementations of MANEMO [33] and the standard PMIP [40].

4.6.1 Proof of concept

A simulation scenario where the ARs are evenly distributed over a road segment, and the vehicle of interest is moving at a constant average speed v_r, is employed for our proof of concept [33]. The vehicle connects through 1-hop and 2-hop paths with the infrastructure, in order to download IP packets from an external application server. In every handover, we consider the worst-case scenario in which every time the vehicle joins a new AR, it first connects through a relay. Hence, the MA-PMIP's authentication is performed before the forwarding of Router Solicitation is completed.

Nodes in the VCN consume a transmission power near 1/10 smaller than the one consumed by ARs, which conveys the asymmetric links between vehicles and ARs. In a free-space path loss environment, the values employed for transmission power lead to radio ranges near 150m and 500m, for vehicles and ARs, respectively. Moreover, we employ the Two-Ray Interference model —an measurement-based enhanced version of the Two Ray ground propagation model for VANETs [95]— for the simulation of radio wave propagation.

The simulation and road traffic parameters are provided in Tables 4.2 and 4.3, respectively. Although we employ a generic 802.11 technology in our simulations, MA-PMIP is agnostic to the WLAN technology employed at the MAC/PHY layer. Furthermore, the downstream throughput and handover delay are evaluated considering three types of traffic: Constant Bit Rate (CBR) bidirectional, Variable Bit Rate (VBR) bidirectional, and VBR unidirectional. The first two account for applications that require a bidirectional link; CBR represents best-effort traffic (low-to-medium data rate), such as Internet browsing or emails fetching, and VBR represents more demanding applications with medium-to-high data rates. Unidirectional VBR traffic requires only a one-way connection for the delivery of UDP packets after the session has been established, such as in video streaming. All simulation results are plotted with the 95% confidence interval.

The throughput comparisons are shown in Fig. 5.17. The performance observed in Fig. 4.9a shows that MA-PMIP (Basic) and MANEMO are almost equivalent in the case of

Table 4.2: MA-PMIP Simulation Parameters

PHY Layer	Frequency 2.4GHz Link rate 5.5Mbps Tx power 2.3mW/25mW (vehicles/AR) Antennas' height 1.5m/3m (vehicles/AR) Sensitivity -80dBm
MAC Layer	802.11 ad hoc mode RTS/CTS disabled SNR threshold 2.6dB
Geo-routing Layer	Beacon rate 1pkt/s Geo-header size 12B
Network Layer	Router Adv rate uniform(0.5s,1.5s)
Application Layer	Bidirectional CBR (best-effort) γ=150Kbps Bidirectional VBR (video-conferencing) γ=384Kbps Unidirectional VBR (streaming) γ=512Kbps VBR σ_γ=0.010s Packet sizes 300B/1024B (CBR/VBR) Session length 600s
Prediction mode	κ=0.875 threshold=4s bufferSize=30KB$\sim$1MB
Network connections	$RTT_{MAG\text{-}LMA}$=10ms $RTT_{PMAG\text{-}NMAG}$=10ms $RTT_{AR\text{-}HA}$=10ms $RTT_{LMA(HA)\text{-}IP\ Server}$=20ms

CBR traffic. This is because with low data rates (1pkt/16ms), the handover delay in the two schemes becomes almost transparent to the flow of packets. However, the extended coverage of the link vehicle→AR, provided by the geo-networking layer, allows for a longer reception of packets and a reduction of 27% of packet losses compared with the standard PMIP. Nevertheless, both MANEMO and MA-PMIP (Basic) suffer from packet losses as soon as the bidirectional link is lost, when the vehicle is unable to connect to a relay that may establish a link toward the infrastructure. Such problem is alleviated by the prediction

Table 4.3: MA-PMIP Road traffic Parameters

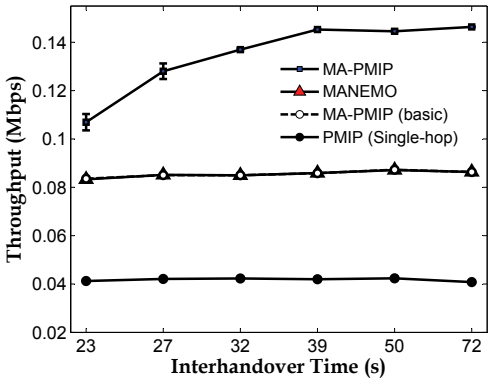
Density	k=30 veh/Km/lane k_j=120v/Km/lane
Velocity	v_r=35Km/h $\sim$ 65Km/h (urban) v_r=80Km/h $\sim$ 110Km/h (highway)
Free-flow speed	v_f=50Km/h (urban) v_f=100Km/h (highway)
Road type	Straight road – two lanes
AR inter-distance	1000m

feature in MA-PMIP. Since packets are buffered at the new location, MA-PMIP allows for a near lossless reception of packets.

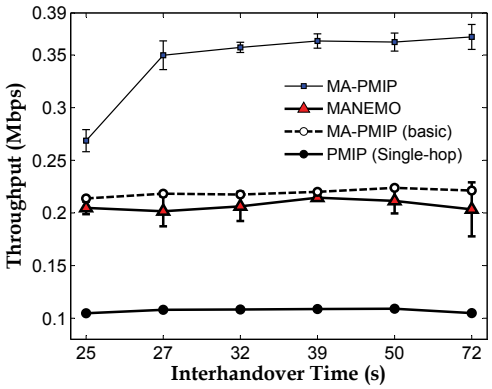
Similar results are obtained with VBR traffic. However, Fig. 4.9b and Fig. 4.9c show that, for increasing data rates, the performance of MA-PMIP (Basic) outperforms the one of MANEMO. This is mainly due to an increase of γ, which is more sensitive to the handover delay. Such handover is illustrated in Fig. 4.10, which shows the average total delay accumulated from the handovers experienced during the simulation runs. As expected, the total delay of all schemes increases with the increase in velocity. This is due to the vehicle traversing the service areas at a higher rate (i.e., there are reduced residence times); hence, the signalling for handover is exchanged more often. Nevertheless, it can be observed that MA-PMIP and MA-PMIP (Basic) result in a reduced delay. MA-PMIP achieves the lowest delay thanks to the proactive signalling, which allows for the resumption of the flow of packets as soon as the Router Solicitation message is forwarded to the MAG in the new service area.

4.6.2 Buffering during predictive handovers

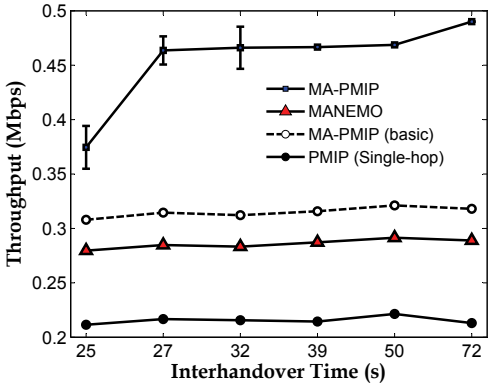
One of the salient features of MA-PMIP is the ability to forward packets in advance to the new service area where the vehicle is roaming. However, this mechanism requires to have storage space for the buffering of packets at the NMAG. Thus, we evaluate the packet



(a) Bidirectional CBR best-effort traffic, γ=150Kbps

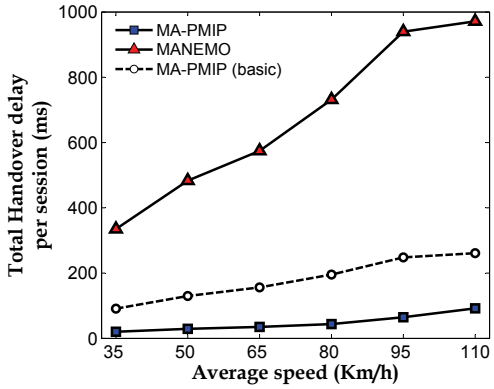


(b) Bidirectional VBR traffic, γ=384Kbps

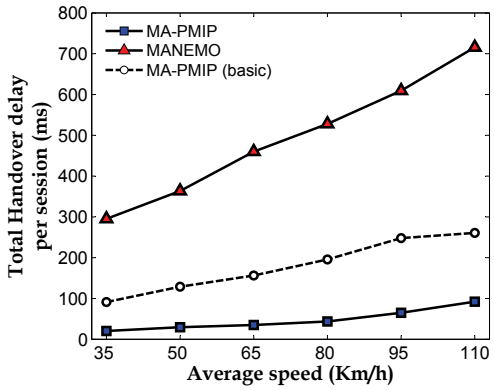


(c) Unidirectional VBR traffic, γ=512Kbps

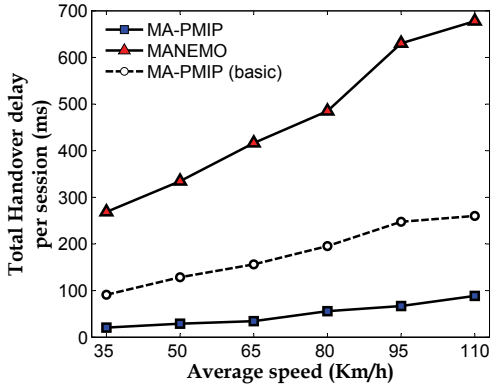
Figure 4.9: Throughput for different types of traffic vs. Inter-handover time



(a) Bidirectional CBR best-effort traffic, γ=150Kbps



(b) Bidirectional VBR traffic, γ=384Kbps



(c) Unidirectional VBR traffic, γ=512Kbps

Figure 4.10: Total handover delay for different types of traffic vs. Average speed

losses due to different buffer sizes in the NMAG, in order to have an insight of the space required for an application to perceive a lossless flow of packets. In our test scenario, the vehicle is moving at an average speed of v_r=50Km/h, and is downloading CBR best-effort traffic at a rate γ=150Kbps.

Fig. 4.11 shows the percentage of packet losses when we limit the NMAG's buffer size from 100 to 5000 packets. In our example application, a buffer of approximately 1500 packets (i.e., 450KB for packet sizes of 300bytes) would be enough to maintain seamless communications. The buffer size employed in real deployments should consider scalability issues when the density is high and several vehicles at a time trigger the predictive handover. Nonetheless, it should be considered that, for real time applications that are sensitive to delay, the predictive handover only helps on reducing the signalling after the vehicle roams to the new service area, since the buffering of real-time application packets is not applicable in this case.

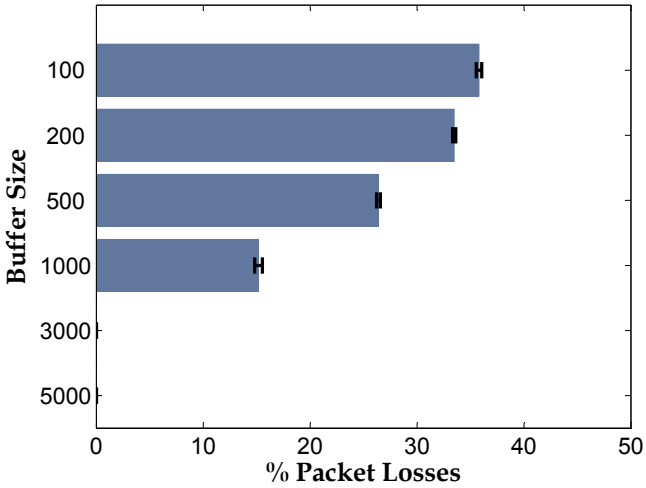


Figure 4.11: MA-PMIP packet losses due to buffer overflow.

4.6.3 A more realistic simulation scenario

After our proof-of-concept of a vehicle moving at a constant average speed, we now employ a more realistic scenario in which all nodes, i.e., vehicles and relays, are traveling at variable speeds on a two-lane highway. The velocity is controlled every Δt according to the formula:

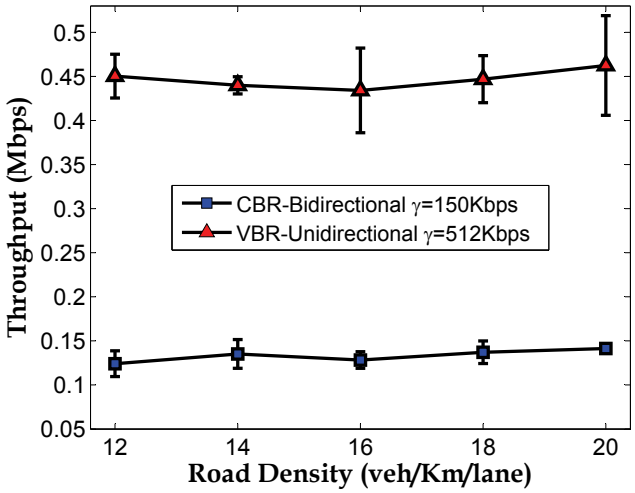
$$v(t + \Delta t) = min[max(v(t) + \Delta v, 0), v_f], \tag{4.19}$$

where $\Delta v =$`uniform`$(-a * \Delta t, a * \Delta t)$. The change of speed is given by the acceleration a, but the resulting speed is always bounded by the maximum speed of the highway v_f [96]. The details of the road traffic parameters employed in this scenario are presented in Table 4.4. We maintain the constrain of 2-hops maximum for the geo-routing layer to forward a packet in the wireless domain.

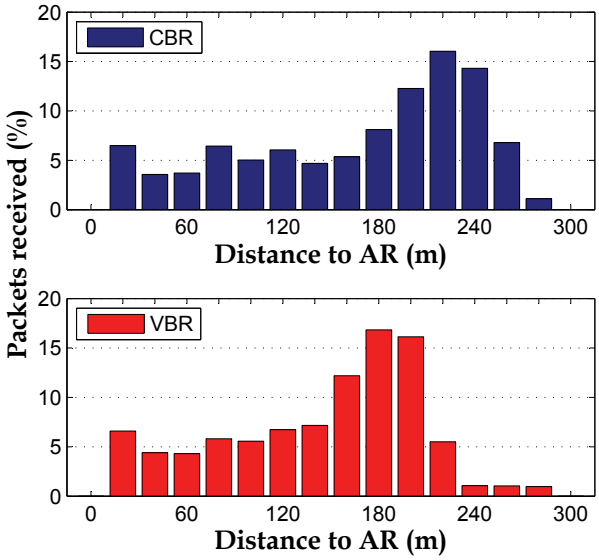
The throughput and handover delay are evaluated for IP applications with CBR bidirectional and VBR unidirectional traffic. By employing different road densities and velocities, we check the effectiveness of delivering packets when the relay selected for forwarding varies from one packet to the next one. The road densities employed during simulations are all classified as non-congested flow conditions, ranging from reasonable free-flow to stable traffic [97]. Fig. 4.12a shows that MA-PMIP still achieves a throughput close to the original downloading data rate γ. We can observe that even when the density plays an important

Table 4.4: MA-PMIP New Road traffic Parameters

Density	k=12 $\sim$ 20 veh/Km/lane k_j=120 veh/Km/lane
Velocity	v_{initial}=80Km/h
Free-flow speed	v_f=100Km/h
Acceleration	10% $* v_f$
Change of lane	disabled
Road type	Straight road – two lanes
AR inter-distance	1000m



(a) Throughput vs. road density



(b) Packets received at different distances, density=20v/Km/lane

Figure 4.12: MA-PMIP in a realistic highway scenario

role in finding available relays to reach the infrastructure, our scheme is able to adapt to the road traffic conditions, specially because the geographical protocol takes the forwarding decision on a per-packet basis. The throughputs shown in Fig. 4.12a, which are obtained

from fully mobile and variable traffic conditions, are consistent with the results obtained in the simulated proof of concept (Fig. 5.17).

Furthermore, Fig. 4.12b illustrates the average percentage of delivered packets for different distances between the vehicle and the AR. It is observed that the majority of packets are delivered when the node is more than one-hop away from the AR (distance$> r$). This is due to the predictive mechanism, in which the buffered packets are delivered as soon as the vehicle handovers through a two-hop connection in the new service area. Since we have limited the multi-hop paths to two hops, there are no packets received for distances larger than 300m.

4.7 Summary

This chapter has presented MA-PMIP, a scheme designed for enabling secure roaming of IP applications in multi-hop vehicular environments. Different from traditional IP mobility management schemes, which consider one-hop connections to the infrastructure, we have allowed for multi-hop communications, and have enhanced the performance of Proxy Mobile IP (PMIP) accordingly. The mobility protocol is coupled with a geo-networking layer for the routing of packets, and it employs the information available in the VCN, such as geographical location and road density, to predict handover events.

Furthermore, we have taken into account the asymmetric links typically encountered in VCN, and have adapted MA-PMIP to detect an react to the directionality of links. In this way, MA-PMIP takes advantage of multi-hop paths to achieve an extended bidirectional communication between vehicles and the infrastructure. Last, but not least, we have provided an efficient authentication mechanism so that IP applications can be securely handover along different IP networks. The results obtained from analytical evaluation, and experimental simulations in realistic highway scenarios, have shown the effectiveness of MA-PMIP to maintain near lossless flows of packets for vehicles with ongoing IP sessions.

One of the main advantages of employing PMIP is that it confines the signalling to the infrastructure side. However, the main limitation is that mobility is only supported within a single PMIP domain, which in practice means mobility within a single administrative domain (e.g., one network operator). Since the VCN is expected to be heterogeneous in nature, in the following chapter we expand the mobility support to a global domain, so that IP sessions are transferable not only through dissimilar access networks but also through different administrative domains.

Chapter 5

Enabling Global Mobility for In-vehicle Networks, Commuters, and Pedestrians

5.1 Preliminaries

In Chapters 3 and 4, we have focused on multi-hop communications based on vehicles that employ other vehicles to enable seamless IP communications in the VCN. However, another type of multi-hop communications appears when the in-vehicle network consist of nodes that are, by definition, also mobile. For example, passengers traveling on a bus may access the Internet from their tablet or laptop, using the local WiFi network available inside the bus. Later, when a passenger leaves the bus, he/she should be able to transfer his/her on-going sessions to the WiFi network at the bus station, even if the latter network belongs to a different IP network (or a different network operator). Such an example belongs to the problem of providing continuous Internet access for nodes in a heterogeneous urban VCN.

An urban VCN typically involves computers and entertainment systems installed in vehicles, buses, or trains, and mobile devices being used by passengers or by people com-

muting between terminal stations. All these mobile devices have communication capabilities, and they use the Internet to access services and applications available in the public network. Thus, one of the major challenges in VCN is to enable the continuity of the communications when the node changes its Internet connection, not only across dissimilar access networks (e.g., from 3G to WiFi), but also across different administrative domains (i.e., from network operator A to network operator B).

Efficient mobility management mechanisms are required to ensure the continuity of communications in VCN. The requirements of such mechanisms are defined depending on: the extension of the area where the mobile node is moving, and the mobility profile of the node (i.e., high, medium, or low mobility). First, if nodes are moving within the same administrative domain, QoS provisioning [98], and fast handovers such as the one introduced in Chapter 4, are desired. Second, when nodes move across different administrative domains, in addition to the described requirements, the mobility management scheme should adapt to support different types of access networks and different administrative policies.

Therefore, in this chapter we discuss the design of a new hybrid scheme for global mobility management in a urban VCN. The standard mobility protocols introduced by the IETF, such as the recently updated Mobility Support in IPv6 (MIP) [24], NEMO Basic Support (NEMO BS) [25], and Proxy Mobile IPv6 (PMIP) [40], were not specifically designed for urban vehicular scenarios. MIP and NEMO BS provide global mobility support, but they tend to use suboptimal routes and to introduce an end-to-end delay that severely affects real-time applications. On the other hand, in previous chapters we have introduced adaptations to the standard PMIP protocol, for making it usable in multi-hop VCN. However, the base protocol is still limited to mobility within a single administrative domain.

Our proposed hybrid scheme aims at enabling the interworking between host-based and network-based mobility support, by means of the interaction between PMIP and the Host Identity Protocol (HIP) [99] (see Section 2.1.3). HIP by itself allows for global mobility, because it defines the signalling for end-to-end mobility when one (or both) peers experience

a change of IP address. However, we aim at taking advantage of the reduced signalling overhead when the localized mobility is managed by PMIP (i.e., when the node is moving within the same administrative domain). Consequently, the design goals of our hybrid global mobility scheme are the following:

- To allow for seamless communications when a change in the global end-to-end routing of packets is caused by nodes in the VCN moving across different administrative domains. The proposed scheme should consider handovers over heterogeneous access networks, mobile devices with multiple wireless interfaces, and nodes with different patterns of mobility (e.g., pedestrian vs. vehicular mobility).

- To reduce the signalling overhead caused by location updates in the global mobility scenario, by enabling efficient transmission of updates when nodes move inside one administrative domain, and by clustering the mobility signalling for (mobile) nodes traveling in the in-vehicle mobile network.

Furthermore, our proposed HIP/PMIP interworking scheme intends to benefit two types of users: legacy nodes that depend on the vehicle's mobile router (MR) to support mobility, and HIP-enabled nodes that manage their own end-to-end mobility. The first type of user represents any computer installed in a vehicle, train, or bus, and travels all the time attached to the same MR. It could also represent a passenger's end device (e.g., a laptop or tablet) that relies on the MR for the support of mobility inside and across different domains. The second type represents end devices of passengers (or pedestrians), which already have HIP for mobility support. The fact that these nodes support HIP gives them more independency in the network. Hence, it is possible for them to move their connections from one MR to another. For example, an HIP-enabled node may transfer an active connection from the MR in a train, to the WiFi access router at the train station. Another example is a passenger switching between two different bus routes, and transferring the IP sessions in his/her tablet, from the first bus' wireless network, to the second bus' wireless network.

In the following sections, we start by providing a brief survey of previous works that address the problem of global IP mobility in vehicular networks (Section 5.2). Next, we describe our system model (Section 5.3), and introduce our hybrid HIP/PMIP interworking scheme (Section 5.4). Finally, we provide performance analysis (Section 5.5) and simulation results (Section 5.6) that demonstrate the improvements and effectiveness achieved by our proposed scheme.

5.2 Related Work

In this section, we describe previous works that extend/improve well-known mobility management protocols for urban vehicular scenarios.

Numerous studies, based on adaptations to MIP, NEMO BS, and PMIP, are proposed to support global mobility for nodes that may eventually leave the mobile network. In MIRON [27], the mobile router uses NEMO BS, whereas the mobility-enabled nodes in the mobile network employ MIP. The mobile node configures a topologically valid Care of Address directly from the infrastructure, to avoid double tunneling of packets (i.e., one to the MR's Home Agent and another to the mobile node's home agent). An address delegation with network access authentication is employed for the care of address assignment, so that mobile nodes can trigger their own route optimization procedure in a secure manner.

A solution for enabling inter-domain handovers with PMIP is proposed in [100]. This solution introduces a new element, the iMAG, which is a normal MAG located between the two different PMIP domains. This iMAG performs a layer 3 inter-domain procedure before the layer 2 inter-domain handover is completed. Hence, by the time the mobile node completes the new L2 connection, the information has already been updated in the new domain. A similar solution that uses a tunnel between LMA's of different domains is presented in [101]. Although the two solutions enable global mobility based on PMIP, they require some pre-agreement between the administrative domains for putting in place the domain-connecting elements. Furthermore, they do not define a mechanism for clustering

the mobility signalling when a number of mobile nodes travel together in a mobile network. The latter problem is addressed in [13], where the authors propose an adaptation to PMIP for the support of mobile networks. The solution focuses on automotive scenarios, and reduces the signalling overhead caused by a number of mobility-enabled nodes traveling in the in-vehicle network. However, N-PMIP does not consider the handover of nodes across different administrative domains.

The problem of broken communications due to changes of IP addresses can also be addressed by separating identifier and locator roles in IP addresses, as proposed by HIP. Since HIP provides a mechanism to maintain the communications independent of changing the IP addresses, it is also considered a global mobility management protocol. A solution to reduce the signalling overhead of HIP in a micro-mobility scenario is presented by Novaczki *et al.* [39]. The authors introduce the Local Rendezvous Servers (LRVS), which are located in every administrative domain and have to translate the mobile node's local IP address to a globally-routable IP address. The mobile node notifies the change of local IP address to the LRVS during an intra-domain handover. Since the global IP address remains the same, no other notifications are required to be sent to correspondent nodes. Conversely, during inter-domain handovers the mobile node first registers with the LRVS in the new domain; in this way the old LRVS can temporarily redirect the packets to the new location. In the meantime, the new LRVS sends notifications to the correspondent nodes updating the location of the mobile node.

There are also proposals that combine protocols from different layers, whether to improve the performance of intra-domain handovers or to enable efficient inter-domain handovers. The protocols presented in [102, 103] show different combinations of HIP with a network layer mobility management protocol. The scheme in [102] enables a micro-mobility solution with less signalling overhead through the combination of HIP and PMIP. However, it is specifically designed for an emergency system, and it does not provide IP mobility for moving networks. On the other hand, HarMoNy [103] provides a global mobility solution that extends HIP for the support of mobile networks by means of NEMO BS.

5.3 System model

We employ the system model illustrated in Fig. 5.1 for nodes moving in a urban VCN. The infrastructure provides the Internet connectivity, and consists of several overlapping heterogeneous access networks. Such networks comprise different areas of coverage, and belong to different administrative domains. In our study, we consider two overlapping access networks: a wide-area wireless network, such as 3G or WiMAX, that covers the entire city, and a WLAN network, such as 802.11b/g/n/p, that provides connection in limited areas.

The infrastructure follows the loose coupling architecture [104], where the WLAN islands do not have a direct connection to the wide-area wireless network. Thus, communications between the overlapping networks happen indirectly through a third party, in this case the Internet. The WLANs are considered low-cost access networks. On the contrary, a wide area wireless network, although it has a better coverage, it also has a higher cost per Kilobyte transmitted/received. Hence, mobile users may choose to connect through the low-cost access networks whenever possible.

Road-side Access Routers (ARs) are available for mobile devices and vehicles to access the Internet. The different network operators support PMIP within the administrative domains. We consider one single LMA per domain, and fixed tunnels from the LMA to each MAG, so that they remain active even when there are no active connections in a given MAG. Nodes obtain access across different administrative domains after an authentication method grants them with such an access. This authentication is performed while the node is establishing the layer 2 connection in a new domain.

The mobile networks are represented by different transportation systems that carry several passengers at a time, such as buses, trains, and private vehicles. Each vehicle is equipped with an on-board unit that has one or more wireless interfaces for connecting to external networks, and WLAN interface to serve as the MR for the in-vehicle network. On the other hand, mobile nodes correspond to mobile end devices, such as laptops, smart

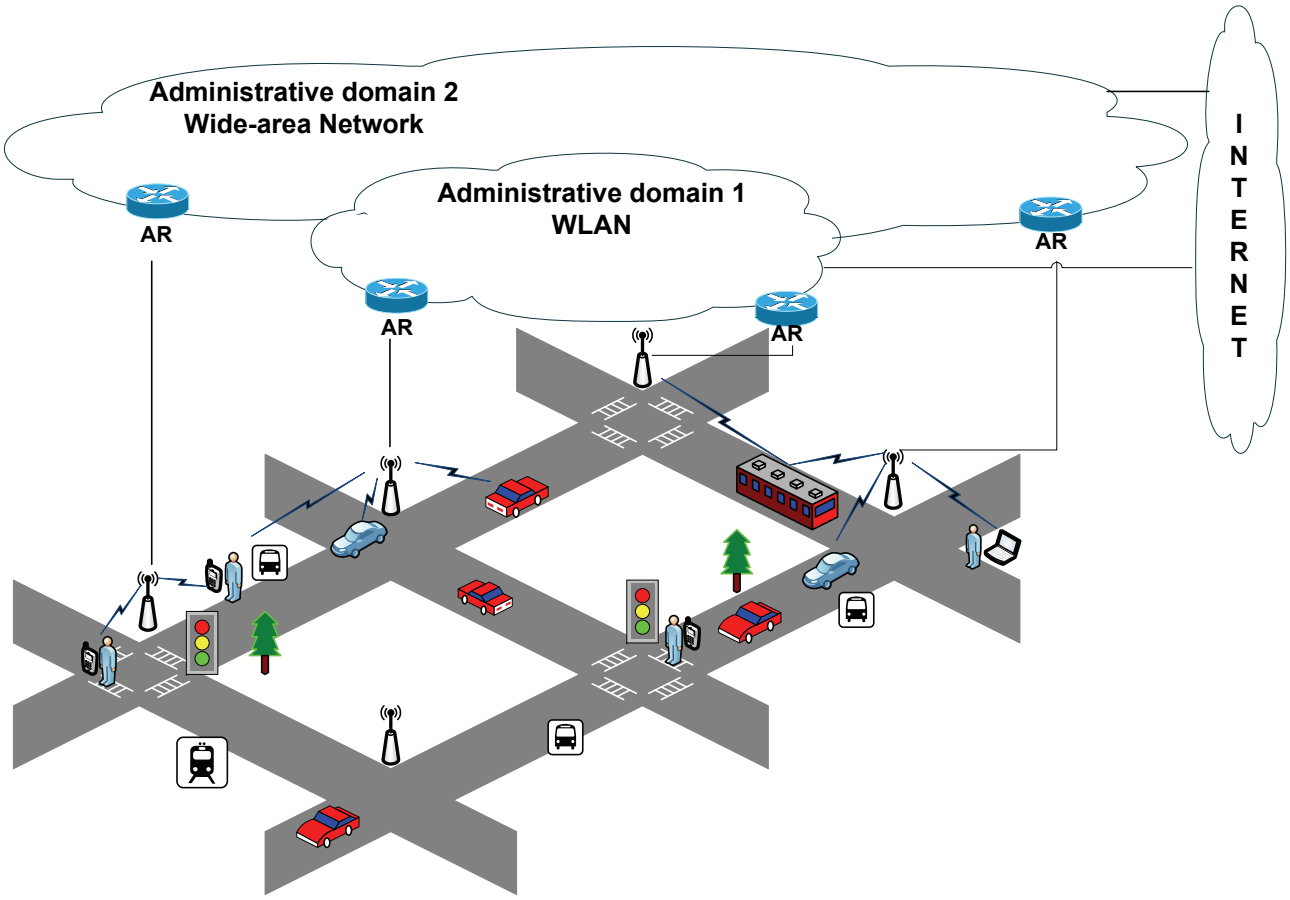


Figure 5.1: Global mobility scheme system model

phones, and tablets, which belong to passengers, drivers, or people commuting between terminal stations. These mobile nodes may connect directly to ARs in the infrastructure, or to vehicles' MRs. The mobile end devices may also have one or more wireless interfaces, although we consider only one active interface at all times.

Nodes in the VCN communicate with correspondent nodes (CN) arbitrarily located in the Internet. These CNs are HIP-enabled or located behind a proxy HIP. Domain name servers (DNS) are available for translating Full Qualified Domain Names (FQDN) to host identities, and from host identities to IP addresses [105]. Rendezvous servers (RVS) are available for redirecting initial solicitations of HIP associations when the mobile node's

location is unknown by the correspondent node. The two servers may be co-located, although this is not strictly necessary.

5.4 Hybrid HIP/PMIP interworking scheme

In this section, we describe the operation of our proposed hybrid HIP/PMIP interworking scheme. We start by describing the initialization phase required by nodes entering the administrative domains for the first time. Next, we describe how end-to-end packets delivery is performed, followed by the description of the signalling required for both intra and inter-domain handovers. In each phase, we include an explanation of the signalling required for legacy mobile nodes (i.e., nodes with no IP mobility support) and HIP-enabled nodes.

5.4.1 Initialization

An illustration of the initialization phase of the hybrid HIP/PMIP scheme is depicted in Fig. 5.2. When an MR enters to a PMIP domain for the first time, it initially follows the regular steps defined in the standard PMIP for new associations [40]. During the layer 2 connection to the serving MAG, the MR completes the authentication procedures in the new network. Next, the MAG notifies the detection of a new connection to the LMA by means of a PBU message. The PBU includes the MR's unique identifier, which is used by the LMA to detect that it corresponds to a new node in the network.

Once the LMA checks that this is the first time the MR registers in the domain, it proceeds to assign it a home network prefix, and to send a PBA back to the MAG. The MAG then advertises the network prefix to the MR in a Router Advertisement (RA) message, and the MR configures an address based on the received home network prefix. In parallel, the MR continuously sends RA messages to nodes in the in-vehicle network. The RAs announce a unique local IPv6 unicast prefix, which allows the nodes to configure

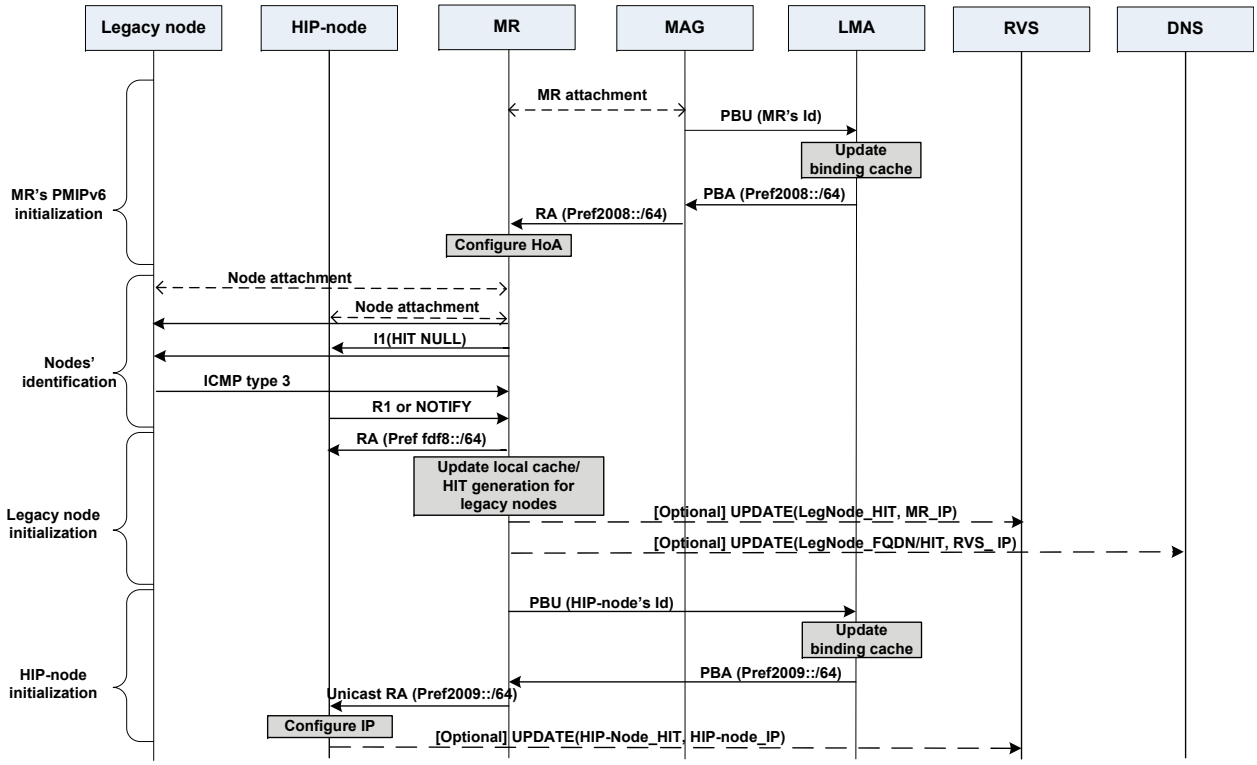


Figure 5.2: Hybrid HIP/PMIP Initialization phase

globally unique addresses that are intended for local communications [106]. All nodes in the in-vehicle network configure addresses from the local unicast prefix.

After the initialization is completed, the MR identifies if there are HIP-enabled nodes in the in-vehicle network. In order to do this identification, the MR sends I1 messages in opportunistic mode (i.e., an I1 with a NULL destination HIT). Only the HIP-enabled nodes will respond to that message, whether with an R1 or a NOTIFY packet [1]. Nodes that are not HIP-enabled will reply with an ICMP destination protocol unreachable packet. Subsequently, the MR completes the initialization in a different way, depending on whether or not mobile node support HIP. The procedures are described as follows.

[1] A NOTIFY reply is sent when the node does not allow for opportunistic mode.

Initialization for legacy nodes

The MR acts as a proxy HIP for the identified legacy nodes. The proxy HIP generates a Host Identity Tag (HIT) for each legacy node, and places this information in a local cache. The cache relates the HIT to the unique local IPv6 address of the legacy node. At this point, the legacy nodes may initialize access to the Internet. As an optional step, the MR may send an UPDATE message to the RVS (`[LegNode_HIT→MR_IP]`), and to the DNS (`[LegNode_FQDN→LegNode_HIT→RVS_IP]`), on behalf of each legacy node. In this way, incoming communications from correspondent nodes to legacy nodes are also enabled.

Initialization for HIP-enabled nodes

The MR acts as a mobile MAG (mMAG) for mobile nodes that have been identified as HIP-enabled [13]. A PBU is sent from the MR to the LMA indicating the unique identifier of the HIP-enabled node, and the LMA sends back a PBA with the IP prefix assigned to the mobile node. The information about the HIP-enabled node is stored in the LMA's binding cache. The stored entry includes the node's identifier, the assigned IP prefix, the serving MAG (i.e., the mMAG), and a flag to indicate the serving MAG is mobile. This flag is necessary to perform recursive lookups when there is incoming traffic directed to the HIP-enabled node, as we later explain in Section 5.4.2.

After completing the PMIP signalling, the MR announces the network prefix in a unicast RA message to the HIP-enabled node [107]. Upon receiving the RA, the node configures an IP address from the new prefix and selects it as the source address for external communications [108]. However, the node also keeps the address initially configured from the local unicast prefix. At this point, HIP-enabled nodes may initialize access to the Internet. An additional UPDATE message, `[HIP-node_HIT→HIP-node_IP]`, can be sent from the HIP-enabled node to the RVS, in order to enable incoming communications. No updates need to be sent to the DNS.

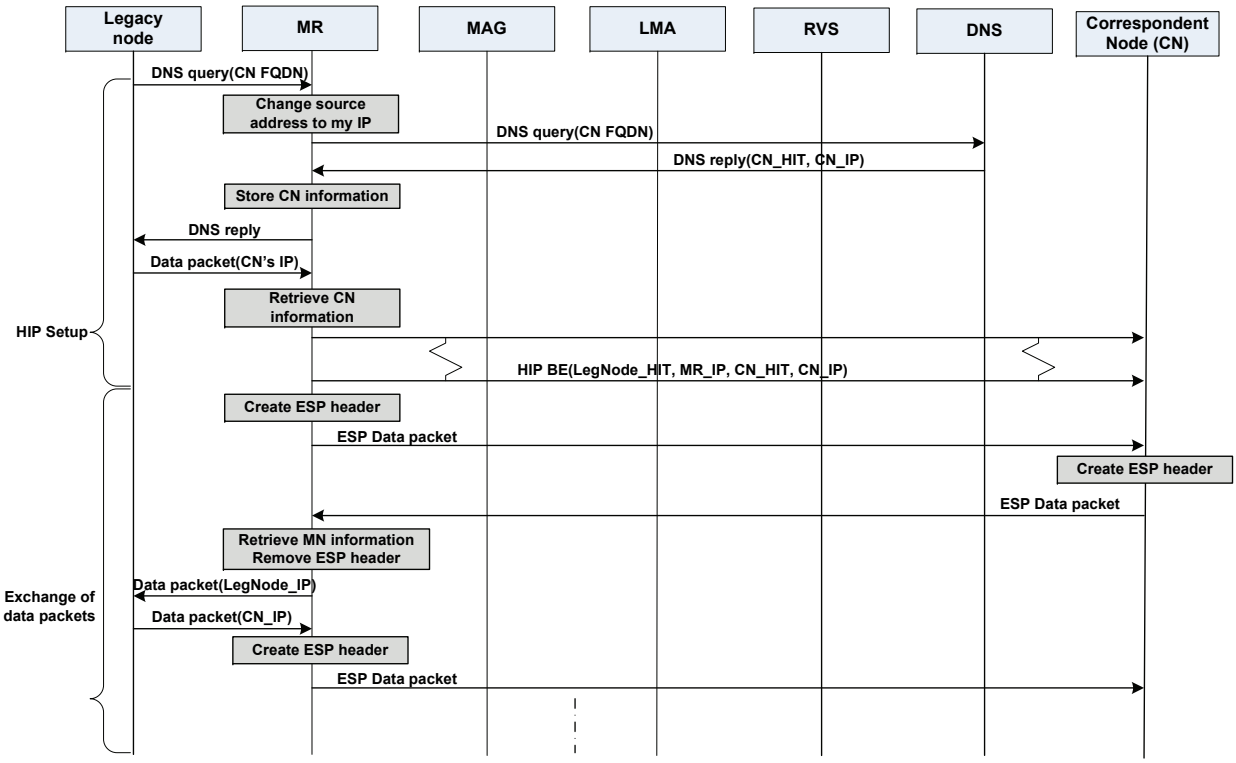


Figure 5.3: End-to-end communications between legacy nodes and correspondent nodes

5.4.2 End-to-end communications

Data packets to/from the Internet are forwarded in a different way depending on the type of node that is transmitting/receiving the packets in urban VCNs. The two procedures are explained below.

Communications from/to legacy nodes

The end-to-end communication between a legacy node and a correspondent node is illustrated in Fig. 5.3. When the legacy node communicates with a correspondent node in an external network, it first sends a DNS query to translate the correspondent node's FQDN to an IP address. The proxy HIP in the MR then intercepts this query, and replaces the packet's source address with its own IP [109]. Once the MR receives a reply from the DNS,

it inspects the packet and stores the correspondent node's information (i.e., the HIT and IP address). The reply packet is then forwarded to the legacy node.

Upon receiving the first legacy node's data packet to be forwarded outside the in-vehicle network, the MR starts an HIP base exchange with the correspondent node. This is a four-way handshake in which the MR and correspondent node establish the required HIP security associations. Consequently, the MR removes the IP header of each packet received from a legacy node, and generates a new header using the Encapsulating Security Payload (ESP) transport format. This new header includes the MR's IP as the packet's source address. When packets arrive from the infrastructure, the MR looks for the correspondent security association, and once it locates the HIT-IP association in its local cache, it removes the packet's ESP encapsulation and forwards it to the legacy node.

Communications from/to HIP-enabled nodes

The end-to-end communication between an HIP-enabled mobile node and a correspondent node is illustrated in Fig. 5.4. Since HIP-enabled nodes manage their communications autonomously, they do not require any action from the MR other than the forwarding of packets. Before transmitting the first data packet, the HIP-node performs the HIP base exchange with the correspondent node. It then encapsulates the packets using the ESP format and forwards them through the outgoing security association. As for the MR, when it receives an ESP-protected packet, it simply forwards it in the proper direction after identifying the packet's destination address.

5.4.3 Intra-domain handovers

Intra-domain handovers involve the change of connection to another AR/MR located in the same administrative domain (i.e., inside the PMIP domain). The procedures for intra-domain handovers for both types of mobile nodes are depicted in Fig. 5.5 and described below.

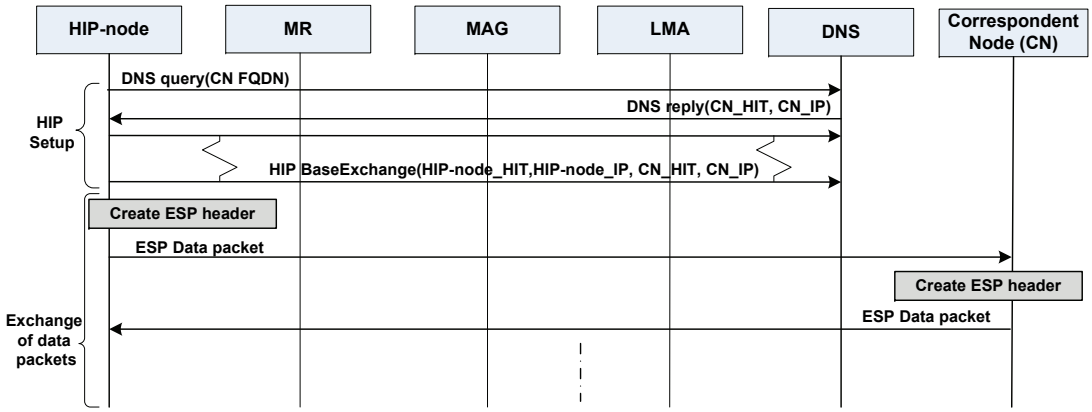


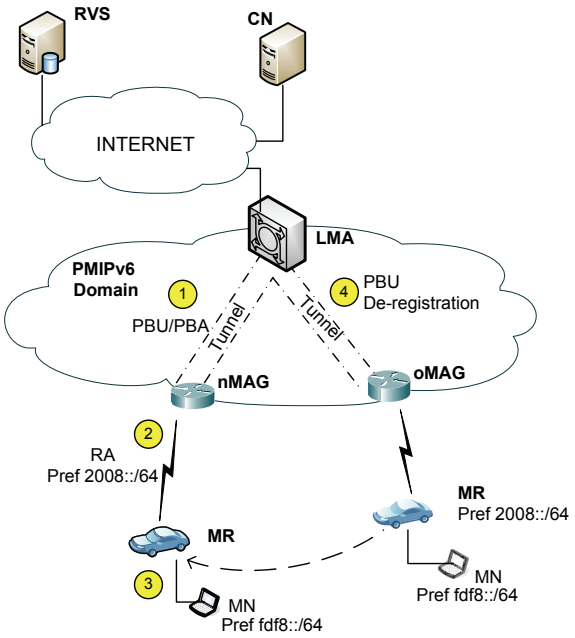
Figure 5.4: End-to-end communications between HIP-enabled nodes and correspondent nodes

Intra-domain handovers for legacy nodes

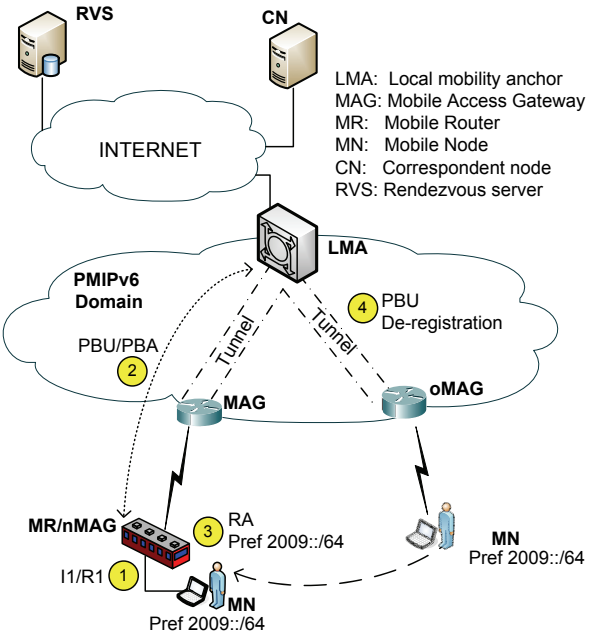
The process of intra-domain handovers for legacy nodes is illustrated in Fig. 5.5a. An intra-domain handover should be the result of a movement of the MR (the one serving the legacy node) to a new AR in the same domain. When this movement occurs, the PMIP functionalities are activated, so that the new MAG detects the new connection and proceeds with the notification to the LMA (Fig. 5.5a-1).

Once the LMA receives the PBU sent by the MAG, it recognizes the MR has been already registered in the domain, and it maintains the same home network prefix assignment. When the new MAG receives the PBA, it announces the same network prefix to the MR (Fig. 5.5a-2). Thus, the MR does not perceive any changes at the network layer. As for the legacy node, the local unicast prefix announced by the MR does not changes (Fig. 5.5a-3), so the intra-domain handover is transparent to the node.

Given that the MR's IP remains the same, the MR does not need to update any of the active HIP sessions. This involves no notifications to correspondent nodes, nor to the RVS or DNS.



(a) Legacy node intra-domain handover



(b) HIP-enabled node intra-domain handover to a new MR

Figure 5.5: Intra-domain handover in Hybrid HIP/PMIP scheme

Intra-domain handovers for HIP-enabled nodes

There are several cases in which an HIP-enabled node may experience an intra-domain handover. The least complex cases are: a) when the vehicle where the HIP-node is located moves the connection to a new AR; and b) when the HIP-node itself moves its connection to a new AR (e.g., a passenger leaving a train and joining the network at the train station). In these cases, the signalling is the same as for the intra-domain handover of a legacy node (Fig. 5.5a).

A more complex situation appears when the HIP-enabled node switches the connection to another MR (e.g., a passenger switching between two bus routes). This process is illustrated in Fig. 5.5b. When the HIP-node joins the network of the new MR, the MR first performs the identification process described in Section 5.4.1. Once the R1 or NOTIFY packets are received as a response from the node (Fig. 5.5b-1), the new MR exchanges the PMIP signalling with the LMA (Fig. 5.5b-2). Since the node has been already registered in the domain, the LMA assigns the same network prefix to it, and the MR proceeds to advertise such a prefix to the node (Fig. 5.5b-3). Once again, none of the active HIP sessions has to be updated, since the HIP-enabled node does not perceive any changes at the network layer.

5.4.4 Inter-domain handovers

Inter-domain handovers involve the change of connection, whether from the node or the MR, to a point of attachment that belongs to a different PMIP domain. The procedures for inter-domain handovers are depicted in Fig. 5.6 and described below.

Inter-domain handovers for legacy nodes

The process of inter-domain handovers for legacy nodes is illustrated in Fig. 5.6a. An inter-domain handover is the result of the MR (the one serving the legacy node) roaming

to a new PMIP domain. When this occurs, the new MAG and LMA exchange the standard PMIP signalling (Fig. 5.6a-1).

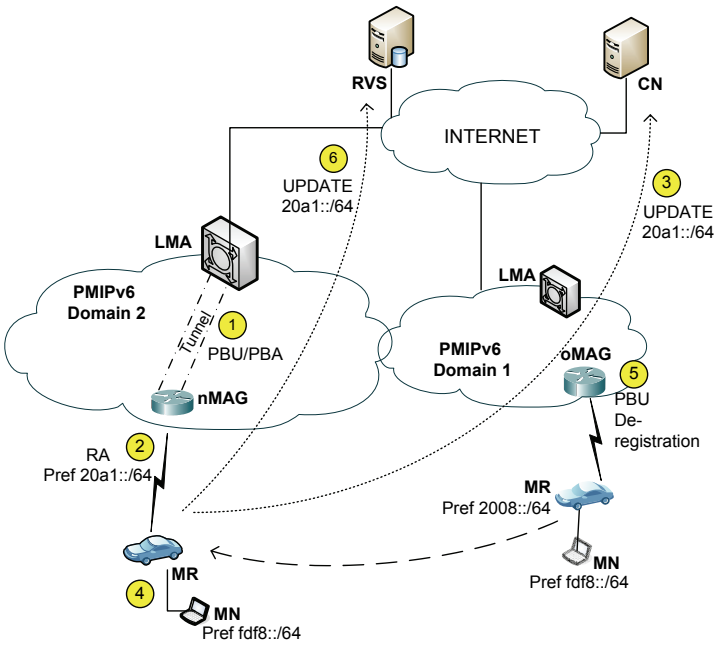
The LMA registers the MR upon reception of the PBU, and proceeds to assign a home network prefix to it (Fig. 5.6a-1). Next, the MAG announces the prefix to the MR (Fig. 5.6a-2). At this point, the MR detects the change of IP network, and starts updating the active HIP communications. Thus, the MR sends UPDATE message to correspondent nodes for which active security associations exist. The UPDATE indicates the newly acquired IP address as the new locator (Fig. 5.6a-3). In the meantime, the legacy node keeps the same local IP address; hence, it does not detect any changes at the network layer (Fig. 5.6a-4). The MR may also send an UPDATE message to the RVS, on behalf of each legacy node, in order to enable incoming communications at the new location (Fig. 5.6a-6).

We employ the Credit-Based Authorization mechanism [110], which allows the correspondent node to securely use the new locator as soon as it receives the UPDATE message. Although the peer's reachability at the address embedded in the locator has not yet been verified, with such an authorization both sides can immediately start using the new address for active communications. Nonetheless, the verification of the new address is later completed with two more UPDATE packets exchanged between the MR and correspondent node, but this verification does not affect the continuity of current communications.

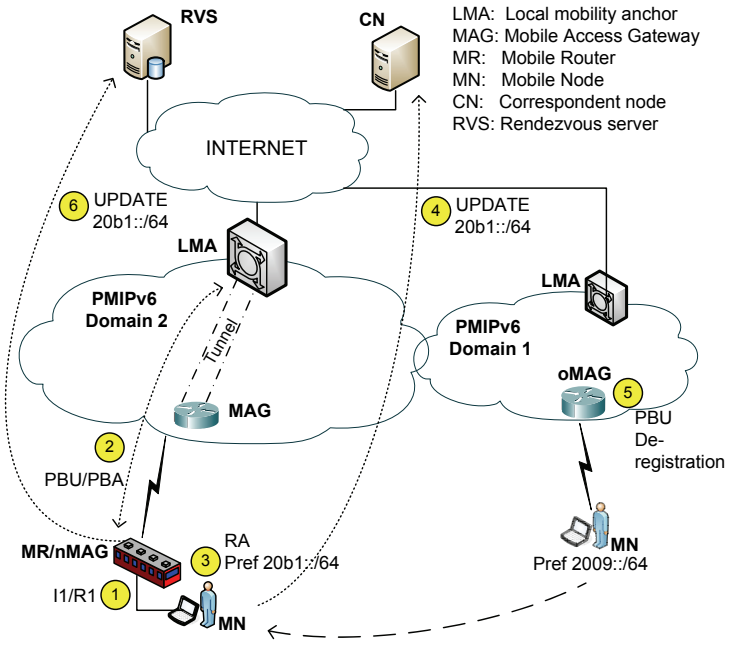
Inter-domain handovers for HIP-enabled nodes

The scenarios considered in Section 5.4.3 are also applicable for inter-domain handovers of HIP-enabled nodes. However, the difference here is that the new point of attachment belongs to a different administrative domain.

If a node transfers its connection from an AR in one domain, to an AR in another domain, the signalling is exactly the same as the one described for inter-domain handovers of a legacy node; except that the update of active sessions is done by the node itself. On the other hand, if the connection is transferred to an MR in a different domain, the MR then



(a) legacy node inter-domain handover



(b) HIP-enabled node inter-domain handover to a MR

Figure 5.6: Inter-domain handover in Hybrid HIP/PMIP scheme

advertises the new IP prefix to the HIP-enabled node, and the HIP-enabled node updates its IP address accordingly (Fig. 5.6b-3). Subsequently, the node sends UDPATE messages for each active security associations established with correspondent nodes (Fig. 5.6b-4). The node may also send an UPDATE message to the RVS, in order to enable new incoming communications (Fig. 5.6b-6).

5.5 Performance Analysis

We evaluate the proposed scheme from the point of view of the in-vehicle mobile network, and from the legacy nodes and HIP-enabled nodes perspective. In the mobile network case, we extend the crossing probability calculation introduced in Section 4.5.1, and quantify the generated signalling load according to the the definition in Section 4.5 for *location update cost* and *packet delivery overhead cost*. The mobile nodes performance is evaluated based on two different criteria: *handover delay* and *expected number of dropped packets*. The latter refers to the expected number of packets the MN is unable to transmit due to the handover process.

5.5.1 Mobile network analysis

The in-vehicle network mobility is described according to a fluid flow model. Using the model, we then calculate the crossing rate at which a vehicle transitions across different ARs (i.e., intra-domain handovers), and across different PMIP domains (i.e., inter-domain handovers). The bases for our mobile network analysis are described as follows:

- A PMIP domain is composed by N subnets. Each subnet is served by one AR. For simplicity, each area covered by an AR is assumed to be square-shaped with perimeter P_S and area A_S.

- The mobility pattern of each mobile network is described with an average velocity v and a uniformly distributed direction in the range $[0, 2\pi]$.
- The intra-domain crossing rate for a mobile network is given by $\mu_{intra} = vP_s/(\pi A_s)$. The inter-domain crossing rate for a mobile network is given by $\mu_{inter} = \mu_{intra}/\sqrt{N}$.
- The subnet residence times and domain residence times are given by general distributions, $f_{intra}(s)$ and $f_{inter}(s)$, with mean $1/\mu_{intra}$ and $1/\mu_{inter}$, respectively. Likewise, the CDF and Laplace transform for the residence time distributions are given by $F_{intra}(s)$, $F_{inter}(s)$, and $f^*_{intra}(s)$ and $f^*_{inter}(s)$, respectively.
- Only incoming data sessions are considered. A data session is assumed to have an average length L (packets). Each node in the mobile network has independent and identically distributed session arrival rates.
- Inter-session arrival times for the mobile network are assumed to be exponentially distributed with rate λ_I.
- N_{intra} and N_{inter} determine the number of subnet crossings and domain crossings, respectively, during an inter-session arrival time. The probabilities of i subnet crossings $P(N_{intra} = i)$, and j domain crossings $P(N_{inter} = j)$ are given by [44]:

$$P(N_{intra} = i) = \alpha(i) = \begin{cases} 1 - \frac{1}{\rho_{intra}}[1 - f^*_{intra}(\lambda_I)] & \text{if } i = 0, \\ \frac{1}{\rho_{intra}}[1 - f^*_{intra}(\lambda_I)]^2[f^*_{intra}(\lambda_I)]^{i-1} & \text{if } i > 0. \end{cases} \tag{5.1}$$

$$P(N_{inter} = j) = \beta(j) = \begin{cases} 1 - \frac{1}{\rho_{inter}}[1 - f^*_{inter}(\lambda_I)] & \text{if } j = 0, \\ \frac{1}{\rho_{inter}}[1 - f^*_{inter}(\lambda_I)]^2[f^*_{inter}(\lambda_I)]^{j-1} & \text{if } j > 0, \end{cases} \tag{5.2}$$

where $\rho_{intra} = \lambda_I/\mu_{intra}$ and $\rho_{inter} = \lambda_I/\mu_{inter}$.

- The following notations are used to indicate packet/header sizes: a is the size of BU/BA (PBU/PBA) message in NEMO BS (PMIP), b is the size of an IP tunnel header, c is the size of an ESP header, and d is the size of an UPDATE message in HIP.

- The distances between network elements (i.e., number of intermediate hops) are represented in Fig. 5.7a by $d1, d2, d3, d4,$ and $d5$. Although the wireless links only have a one-hop distance, a weight factor ω is included to indicate their high cost compared to the links in the infrastructure.

- Given m legacy nodes and n mobility-enabled nodes in the mobile network, the total signalling cost is calculated as follows:

$$C_T(m,n) = C_{BU}(m,n) + C_{PD}(m,n), \tag{5.3}$$

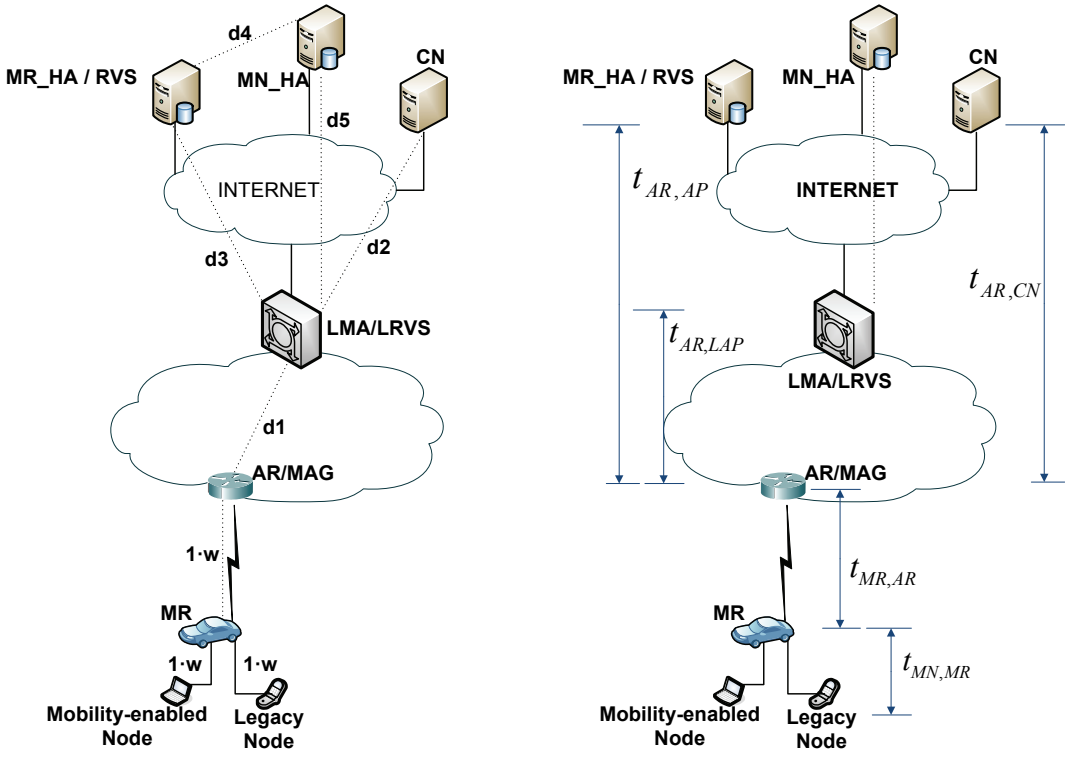
  where $C_{BU}(m,n)$ is the average location updates signalling cost during an inter-session arrival time, and $C_{PD}(m,n)$ is the total packet delivery overhead cost incurred during the same period. Moreover, $C_{BU}(m,n)$ is calculated as follows:

$$C_{BU}(m,n) = \sum_{j}\sum_{i} C_{BU}(m,n|i,j)\cdot\alpha(i)\cdot\beta(j). \tag{5.4}$$

The mobile network performance of our proposed Hybrid HIP/PMIP scheme is compared with the standard NEMO BS. There are two types of nodes in NEMO BS: Local Fixed Nodes (LFN) and Visitor Mobile Nodes (VMN). LFN rely on the MR for the support of mobility, whereas VMN use their own home agent to register the changes of location.

Location updates in NEMO BS

According to the signalling defined in the standard NEMO BS [44], the location udpate cost is computed by:



(a) Mobile network performance analysis **(b)** Mobile node performance analysis

Figure 5.7: Hybrid HIP/PMIP scheme performance analysis

$$C_{BU}^{\text{NEMO}}(m,n) = \sum_{i} C_{BU}^{\text{NEMO}}(m,n|i) \cdot \alpha(i), \tag{5.5}$$

$$C_{BU}^{NEMO}(m,n|i) = i \cdot B^{\text{MR-HA}}, \tag{5.6}$$

where $B^{\text{MR-HA}} = a \cdot (w + d1 + d3)$. Note that NEMO BS does not have the concept of domains. Moreover, when the MR performs a handover, only its own care-of-address changes, and therefore none of the local nodes have to update their locations. Consequently, there is additional signalling cost from LFNs or VMNs.

Location updates in the Hybrid HIP/PMIP scheme

When the mobile network performs an intra-domain handover, there is an exchange of PBU/PBA messages to maintain the network prefix assigned to the MR. On the other hand, when an inter-domain handover occurs, the MR has to additionally notify the change of address to the correspondent nodes, on behalf of legacy nodes. Similarly, HIP-enabled nodes also update the correspondent nodes about the new location. Consequently, the location update cost for HIP/PMIP, given that i subnets and j domains are crossed, is calculated as follows:

$$C_{BU}^{\text{HYBRID}}(m,n|i,j) = i \cdot B^{\text{MAG-LMA}} + j \cdot (B^{\text{MAG-LMA}} + m \cdot (B^{\text{MR-CN}} + B^{\text{MR-RVS}})) \\ + n \cdot j \cdot (B^{\text{MR-LMA}} + B^{\text{MN-CN}} + B^{\text{MN-RVS}}), \tag{5.7}$$

where $B^{\text{MAG-LMA}} = a \cdot d1$, $B^{\text{MR-CN}} = d \cdot (w + d1 + d2)$, $B^{\text{MR-RVS}} = d \cdot (w + d1 + d3)$, $B^{\text{MR-LMA}} = a \cdot (w + d1)$, $B^{\text{MN-CN}} = d \cdot (2 \cdot w + d1 + d2)$, and $B^{\text{MN-RVS}} = d \cdot (2 \cdot w + d1 + d3)$. Note that we have included the optional updates to the RVS, in order to enable incoming communications to the mobile network after a handover.

Packet delivery overhead in NEMO BS

According to [44], the packet delivery cost of NEMO BS is calculated as follows:

$$C_{PD}^{\text{NEMO}}(m,n) = L \cdot \left(\frac{m}{m+n} \cdot C^{\text{LFN-PD}} + \frac{n}{m+n} \cdot C^{\text{VMN-PD}} \right), \tag{5.8}$$

where $C^{\text{LFN-PD}} = b \cdot (d3 + d1 + w)$ and $C^{\text{VMN-PD}} = b \cdot d5 + 2b \cdot (d3 + d1 + w) + b \cdot w$. Packets destined to a VMN require an extra tunnel between the MR's home agent and the MR.

Parameter	Value	Parameter	Value
m	3	L_s	2800m
n	2	A_s	490Km^2
$1/\lambda_I$	400s–900s	N_{intra}	20
γ	150Kbps	Packet size	512Bytes
$d1$	3hops	N_{inter}	4
$d2$	8hops	a	124bytes
$d3$	4hops	b	40bytes
$d4$	8hops	c	20bytes
$d5$	4hops	d	80bytes
w	2	v	30Km/h–65Km/h

Table 5.1: Parameters for Hybrid HIP/PMIP mobile network analysis

Packet delivery overhead in the Hybrid HIP/PMIP scheme

The packet delivery overhead of HIP/PMIP is derived as follows:

$$C_{PD}^{\text{HYBRID}}(m,n) = L \cdot \left(\frac{m}{m+n} \cdot C^{\text{LegNode-PD}} + \frac{n}{m+n} \cdot C^{\text{HipNode-PD}} \right), \tag{5.9}$$

where $C^{\text{LegNode-PD}} = c \cdot d2 + (c+b) \cdot d1 + c \cdot w$. Packets destined to legacy nodes travel directly from the correspondent node to the PMIP domain, with an extra tunnel added between the LMA and the serving MAG. When the MR receives a packet, it removes the ESP encapsulation and forwards a normal IP packet to the legacy node. On the other hand, the packet delivery overhead for an HIP-enabled node is $C^{\text{HipNode-PD}} = c \cdot d2 + (c + 2b) \cdot d1 + (c+b) \cdot w + c \cdot w$. In this case, an extra IP tunnel is employed to forward packets to the mobile MAG. Also, the ESP encapsulation is removed only when the packet arrives to the HIP-enabled node.

Mobile network numerical results

The values used to quantify the equations for the mobile network analysis are specified in Table 5.1. $f_{\text{intra}}(s)$ and $f_{\text{inter}}(s)$ follow a Gamma distribution with mean $1/\mu_{\text{intra}}$ and $1/\mu_{\text{inter}}$, respectively. The γ parameter of the Gamma distribution is given by $\gamma = 1/(V_{\text{intra}}\mu_{\text{intra}}^2)$ and $\gamma = 1/(V_{\text{inter}}\mu_{\text{inter}}^2)$. Althoguh variances of the intra and inter-domain residence distributions are assumed to be $V_{\text{intra}} = 1/\mu_{\text{intra}}^2$ and $V_{\text{inter}} = 1/\mu_{\text{inter}}^2$, respectively, the impact of other calculated variances can be found in [44].

The Laplace transforms to be used in (5.1) and (5.2) are then given by:

$$f^*_{\text{intra}}(\lambda_I) = \left(\frac{\mu_{\text{intra}}\gamma}{\lambda_I + \mu_{\text{intra}}\gamma}\right)^{\gamma} = \left(\frac{\mu_{\text{intra}}}{\lambda_I + \mu_{\text{intra}}}\right). \tag{5.10}$$

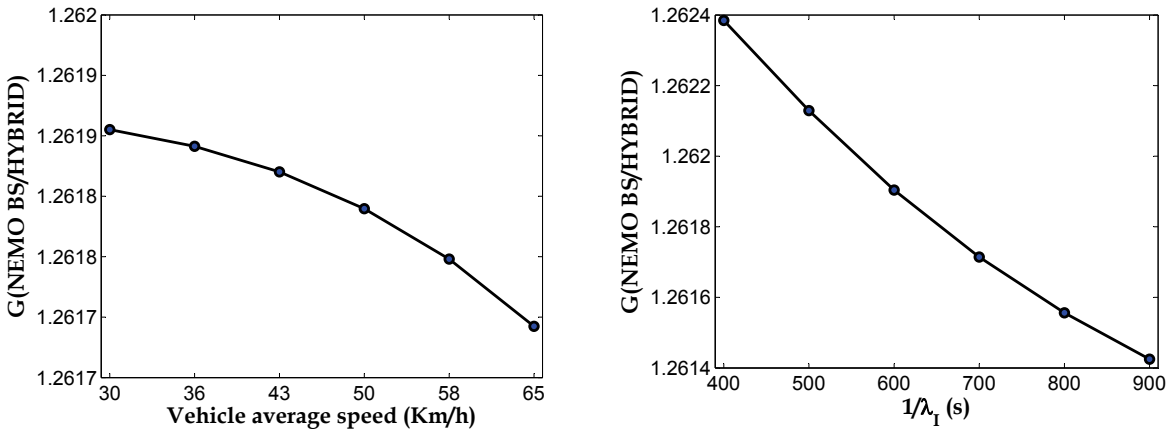
$$f^*_{\text{inter}}(\lambda_I) = \left(\frac{\mu_{\text{inter}}\gamma}{\lambda_I + \mu_{\text{inter}}\gamma}\right)^{\gamma} = \left(\frac{\mu_{\text{inter}}}{\lambda_I + \mu_{\text{inter}}}\right). \tag{5.11}$$

Furthermore, $C_T^{\text{NEMO}}(m,n)$ is obtained by replacing (5.5) and (5.8) in (5.3). In a similar way, $C_T^{\text{HYBRID}}(m,n)$ is obtained by replacing (5.7) and (5.9) in (5.3).

To compare both schemes, the gain G is defined as the total relative cost gain:

$$G = \frac{C_T^{\text{NEMO}}(m,n)}{C_T^{\text{HYBRID}}(m,n)} \tag{5.12}$$

Fig. 5.8a and Fig. 5.8b show the impact of different average speeds and different session lengths, respectively. In both scenarios, a vehicle with 5 in-vehicle network nodes ($m = 3, n = 2$) is employed. The average speeds are set according to measurements in urban scenarios [111]. Due to limitations in the fluid flow model, it is not possible to describe "stop-and-go" patterns due to traffic lights, which are normally seen in urban



(a) G for different urban average speeds

(b) G for different inter-session arrival times

Figure 5.8: Cost gain analysis of NEMO BS vs. Hybrid PMIP/HIP scheme

roads. However, given the average speeds, the analysis helps understand the advantages of using our hybrid scheme instead over employing the standard NEMO BS.

As observed in both figures, although the gain decreases for increasing speeds or inter-session arrival times, the reduced amount is small, which makes the scheme to outperform NEMO BS almost with a constant gain. The decreasing gain observed in Fig. 5.8a is caused by the increased vehicular mobility, which triggers more inter-domain handovers. Our hybrid scheme, in comparison to NEMO BS, has a costly location update process because it involves updates to each correspondent node. A similar effect is observed in Fig. 5.8b by considering longer session lengths. However, the high location update cost of our hybrid scheme is compensated by the low overhead packet delivery cost. In our scheme, packets go directly between correspondent node and LMA, as opposed to the packet delivery in NEMO BS. Therefore, on average, packets traverse less hops in the hybrid scheme than in NEMO BS.

5.5.2 Mobile nodes analysis

The *handover delay* and the *expected number of dropped packets* are derived separately for legacy and HIP-enabled nodes. The notations used for the analysis are illustrated in Fig. 5.7b and explained as follows:

- T_{HDI}= Total handover delay for intra-domain handover.
- T_{HDE}= Total handover delay for inter-domain handover.
- P_{HDI}= Expected number of dropped packets during an intra-domain handover.
- P_{HDE}= Expected number of dropped packets during an inter-domain handover.
- T_{L2HD}= Layer 2 handover delay. The time between the node's disconnection from the AR, and the layer 2 connection to a new point of attachment. It includes the time for authenticating the node in the new network.
- $t_{\mathrm{MR,AR}}$= Time required to transmit a packet from the MR to the road-side AR.
- $t_{\mathrm{MN,MR}}$= Time required to transmit a packet from the MN to the MR.
- $t_{\mathrm{AR,LAP}}$= Time required to transmit a packet from the road-side AR to the Local Anchor Point (for instance an LMA or LRVS) located in the same domain.
- $t_{\mathrm{AR,CN}}$= Time required to transmit a packet from the road-side AR to a node in the Internet.
- $t_{\mathrm{AR,AP}}$= Time required to transmit a packet from the road-side AR to an Anchor Point (for instance a HA).
- a= Processing time due to the updating of a local binding cache.

In this analysis, we compare our scheme with four additional protocols that also provide global mobility support: MIPv6 [24], NEMO BS [25], HIP [99], and Novaczki's micro-mobility solution for HIP [39]. The bases for our mobile node analysis are described as follows:

- All wireless links are symmetric.
- For simplicity, we consider the mobile node is communicating only with one correspondent node at the moment of handover.
- The layer 2 handover delay is the same for all the compared protocols.
- The movement detection at the network side is given by the reception of a RS message. Nodes detect a change of network when they receive RA messages as a response to the solicitation.
- The time required to generate an IP address is negligible. This generation should happen right after the node receives an RA message.
- We do not consider the time for Duplicate Address Detection (DAD) in any of the compared schemes.
- In the HIP-related protocols, including our hybrid scheme, we do not perform rekeying of the security associations after a change of IP address.
- It is assumed the mobility-enabled nodes are able to obtain IP addresses directly from the infrastructure. This means that, if the node is connected through an MR, the MR forwards the RS/RA messages between the MN and the AR to allow for IP address configuration. This assumption holds for MIPv6, HIP, and Novackzi's scheme. The assumption does not hold for our proposed scheme, since the mobile MAG already allows for this configuration. In addition, we only consider the case in which a mobility-enabled node handovers to an MR, since in our scheme this is the worst case signalling load scenario.

In general, the handover delay T_{HD} comprises the layer 2 handover delay, the movement detection delay, the IP addressing configuration, and location update delay. The derivation of this metric is explained below.

Handover delay in MIPv6/NEMO BS

MIPv6 and NEMO BS work in a similar manner. The former supports single nodes, whereas the latter supports mobile networks. Consequently, NEMO BS is employed for legacy nodes moving in the in-vehicle network, whereas MIPv6 is employed by mobility-enabled nodes. NEMO BS requires an update to be sent to the home agent every time the MR experiences a handover. We consider the HA to be arbitrarily located in the Internet, hence T_{HD}^{NEMO} is expressed as follows:

$$T_{HD}^{\text{NEMO}} = T_{\text{L2HD}} + 2t_{\text{MR,AR}} + 2(t_{\text{MR,AR}} + t_{\text{AR,HA}}) + a_{\text{HA}}. \tag{5.13}$$

Similarly, MIPv6 requires the node to update the home agent whenever it acquires a new care-of-address. Moreover, MIPv6 defines an optimized version in which the node is able to notify the change directly to the correspondent node. Therefore, we calculate T_{HD}^{MIPv6} as follows:

$$T_{HD}^{\text{MIPv6}} = T_{\text{L2HD}} + 2(t_{\text{MN,MR}} + t_{\text{MR,AR}}) + 2(t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,CN}}) + a_{\text{CN}}. \tag{5.14}$$

Since NEMO BS and MIPv6 are not limited to domains, there is no separated calculation for intra and inter-domain handovers.

Handover delay in standard HIP

When an HIP node travels in the in-vehicle network, it expects the MR to announce the change of IP addresses every time the vehicle roams to a different IP network. Thus, after

the node reconfigures its address, it has to send an UPDATE to the correspondent node. As a result, $T_{\text{HIP-a}}$ is calculated as follows:

$$T_{HD}^{\text{HIP-a}} = T_{\text{L2HD}} + 2t_{\text{MR,AR}} + t_{\text{MN,MR}} + (t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,CN}}) + a_{\text{CN}}. \tag{5.15}$$

On the other hand, when the HIP node transfers a connection to an MR, it updates the correspondent node right after acquiring the new IP address. Thus, $T_{HDI}^{\text{HIP-b}}$ is calculated as follows:

$$T_{HD}^{\text{HIP-b}} = T_{\text{L2HD}} + 2(t_{\text{MN,MR}} + t_{\text{MR,AR}}) + (t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,CN}}) + a_{\text{CN}} \tag{5.16}$$

Since the standard HIP is not limited to domains, there is no separated calculation for intra and inter-domain handovers.

Handover delay in Novaczki's scheme

In this scheme, when the node performs an intra-domain handover, it updates the new location only with the LRVS [39]. The improvement to the normal HIP is given by the fact that no updates have to be sent to correspondent nodes. As a result, the calculations for handover delay differ from (5.15) and (5.16) only in the destination for the UPDATE message, as indicated below:

$$T_{HD-\text{intra}}^{\text{NOV-a}} = T_{\text{L2HD}} + 2t_{\text{MR,AR}} + t_{\text{MN,MR}} + (t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,LRVS}}) + a_{\text{LRVS}}, \tag{5.17}$$

$$T_{HD-\text{intra}}^{\text{NOV-b}} = T_{\text{L2HD}} + 2(t_{\text{MN,MR}} + t_{\text{MR,AR}}) + (t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,LRVS}}) + a_{\text{LRVS}}. \tag{5.18}$$

Novaczki's scheme is more complex for inter-domain handovers. Given that the LRVS operates as the anchor point and address translator for mobile nodes, every time a node moves to a different domain, it has to register with a new LRVS. A registration with the LRVS is an HIP base exchange (i.e., a four-way handshake). Additionally, once the MN finishes the registration at the new domain, it updates the previous LRVS with the information of its new location. In this way, the old LRVS can redirect the incoming packets to the new domain. The old LRVS is used as a temporary relay only while the new LRVS completes the updates to the correspondent nodes.

When the mobile node sends the UPDATE message to the old LRVS, we consider the old LRVS to be arbitrarily located in the Internet. As a result, the inter-domain handover for Novaczki's scheme is calculated as follows:

$$\begin{aligned} T_{HD-\text{inter}}^{\text{NOV-a}} &= T_{\text{L2HD}} + 2t_{\text{MR,AR}} + t_{\text{MN,MR}} + 4(t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,nLRVS}}) \\ &\quad + a_{\text{nLRVS}} + (t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,CN}}) + a_{\text{oLRVS}}, \qquad (5.19) \\ T_{HD-\text{inter}}^{\text{NOV-b}} &= T_{\text{L2HD}} + 2(t_{\text{MN,MR}} + t_{\text{MR,AR}}) + 4(t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,nLRVS}}) \\ &\quad + a_{\text{nLRVS}} + (t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,CN}}) + a_{\text{oLRVS}}. \qquad (5.20) \end{aligned}$$

Handover delay in the Hybrid HIP/PMIP interworking scheme

Intra-domain handovers of legacy nodes are managed by the MR. When the in-vehicle mobile network handovers, a regular PMIP location update is performed as soon as the new MAG receives the router solicitation. As a result, $T_{HD-\text{intra}}^{\text{HYBRID-a}}$ follows the signalling presented in Fig. 5.5a, and is expressed by:

$$T_{HD-\text{intra}}^{\text{HYBRID-LegNode}} = T_{\text{L2HD}} + t_{\text{MR,AR}} + 2t_{\text{AR,LMA}} + a_{\text{LMA}} \qquad (5.21)$$

On the other hand, the intra-domain handover of an HIP-enabled node involves ad-

ditional identification signalling, given that the connection is transferred to an MR. The intra-domain handover in such a case follows the signalling presented in Fig. 5.5b, and the delay is calculated as follows:

$$T_{HD-\text{intra}}^{\text{HYBRID-HipNode}} = T_{\text{L2HD}} + 2t_{\text{MN,MR}} + 2(t_{\text{MR,AR}} + t_{\text{AR,LMA}}) + a_{\text{LMA}} + t_{\text{MN,MR}}. \quad (5.22)$$

Likewise, two different calculations are provided for inter-domain handover delay. When a legacy node moves to a different administrative domain, that means the serving MR has moved. The only difference with the intra-domain handover case is that the MR has to notify the correspondent node about the change of location. The handover follows the signalling presented in Fig. 5.6a, and its delay is expressed by:

$$T_{HD-\text{inter}}^{\text{HYBRID-LegNode}} = T_{\text{L2HD}} + t_{\text{MR,AR}} + 2t_{\text{AR,LMA}} + a_{\text{LMA}} + (t_{\text{MR,AR}} + t_{\text{AR,CN}}) + a_{\text{CN}}. \quad (5.23)$$

When an HIP-enabled node transfer its connection to an MR in a different domain, the calculations are similar to ones for intra-domain handover, except that the UPDATE notification is delivered to the correspondent node. In such a case, the handover signalling is depicted in Fig. 5.6b, and the delay is calculated as follows:

$$\begin{aligned} T_{HD-\text{inter}}^{\text{HYBRID-HipNode}} &= T_{\text{L2HD}} + 2t_{\text{MN,MR}} + 2(t_{\text{MR,AR}} + t_{\text{AR,LMA}}) + a_{\text{LMA}} \\ &\quad + t_{\text{MN,MR}} + (t_{\text{MN,MR}} + t_{\text{MR,AR}} + t_{\text{AR,CN}}) + a_{\text{CN}}. \end{aligned} \quad (5.24)$$

Expected number of dropped packet

To quantify this metric, we count all the packets that are generated after the initial link breaks down, and until the network-layer handover is completed. Although real-time

applications are more sensitive to packets losses, non real-time applications can still be throughput-sensitive, and hence, they can also be affected by packets discarded during the handover.

This calculation of this metric employs the average downloading data rate γ. The packets losses are then calculated as follows:

$$PL_{HD-\text{intra}} = T_{HD-\text{intra}} \times \gamma \tag{5.25}$$

$$PL_{HD-\text{inter}} = T_{HD-\text{inter}} \times \gamma \tag{5.26}$$

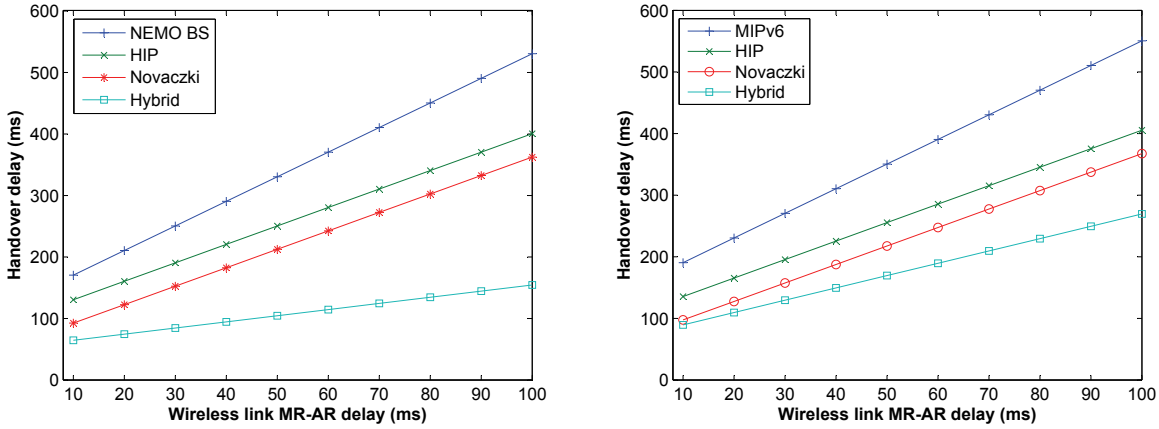
Mobile node numerical results

The parameters used for mobile node analysis are presented in Table 5.2.

Parameter	Value	Condition
T_{L2HD}	50ms	Same for all schemes
$t_{MR,AR}$	10ms	802.11 access network
$t_{MN,MR}$	5ms	54Mbps WLAN access network
$t_{AR,LAP}$	2ms	100Mbps local area network link
$t_{AR,CN}$	40ms	Average times to reach a node in the Internet
a	0.5ms	Same for anchor points of all schemes
γ	50packets/sec	UDP traffic for VoIP using G.729 codec

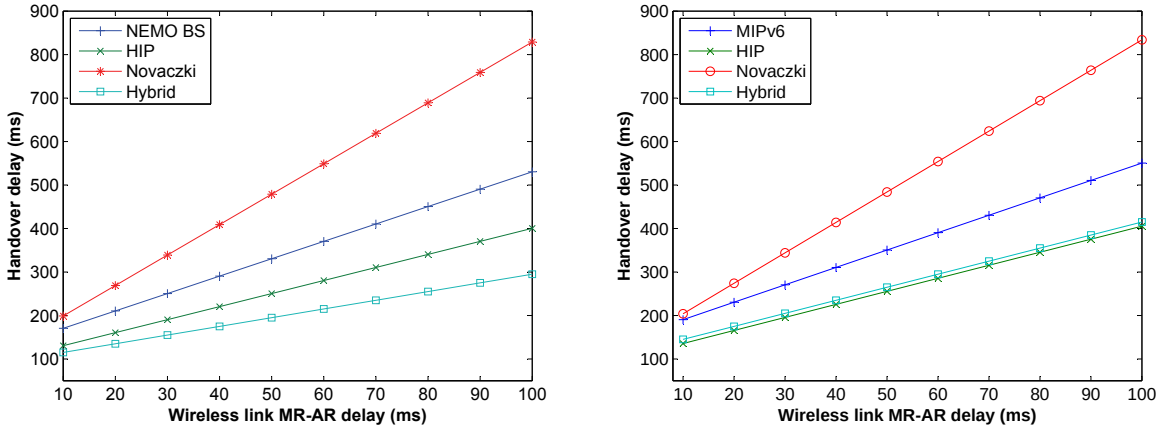
Table 5.2: Parameters for mobile node performance analysis

Fig 5.9 shows the impact of wireless access delays during intra-domain handovers. Different access delays are due to the dissimilar radio access technologies that a node may employ for accessing the Internet in the VCN. Although the handover delay is sensitive to a high-delay access network, the hybrid HIP/PMIP scheme is observed to outperform the other schemes. This is due to the fact the same network prefix is assigned when the node or the MR are moving inside the PMIP domain. The high delay experienced by the other reported schemes is the result of changing the network prefix (or care-of-address)



(a) Legacy node intra-domain handover **(b)** Mobility-enabled node intra-domain handover

Figure 5.9: Impact of wireless access delay on intra-domain handovers

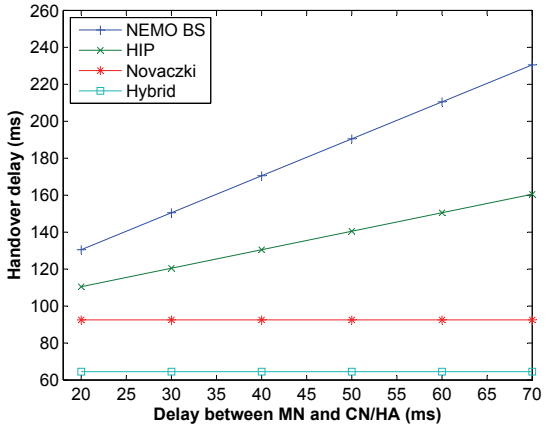


(a) Legacy node inter-domain handover **(b)** Mobility-enabled node inter-domain handover

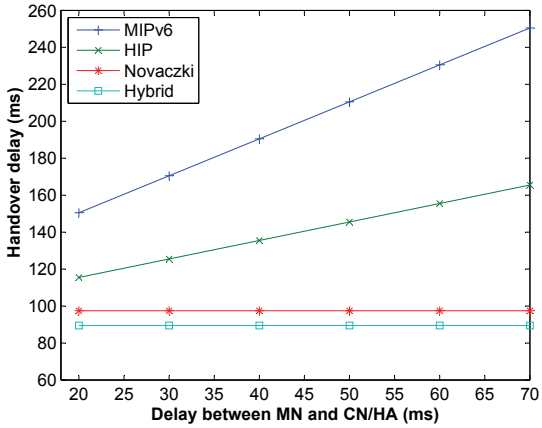
Figure 5.10: Impact of wireless access delay on inter-domain handovers

in every handover. The slight increment observed for HIP-enabled nodes in the hybrid scheme (Fig. 5.9b) is due to the node identification process.

Similarly, Fig. 5.10 illustrates the impact of wireless access delays during inter-domain handovers. It is observed that only Novaczki's and our hybrid scheme present a different behaviour compared with the intra-domain handover. The temporary use of old LRVS for redirection of packets in Novaczki's scheme, increases the handover delay to the point that

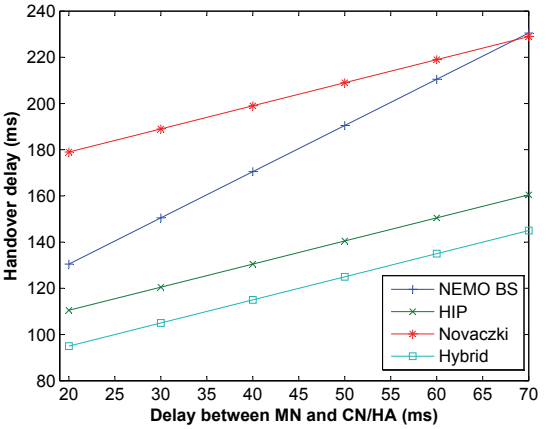


(a) Legacy node intra-domain handover

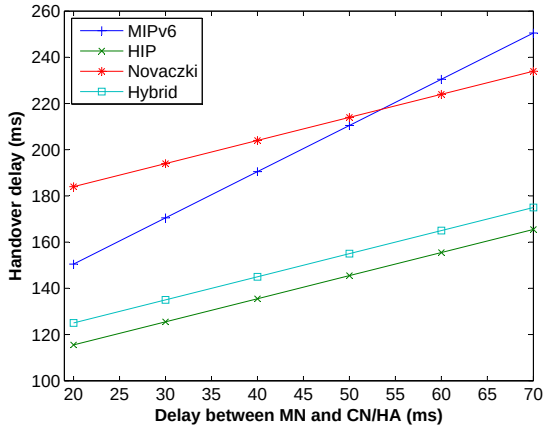


(b) Mobility-enabled node intra-domain handover

Figure 5.11: Impact of end-to-end delay on intra-domain handovers



(a) Legacy node inter-domain handover



(b) Mobility-enabled node inter-domain handover

Figure 5.12: Impact of end-to-end delay on inter-domain handovers

makes it impractical during inter-domain handovers. In the case of our hybrid scheme, it presents a performance comparable to that of HIP. The increased delay observed by HIP-enabled nodes (Fig. 5.10b) is due to the MR's exchange of PMIP signalling before being able to advertise the new prefix to the node.

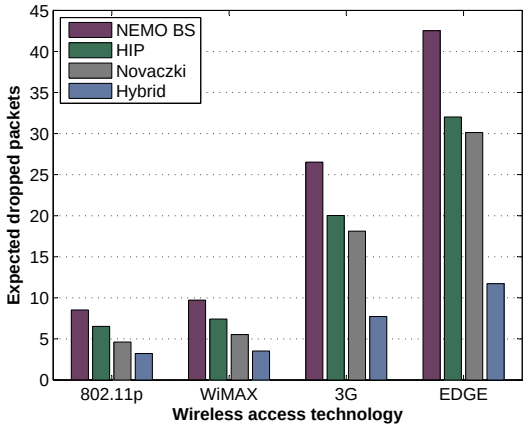
Fig 5.11 and Fig. 5.12 show the impact of different end-to-end delays between the mobile node and the correspondent node (or the home agent in the case of NEMO BS). When

the correspondent node or home agent are located far away from the VCN, the delay for end-to-end communications increases. In addition, the handover is affected because the location update takes longer. Such a behavior severly affects NEMO BS and MIPv6 schemes. On the other hand, during inter-domain handovers, we observe an increased delay of our hybrid scheme compared with HIP (Fig. 5.12b). Despite this increase, the hybrid scheme has the additional advantage of supporting legacy mobile nodes, whereas the standard HIP requires all the nodes to be HIP-enabled.

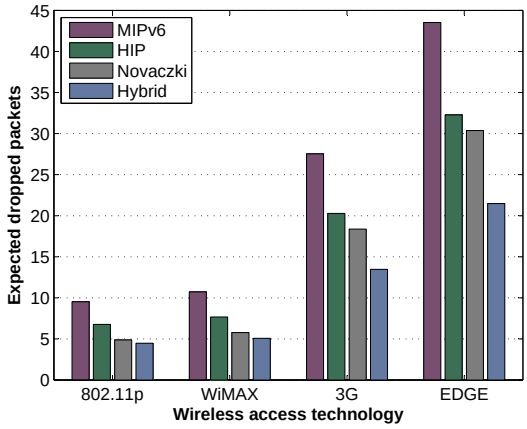
Fig. 5.13 and Fig. 5.14 illustrate the expected number of dropped packets during intra-domain and inter-domain handovers, respectively. We have selected four different radio access technologies the user (whether vehicle or pedestrian/commuter) may join at the moment of handover, which allows us to evaluate intra and inter-technology handovers. Each technology is represented by a different wireless access delay: 10ms for 802.11p, 16ms for WiMAX, 100ms for 3G and 180ms for EDGE.

The figures show that our proposed scheme achieves the best results compared to the other reported protocols, except for the scenario in Fig. 5.14b, where the hybrid scheme packet losses are comparable to HIP. The increased packet losses are the consequence of the MR's identification of HIP-enabled nodes. Nevertheless, we observe in Fig. 5.13a that the hybrid scheme may achieve as little as 3 dropped packets. These packet losses correspond to a gap of nearly 60ms according to the real-time application employed for this evaluation. Therefore, our hybrid scheme should be suitable even for highly demanding application, such as VoIP.

Our analysis highlights the following advantages: 1) the hybrid scheme achieves a reduced handover delay, which is the result of using PMIP for the localized mobility; 2) by clustering the signalling overhead from mobile nodes, even for those that are mobility-enabled, the hybrid scheme reduces the load over the MR→AR link; and 3) our interworking scheme allows for seamless communications of legacy and mobility-enabled nodes in the VCN.

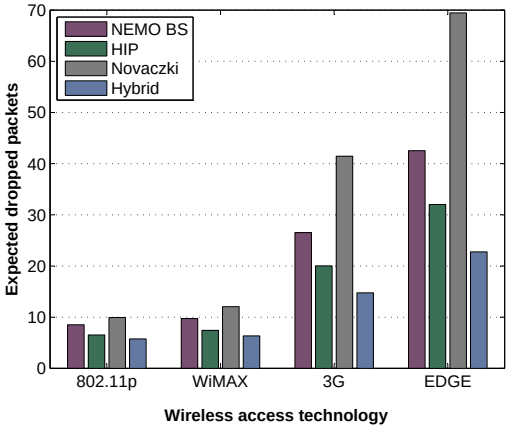


(a) Legacy node intra-domain handover

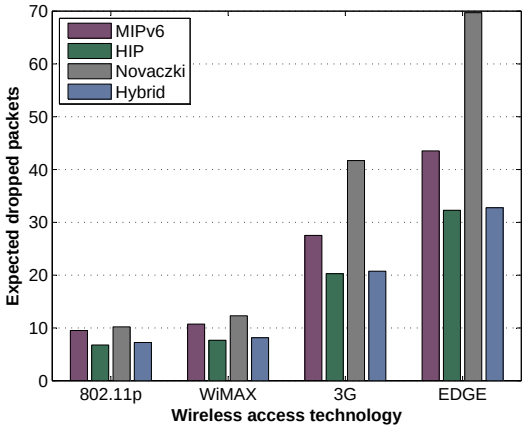


(b) Mobility-enabled node intra-domain handover

Figure 5.13: Expected number of dropped packets for intra-domain handovers



(a) Legacy node inter-domain handover



(b) Mobility-enabled node inter-domain handover

Figure 5.14: Expected number of dropped packets for inter-domain handovers

5.6 Simulation results

In order to evaluate the performance of the Hybrid HIP/PMIP scheme, we have performed simulations in a realistic urban scenario. A typical commuter is simulated traveling to his/her workplace.

Pedestrian Mobility	Minimum speed	0.7m/s
	Mean speed	1 m/s
	Speed standard deviation	0.1 m/s
	Max. pause time	10s
	Pause probability	0.15
	Speed change probability	0.1
	Turn probability	0.25
Vehicular Mobility	Minimum speed	7m/s
	Mean speed	13 m/s
	Speed standard deviation	1 m/s
	Max. pause time	20s
	Pause probability	0.15
	Speed change probability	0.1
	Turn probability	0.3
Intra-domain Set #1	p^{w-c}	0.3
	p^{w-w}	0.3
	p^{c-w}	0.3
	p^{c-c}	0.9
Intra-domain Set #2	p^{w-c}	0.7
	p^{w-w}	0.7
	p^{c-w}	0.7
	p^{c-c}	0.9
Handover delay (HD)	HD^{c-w}_{intra}, HD^{w-w}_{intra}	98ms
	HD^{c-c}_{intra}, HD^{w-c}_{intra}	290ms
	HD^{c-w}_{inter}, HD^{w-w}_{inter}	150ms
	HD^{c-c}_{inter}, HD^{w-c}_{inter}	450ms

Table 5.3: Hybrid HIP/PMIP Scheme Simulation Parameters

The commuter has an HIP-enabled mobile device, which is employed for Internet access during the journey. Initially, the commuter walks toward the nearest bus station, and from there, he/she takes a bus ride toward the destination bus stop. In the last segment, the commuter walks from the bus stop to the workplace. The total commuting time has been set to 26 minutes, according to the average travel times that Canadian commuters take for

going to work on a typical day in 2010 [112].

To recreate the city scenario, the commuter and the bus move according to the Manhattan Grid mobility model, on a grid of 4000 Km^2 and with 100m×100m-blocks that emulate the city blocks. The mobility traces are generated with the BonnMotion tool [113], and the parameters employed for generating pedestrian and vehicular traces are described in Table 5.3. We have employed the ChainScenario, provided by BonnMotion, in order to concatenate the different mobility patterns (i.e., walking–bus riding–walking) in a single 26-minute trip. During both pedestrian and vehicular movements, the node stops at random times to simulate the red traffic lights it may encounter during the journey.

Then, we calculate the residence times in every 50m×50m-cell along the path employed by the node during the simulation. The residence times are illustrated in Fig. 5.15. The figure indicates the two areas where the commuter is walking, and the rest of the movements happen during the bus ride. Note that, although the randomness in direction's selection of the Manhattan Grid model causes a few loops in the path, in general this does not affect the results obtained for dwell times.

Based on this information, we proceed to simulate the HIP/PMIP scheme in Matlab. A 3G network is assumed to cover all the simulated area, whereas WiFi hotspots provide limited coverage. The ratio of coverage of WiFi to 3G in the simulated area is indicated by Δ, which varies from 0 (only 3G coverage available) to 1 (double coverage always available). When roaming through the cells along the path, the node decides with a probability $1-\Delta$ to switch between networks. If a switching occurs, the type of intra-domain handover is determined by the transition probabilities $p^{w-w}, p^{w-c}, p^{c-w}$, and p^{c-c}, where p^{a-b} indicates a handover from technology a to technology b, w indicates WiFi, and c indicates 3G cellular network. An inter-domain handover in each case occurs with probability $1-p^{a-b}$. Once the type of handover has been determined, the simulation calculates the throughput per cell considering the residence time (i.e., time available for receiving data) and the handover delay (i.e., time unavailable for receiving data). Note we do not consider unavailability due to link layer collisions or weak channel conditions.

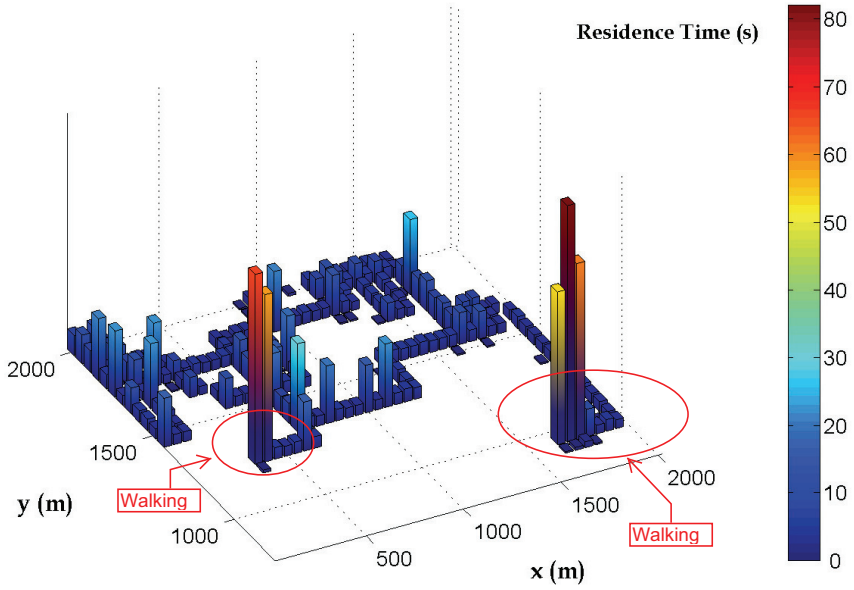
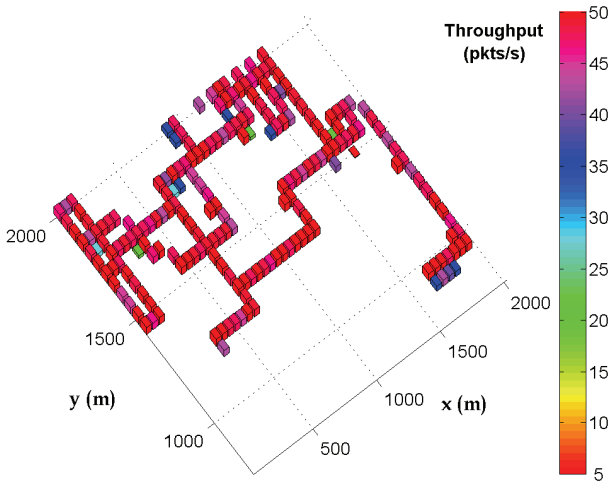


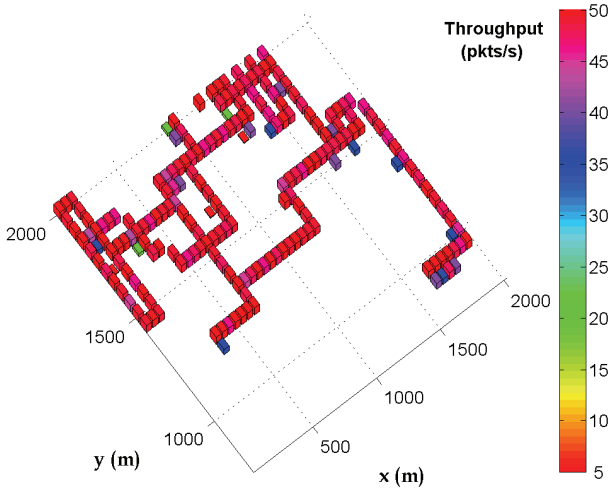
Figure 5.15: Residence times of a commuter during a journey to work

Two sets of probabilities have been used during simulations (Table 5.3). The Set #1 represents a loosely coupled architecture, where inter-domains handovers happen frequently, except for cellular-to-cellular transitions. The 90% of the time, a cellular-to-cellular transition result in intra-domain handover, because a single cellular operator typically provides a large coverage. The Set #2 represents an architecture in which more access networks belong to the same provider, resulting in intra-domain handovers happening more frequently than in Set #1. The delays caused by intra and inter-domain handovers, to WiFi and 3G technologies, have been calculated from the analysis presented in Section 5.5.2. In addition, the node is actively receiving data from the Internet during the whole duration of the journey, at a rate γ=50 packets/s. Examples of the throughput obtained for each simulated set, given Δ=0.5, are illustrated in Fig. 5.16

In order to verify the behavior of our scheme for different ratios of coverage, we have run both sets 30 times for each Δ value. The results are plotted in Fig. 5.17 with the 95% confidence interval. It is observed that Set #2 suffers from less packet losses than Set #1.



(a) Set #1



(b) Set #2

Figure 5.16: Hybrid HIP/PMIP throughput examples in loosely-coupled network architectures

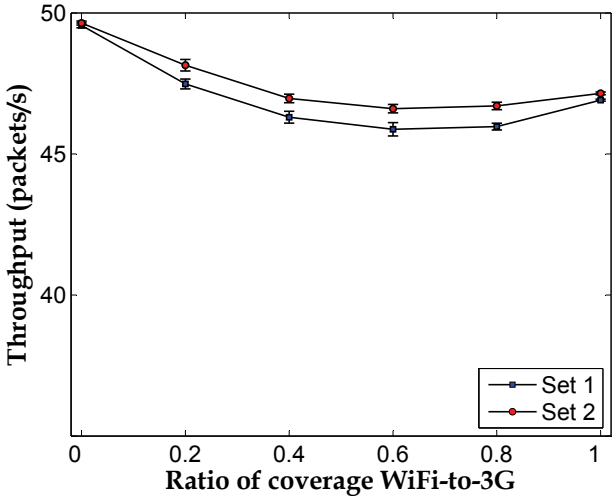


Figure 5.17: Hybrid HIP/PMIP throughput in a city scenario

This is explained because the inter-domain handover delay is higher compared with the intra-domain handover. Therefore, the more loosely-coupled the architecture is (Set #1), the more inter-domain handovers the commuter's mobile device has to experienced. Nevertheless, for both scenarios the performance of our hybrid scheme achieves throughputs ranging from 90% to 98% of the total packets sent. The results are promising considering that we have employed the highest HIP-enabled node's handover delays calculated in Section 5.5.2.

These results are a good approximation for demonstrating that our hybrid HIP/PMIP scheme allows for a seamless transferring of IP sessions, despite of the different patterns of mobility found in urban vehicular scenarios, and the heterogeneity of the supporting vehicular network infrastructure.

5.7 Summary

This chapter has presented a novel hybrid HIP/PMIP interworking scheme, which enables a global mobility management mechanism for users of vehicular networks in urban scenarios.

The scheme considers Internet access from devices installed in cars, buses or trains, as well as passenger's legacy devices with no mobility support. In addition, it considers users that commute between vehicles and terminal stations, which employ HIP to support mobility in an autonomous manner. The key advantage of this scheme is that it allows for intra and inter-domain handovers of nodes, as well as intra and inter-technology handovers, over loose coupling architectures. That means that nodes employing the hybrid HIP/PMIP scheme, can maintain seamless communications regardless of roaming agreements between network operators.

Our performance analysis has shown that the proposed scheme outperforms other protocols, such as the optimized version of MIPv6, NEMO BS, the standard HIP, and Novaczki's micro-mobility scheme for HIP. Furthermore, we have carried out simulations in a realistic urban vehicular scenario, in which pedestrian and vehicular mobility traces are combined to recreate a commuter's journey to his/her workplace. The simulations have considered different coupling levels among the network operators, as well as two different technologies (i.e., WiFi and 3G). The experimental results have confirmed that by employing HIP/PMIP scheme, a mobility-enabled mobile device is able to maintain seamless communications for Internet access in city scenarios.

Chapter 6

Conclusions and Future Work

In this chapter, we summarize the main research results and discuss further work.

6.1 Major Research Results

The objective of this research is to investigate the seamless support of IP services in multi-hop vehicular communications networks (VCN). However, the multi-hop VCN pose several research challenges including, but not limited to, a highly heterogeneous communications infrastructure with variable market penetration, a limited support of IP applications in some of the current standards for vehicular environments, a lack of robust IP mobility support mechanisms that operate over relayed communications, an asymmetric wireless network, and a lack of motivation at intermediate vehicles for relaying external IP traffic. In what follows, we provide a summary of the main research results, in which we have addressed those challenges and proposed effective solutions.

In Chapter 3, we have proposed VIP-WAVE as a novel framework for the support of IP communications in 802.11p/WAVE vehicular networks. In VIP-WAVE, we have defined mechanisms for IP addressing and IP mobility over one-hop and two-hop communications with 802.11p/WAVE technologies. Such mechanisms have been designed to provide

differentiated treatment for extended and non-extended services, and they consider the intricacies and special characteristics of 802.11p/WAVE networks. Moreover, an accurate model that involves mobility, channel conditions, MAC layer collisions, and handover delays, has been provided for evaluating the performance of our proposed framework. The simulations results have confirmed the accuracy of our model and the effectiveness of VIP-WAVE, not only for providing seamless IP communications in the multi-hop VCN, but also for improving the performance in a low presence of infrastructure.

In Chapter 4, we have studied the deployment of IP applications over a multi-hop VCN, powered by off-the-shelf WLAN technologies such as 802.11b/g/n. The main characteristic we have considered is the presence of asymmetric links due to dissimilar trasmission powers. Consequently, we have enhanced the performance of Proxy Mobile IP (PMIP) over multi-hop asymmetric VCN. Such an enhanced version integrates a predictive handover mechanism, and considers the security issues of employing I2V2V communications. MA-PMIP employs road traffic information and adapts itself to the link's directionality, in order to achieve near lossless flows of packets over multi-hop communications. Moreover, it provides an authentication mechanism to guarantee that vehicles relaying traffic during handover events, are not deceived by possible attackers. We have compared the performance of MA-PMIP with previous works, and have demonstrated it outperforms the previous schemes. Furthermore, the simulation results in realistic vehicular scenarios have demonstrated that MA-PMIP achieves seamless IP communications.

Finally, in Chapter 5, we have enabled an extended IP mobility support over the heterogeneous infrastructure that supports the vehicular communications. That means, we have considered the integration of other wireless technologies, such as 3G/4G cellular networks, besides the already addressed WLAN technologies. In order to achieve the global mobility, we have proposed a hybrid scheme that combines host-based and network-based mobility. The scheme allows for the transferring of on-going IP sessions across dissimilar access networks and administrative domains, at vehicular and pedestrians speeds. Thus, we have expanded the VCN, and in addition to consider vehicular communications, we have

also consider connections coming from commuters and pedestrians. By means of analytical evaluations and simulations of realistic urban vehicular scenarios, we have shown that our hybrid scheme can achieve seamless communications for Internet access in city scenarios.

6.2 Future work

This research has investigated the provision of seamless IP services over multi-hop vehicular communications networks. Nevertheless, there are still several research directions in which this research can be extended.

- **Adapted Distributed Mobility**

  Predictive handovers within the vehicular communications network have been studied in this thesis by considering important online measurements, such as velocity, geographic location, and road traffic conditions. A higher response from the IP mobility mechanism is achievable if more information from the applications side can be properly used. In fact, an adaptive mobility management scheme could be developed in which, by properly specifying the type of application, the mobility protocol determines if the IP addresses employed for the communications should be or not transferrable to other access networks. In this way, mobility signalling is reduced when the granularity of the IP prefix assignment allows for an also granular IP mobility provision. Consequently, as opposed to by default employing one single prefix for all communications, and generating oftentimes needless signalling for IP mobility, the mobility scheme will address only prefixes of applications for which mobility is a requisite.

- **Multipath TCP as a supporting mechanism to improve the IP mobility performance**

  We have previously considered an urban vehicular communications network where different network operators provide overlapping service areas. Accordingly, the ve-

hicle's mobile router may have several wireless interfaces to connect to those access networks. In such scenario, the vehicle is entitled to maintain simultaneous connections over dissimilar access networks. Thus, the mobile router is converted to a multi-homed mobile router. In the traditional implementation of TCP, it only permits the use of a single pair of source-destination identifiers for each established TCP session. As a consequence, most of the efforts to exploit multi-homing have been dedicated to network layer mechanisms for load balancing and traffic engineering. Moreover, additional complexity is added when the router is mobile and multi-homed simultaneously. Recently, there have been efforts to adapt the existent mobility protocols for scenarios where the node is also multi-homed [110, 114]. However, in the traditional Internet, only one path between a source-destination pair of addresses is exposed to the transport layer, and hence the use of different paths can not be exploited by the upper layers.

To address this limitation, the IETF is working on Multipath TCP [115], as an extension to the traditional TCP. This protocol aims at making TCP able to use multiple paths when they are available between the two peers. By using multiple paths for a single TCP session, metrics such as the resilience to outages, the throughput, and the efficient use of bandwidth could be improved. However, benefits in terms of improving the mobility performance remain unexplored. Therefore, we will investigate the impact of using multiple paths at the transport layer, for nodes in the highly dynamic vehicular network. The objective is to study how the use of Multipath TCP could boost the mobility performance for multi-homed in-vehicle mobile networks.

APPENDICES

Appendix A

Author's Related Publications

Journals & Magazines

J.1 **S. Céspedes**, N. Lu, X. (Sherman) Shen. "VIP-WAVE: On the Feasibility of IP Communications in 802.11p Vehicular Networks", *IEEE Transactions on Intelligent Transportation Systems*, accepted for publication, pp.1-16, 2012.

J.2 **S. Céspedes**, X. (Sherman) Shen, C. Lazo. "IP Mobility Management for Vehicular Communication Networks: Challenges and Solutions", *IEEE Communications Magazine - [Topics in automotive networking].* vol.49, no.5, pp.187-194, May 2011.

J.3 **S. Céspedes**, S. Taha, X. (Sherman) Shen. "A Multi-hop Authenticated Proxy Mobile IP Scheme for Asymmetric VANET", *in preparation for submission.*

Conferences

C.1 **S. Céspedes**, X. (Sherman) Shen. "Enabling Relay-aided IP Communications in 802.11p/WAVE Networks", *IEEE GLOBECOM'12*, accepted for publicacion, 2012.

C.2 S. Taha, **S. Céspedes**, X. (Sherman) Shen. "EM^{3}A: Efficient Mutual Multi-hop Mobile Authentication Scheme for PMIP Networks", in *Proc. IEEE ICC'12*, June 2012, pp. 1–5.

C.3 **S. Céspedes**, X. (Sherman) Shen. "A Framework for Ubiquitous IP Communications in Vehicle to Grid Networks", in *Proc. IEEE GLOBECOM Wkshps'11 - UbiCoNet*. Dec. 2011, pp. 1231–1235.

C.4 M. Asefi, **S. Céspedes**, X. (Sherman) Shen., J. Mark. "A Seamless Quality-Driven Multi-Hop Data Delivery Scheme for Video Streaming in Urban VANET Scenarios", in *Proc. IEEE ICC'11*. June 2011, pp. 1–5.

C.5 **S. Céspedes**, X. (Sherman) Shen. "An Efficient Hybrid HIP-PMIPv6 Scheme for Seamless Internet Access in Urban Vehicular Scenarios", in *Proc. IEEE GLOBECOM'10*. Dec. 2010, pp. 11–5.

References

[1] European Telecommunications Standards Institute (ETSI), "Intelligent Transport Systems (ITS); Communications Architecture," *European Standard (Telecommunications series)*, vol. 1, no. EN 302 665, pp. 1–44, 2010.

[2] S. Y. Wang, "On the Effectiveness of Distributing Information among Vehicles Using Inter-Vehicle Communication," in *Proc. IEEE Intelligent Transportation Systems Conference*, vol. 2, Oct. 2003, pp. 1521–1526.

[3] A. Abdrabou and W. Zhuang, "Probabilistic Delay Control and Road Side Unit Placement for Vehicular Ad Hoc Networks with Disrupted Connectivity," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 129–139, Jan. 2011.

[4] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment: Wireless Access in Vehicular Environments," *IEEE Unapproved Draft Std P802.11p /D11.0*, p. 45, Mar. 2010.

[5] "1609.3-2010 IEEE Standard for Wireless Access in Vehicular Environments (WAVE) - Networking Services," *IEEE Std 1609.3-2010 (Revision of IEEE Std 1609.3-2007)*, pp. 1–144, Dec. 2010.

[6] "1609.4-2010 IEEE Standard for Wireless Access in Vehicular Environments (WAVE)Multi-channel Operation," *IEEE Std 1609.4-2010 (Revision of IEEE Std 1609.4-2006)*, pp. 1–89, Feb. 2011.

[7] A. Balasubramanian, R. Mahajan, A. Venkataramani, B. N. Levine, and J. Zahorjan, "Interactive WiFi Connectivity for Moving Vehicles," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, pp. 427–438, Oct. 2008.

[8] J. Eriksson, H. Balakrishnan, and S. Madden, "Cabernet: Vehicular Content Delivery using WiFi," in *Proc. ACM MobiCom*, Sept. 2008, pp. 199–210.

[9] M. Tsukada, I. B. Jemaa, H. Menouar, W. Zhang, M. Goleva, and T. Ernst, "Experimental Evaluation for IPv6 over VANET Geographic Routing," in *Proc. International Wireless Communications and Mobile Computing Conference*, June 2010, pp. 736–741.

[10] S. Pack, X. Shen, J. W. Mark, and J. Pan, "Mobility Management in Mobile Hotspots with Heterogeneous Multihop Wireless Links," *IEEE Commun. Mag.*, vol. 45, no. 9, pp. 106–112, Sept. 2007.

[11] J.-H. Lee, T. Ernst, and N. Chilamkurti, "Performance Analysis of PMIPv6-Based NEtwork MObility for Intelligent Transportation Systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 74–85, Jan. 2012.

[12] K. Andersson, C. Ahlund, B. S. Gukhool, and S. Cherkaoui, "Mobility Management for Highly Mobile Users and Vehicular Networks in Heterogeneous Environments," in *Proc. IEEE LCN*, Oct. 2008, pp. 593–599.

[13] I. Soto, C. J. Bernardos, M. Calderon, A. Banchs, and A. Azcorra, "Nemo-enabled Localized Mobility Support for Internet Access in Automotive Scenarios," *IEEE Commun. Mag.*, vol. 47, no. 5, pp. 152–159, May 2009.

[14] K. Seada, "Insights from a Freeway Car-to-Car Real-World Experiment," in *Proc. ACM WiNTECH*, Sept. 2008, pp. 49–56.

[15] M. Fiore and J. Härri, "The Networking Shape of Vehicular Mobility," in *Proc. ACM MobiHoc*, May 2008, pp. 261–272.

[16] D. Kotz, C. Newport, R. S. Gray, J. Liu, Y. Yuan, and C. Elliott, "Experimental Evaluation of Wireless Simulation Assumptions," in *Proc. ACM MSWiM*, Oct. 2004, pp. 78–82.

[17] F. Bai and B. Krishnamachari, "Spatio-Temporal Variations of Vehicle Traffic in VANETs," in *Proc. ACM VANET*, Sept. 2009, pp. 43–52.

[18] J. Yoo, B. S. C. Choi, and M. Gerla, "An opportunistic relay protocol for vehicular road-side access with fading channels," in *Proc. IEEE ICNC*, Oct. 2010, pp. 233–242.

[19] M. E. Mahmoud and X. Shen, "PIS: A Practical Incentive System for Multihop Wireless Networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 4012–4025, Oct. 2010.

[20] N. Lu, T. H. Luan, M. Wang, X. Shen, and F. Bai, "Capacity and Delay Analysis for Social-Proximity Urban Vehicular Networks," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 1476–1484.

[21] S.-T. Cheng, G.-J. Horng, and C.-L. Chou, "Adaptive Vehicle to Vehicle Heterogeneous Transmission in Cooperative Cognitive Network VANETs," *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 2, pp. 1263–1274, Feb. 2012.

[22] J. Santa, M. Tsukada, T. Ernst, O. Mehani, and A. F. G. Skarmeta, "Assessment of VANET multi-hop routing over an experimental platform," *International Journal of Internet Protocol Technology*, vol. 4, no. 3, p. 158, Sept. 2009.

[23] T. Chen, L. Wu, F. Wu, and S. Zhong, "Stimulating Cooperation in Vehicular Ad Hoc Networks: A Coalitional Game Theoretic Approach," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 566–579, Feb. 2011.

[24] C. Perkins, D. Johnson, and J. Arkko, "Mobility Support in IPv6," *IETF Secretariat, RFC 6275*, July 2011.

[25] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol," *IETF Secretariat, RFC 3963*, Jan. 2005.

[26] H. Petander, E. Perera, K.-C. Lan, and A. Seneviratne, "Measuring and Improving the Performance of Network Mobility Management in IPv6 Networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 9, pp. 1671–1681, Sept. 2006.

[27] M. Calderon, C. J. Bernardos, M. Bagnulo, I. Soto, and A. de la Oliva, "Design and Experimental Evaluation of a Route Optimization Solution for NEMO," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 9, pp. 1702–1716, Sept. 2006.

[28] T. Manabu, M. Olivier, and E. Thierry, "Simultaneous Usage of NEMO and MANET for Vehicular Communication," in *Proc. TridentCom*, Mar. 2008, pp. 1–8.

[29] M. Watari, R. Wakikawa, T. Ernst, and J. Murai, "Optimal Path Establishment for Nested Mobile Networks," in *IEEE 62nd VTC-2005-Fall*, vol. 4, Sept. 2005, pp. 2302–2306.

[30] K. Okada, R. Wakikawa, and J. Murai, "MANET and NEMO Converged Communication," in *Technologies for Advanced Heterogeneous Networks II*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2006, vol. 4311, pp. 235–251.

[31] C. J. Bernardos, I. Soto, M. Calderón, F. Boavida, and A. Azcorra, "VARON: Vehicular Ad hoc Route Optimisation for NEMO," *Computer Communications*, vol. 30, no. 8, pp. 1765–1784, June 2007.

[32] H.-J. Lim, M. Kim, J.-H. Lee, and T. M. Chung, "Route Optimization in Nested NEMO: Classification, Evaluation, and Analysis from NEMO Fringe Stub Perspective," *IEEE Trans. Mobile Comput.*, vol. 8, no. 11, pp. 1554–1572, Nov. 2009.

[33] R. Baldessari, W. Zhang, A. Festag, and L. Le, "A MANET-centric Solution for the Application of NEMO in VANET Using Geographic Routing," in *Proc. TridentCom*, Mar. 2008, pp. 1–7.

[34] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson, "Host Identity Protocol," *IETF Secretariat, RFC 5201*, Apr. 2008.

[35] A. Gurtov, M. Komu, and R. Moskowitz, "Host Identity Protocol: Identifier/Locator Split for Host Mobility and Multihoming," *The Internet Protocol Journal, IPJ*, vol. 12, no. 1, pp. 27–32, Mar. 2009.

[36] J. Laganier, T. Koponen, and L. Eggert, "Host Identity Protocol (HIP) Registration Extension," *IETF Secretariat, RFC 5203*, Apr. 2008.

[37] P. Jokela, T. Rinta-aho, T. Jokikyyny, J. Wall, M. Kuparinen, H. Mahkonen, J. Melen, T. Kauppinen, and J. Korhonen, "Handover performance with HIP and MIPv6," in *Proc. 1st Int. Symp. on Wireless Comm. Systems*, Sept. 2004, pp. 324–328.

[38] J. Y. H. So and J. Wang, "Micro-HIP A HIP-Based Micro-Mobility Solution," in *Proc. IEEE ICC Workshops*, May 2008, pp. 430–435.

[39] S. Novaczki, L. Bokor, and S. Imre, "Micromobility Support in HIP: Survey and Extension of Host Identity Protocol," in *Proc. IEEE MELECON*, May 2006, pp. 651–654.

[40] A. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, "Proxy Mobile IPv6," *IETF Secretariat, RFC 5213*, Aug. 2008.

[41] A. Diab and A. Mitschele-Thiel, "Comparative Analysis of Proxy MIPv6 and Fast MIPv6," in *Proc. of ACM MobiWAC*, Oct. 2009, pp. 26–33.

[42] H. Yokota, K. Chowdhury, B. Patil, and F. Xia, "Fast Handovers for Proxy Mobile IPv6," *IETF Secretariat, RFC 5949*, Sept. 2010.

[43] M. Olsson, S. Sultana, S. Rommer, L. Frid, and C. Mulligan, *SAE and the Evolved Packet Core*. Great Britain: Elsevier, 2009, p. 440.

[44] S. Pack, T. Kwon, Y. Choi, and E. K. Paik, "An Adaptive Network Mobility Support Protocol in Hierarchical Mobile IPv6 Networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, p. 3627, Sept. 2009.

[45] T. H. Luan, X. Ling, and X. Shen, "MAC in Motion: Impact of Mobility on the MAC of Drive-Thru Internet," *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 305 – 319, Feb. 2012.

[46] H. Liang and W. Zhuang, "Double-loop receiver-initiated MAC for cooperative data dissemination via roadside WLANs," *IEEE Trans. Commun., to be published.*

[47] T. Zhou, H. Sharif, M. Hempel, P. Mahasukhon, W. Wang, and T. Ma, "A Novel Adaptive Distributed Cooperative Relaying MAC Protocol for Vehicular Networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 72–82, Jan. 2011.

[48] H. Shan, W. Zhuang, and Z. Wang, "Distributed Cooperative MAC for Multihop Wireless Networks," *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 126–133, Feb. 2009.

[49] M. Asefi, S. Céspedes, X. Shen, and J. W. Mark, "A Seamless Quality-Driven Multi-Hop Data Delivery Scheme for Video Streaming in Urban VANET Scenarios," *Proc. IEEE ICC*, pp. 1–5, June 2011.

[50] M. Felegyhazi, J.-P. Hubaux, and L. Buttyan, "Nash equilibria of packet forwarding strategies in wireless ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 5, no. 5, pp. 463–476, May 2006.

[51] N. B. Salem, L. Buttyán, J.-P. Hubaux, and M. Jakobsson, "A charging and rewarding scheme for packet forwarding in multi-hop cellular networks," in *Proc. ACM MobiHoc*, June 2003, p. 13.

[52] C. Wewetzer, M. Caliskan, K. Meier, and A. Luebke, "Experimental Evaluation of UMTS and Wireless LAN for Inter-Vehicle Communication," in *Proc. 7th International Conference on ITS Telecommunications*, June 2007, pp. 1–6.

[53] E. Baccelli, T. Clausen, and R. Wakikawa, "IPv6 Operation for WAVE - Wireless Access in Vehicular Environments," in *Proc. IEEE VNC*, 2010, pp. 160–165.

[54] T. Narten, E. Nordmark, W. Simpson, and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)," *IETF Secretariat, RFC 4861*, 2007.

[55] M. Grossglauser and D. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, Aug. 2002.

[56] European Telecommunications Standards Institute (ETSI), "Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 6: Internet Integration; Sub-part 1: Transmission of IPv6 Packets over GeoNetworking Protocols," *Technical Specification*, vol. 1, no. TS 102 636-6-1, pp. 1–45, Nov. 2011.

[57] M. Fazio, C. E. Palazzi, S. Das, and M. Gerla, "Automatic IP Address Configuration in VANETs," in *Proc. ACM VANET*, Sept. 2006, pp. 100–101.

[58] R. Baldessari, C. J. Bernardos, and M. Calderon, "GeoSAC - Scalable Address Autoconfiguration for VANET Using Geographic Networking Concepts," in *Proc. IEEE PIMRC*, Sept. 2008, pp. 1–7.

[59] T. Kato, K. Kadowaki, T. Koita, and K. Sato, "Routing and Address Assignment Using Lane/Position Information in a Vehicular Ad Hoc Network," pp. 1600–1605, 2008.

[60] S. Pack, Y. Kim, K. Kim, and W. Lee, "On Session Handoff Probability in NEMO-Based Vehicular Environments," in *Proc. IEEE CCNC*, Jan. 2010, pp. 1–5.

[61] A. Prakash, S. Tripathi, R. Verma, N. Tyagi, R. Tripathi, and K. Naik, "Vehicle Assisted Cross-Layer Handover Scheme in NEMO-based VANETs (VANEMO)," *Int. J. of Internet Protocol Technology*, vol. 6, no. 1/2, p. 83, June 2011.

[62] B. Azzedine, Z. Zhenxia, and X. Fei, "Reducing Handoff Latency for NEMO-Based Vehicular Ad Hoc Networks," in *Proc. IEEE Globecom*, Dec. 2011, pp. 1–5.

[63] S. Céspedes, X. Shen, and C. Lazo, "IP Mobility Management for Vehicular Communication Networks: Challenges and Solutions," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 187–194, May 2011.

[64] S.-Y. Wang, H.-L. Chao, K.-C. Liu, T.-W. He, C.-C. Lin, and C.-L. Chou, "Evaluating and Improving the TCP/UDP Performances of IEEE 802.11(p)/1609 Networks," in *Proc. IEEE ISCC*, vol. 11, July 2008, pp. 163–168.

[65] Y. Wu, Y. Zhu, and B. Li, "Trajectory Improves Data Delivery in Vehicular Networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 2183–2191.

[66] S. C. Ng, W. Zhang, Y. Zhang, and Y. Yang, "Analysis of Access and Connectivity Probabilities in Vehicular Relay Networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 140–150, Jan. 2011.

[67] C. Han, M. Dianati, R. Tafazolli, R. Kernchen, and X. Shen, "Analytical Study of the IEEE 802.11p MAC Sublayer in Vehicular Networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 873 – 886, June 2012.

[68] M. J. Khabbaz, W. F. Fawaz, and C. M. Assi, "Probabilistic Bundle Relaying Schemes in Two-Hop Vehicular Delay Tolerant Networks," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 281–283, Mar. 2011.

[69] Y. Du, L. Zhang, Y. Feng, Z. Ren, and Z. Wang, "Performance Analysis and Enhancement of IEEE 802.11 p/1609 Protocol Family in Vehicular Environments," in *IEEE ITSC*, Sept. 2010, pp. 1085–1090.

[70] S. Eichler, "Performance Evaluation of the IEEE 802.11p WAVE Communication Standard," in *Proc. IEEE VTC'07 Fall*, Sept. 2007, pp. 2199–2203.

[71] W. L. Tan, W. C. Lau, O. Yue, and T. H. Hui, "Analytical Models and Performance Evaluation of Drive-thru Internet Systems," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 207–222, Jan. 2011.

[72] O. Tonguz, W. Viriyasitavat, and F. Bai, "Modeling urban traffic: A cellular automata approach," *IEEE. Commun. Mag.*, vol. 47, no. 5, pp. 142–150, 2009.

[73] A. Varga and R. Hornig, "An Overview of the OMNeT++ Simulation Environment," in *Proc. Simutools*, Mar. 2008, pp. 60:1–60:10.

[74] A. Ariza-Quintana, E. Casilari, and A. Cabrera Triviño, "Implementation of MANET Routing Protocols on OMNeT++," in *Proc. Simutools*, Mar. 2008, pp. 80:1–80:4.

[75] H. T. Cheng, H. Shan, and W. Zhuang, "Infotainment and road safety service support in vehicular networking: From a communication perspective," *Mechanical Systems and Signal Processing*, vol. 25, no. 6, pp. 2020–2038, Aug. 2011.

[76] P. Mitra and C. Poellabauer, "Asymmetric Geographic Forwarding," *International Journal of Embedded and Real-Time Communication Systems*, vol. 2, no. 4, pp. 46–70, Jan. 2011.

[77] A. Amoroso, G. Marfia, M. Roccetti, and C. E. Palazzi, "A Simulative Evaluation of V2V Algorithms for Road Safety and In-Car Entertainment," in *Proc. ICCCN*, July 2011, pp. 1–6.

[78] Y.-J. Kim, R. Govindan, B. Karp, and S. Shenker, "Geographic routing made practical," in *Proc. ACM NSDI*, 2005, pp. 217–230.

[79] I. Ben Jemaa, M. Tsukada, H. Menouar, and T. Ernst, "Validation and Evaluation of NEMO in VANET Using Geographic Routing," in *Proc. ITST*, Nov. 2010, p. 6.

[80] S. Jeon and Y. Kim, "Cost-Efficient Network Mobility Scheme over Proxy Mobile IPv6 Network," *IET Communications*, vol. 5, no. 18, p. 2656, Dec. 2011.

[81] C. Tang and D. O. Wu, "An Efficient Mobile Authentication Scheme for Wireless Networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1408–1416, Apr. 2008.

[82] B. Xie, A. Kumar, S. Srinivasan, and D. P. Agrawal, "GMSP: A Generalized Multi-hop Security Protocol for Heterogeneous Multi-hop Wireless Network," in *Proc. IEEE WCNC*, vol. 2, Apr. 2006, pp. 634–639.

[83] A. Al Shidhani and V. C. M. Leung, "Secure and Efficient Multi-Hop Mobile IP Registration Scheme for MANET-Internet Integrated Architecture," in *Proc. IEEE WCNC*, Apr. 2010, pp. 1–6.

[84] Y. Jiang, C. Lin, X. Shen, and M. Shi, "Mutual Authentication and Key Exchange Protocols for Roaming Services in Wireless Mobile Networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2569–2577, Sept. 2006.

[85] T. Heer, S. Götz, O. G. Morchon, and K. Wehrle, "ALPHA: An Adaptive and Lightweight Protocol for Hop-by-hop Authentication," in *Proc. ACM CoNEXT*, Dec. 2008, pp. 23:1–23:12.

[86] N. Ristanovic, P. Papadimitratos, G. Theodorakopoulos, J.-P. Hubaux, and J.-Y. Le Boudec, "Adaptive Message Authentication for Multi-hop Networks," in *Proc. WONS*, Jan. 2011, pp. 96–103.

[87] J. Choi, Y. Khaled, M. Tsukada, and T. Ernst, "IPv6 Support for VANET with Geographical Routing," in *Proc. ITST*, Oct. 2008, pp. 222–227.

[88] X. Zhou, J. Korhonen, C. Williams, S. Gundavelli, and C. Bernardos, "Prefix Delegation for Proxy Mobile IPv6," *IETF Secretariat, Internet draft (work in progress)*, Mar. 2012. [Online]. Available: http://tools.ietf.org/html/draft-ietf-netext-pd-pmip-02

[89] A. Petrescu, M. Boc, and C. Janneteau, "Network Mobility with Proxy Mobile IPv6," *IETF Secretariat, Internet draft (work in progress)*, Mar. 2012. [Online]. Available: http://tools.ietf.org/id/draft-petrescu-netext-pmip-nemo-00.txt

[90] S. Taha, S. Céspedes, and X. Shen, "EMˆ3A: Efficient Mutual Multi-hop Mobile Authentication Scheme for PMIP Networks," in *Proc. IEEE ICC*, June 2012, pp. 1–5.

[91] A. Gupta, A. Mukherjee, B. Xie, and D. P. Agrawal, "Decentralized Key Generation Scheme for Cellular-based Heterogeneous Wireless Ad hoc Networks," *J. Parallel Distrib. Comput.*, vol. 67, no. 9, pp. 981–991, Sept. 2007.

[92] K. R. C. Pillai and M. P. Sebastain, "A Hierarchical and Decentralized Key Establishment Scheme for End-to-End Security in Heterogeneous Networks," in *Proc. IEEE IMSAA*, Dec. 2009, pp. 1–6.

[93] A. Köpke, M. Swigulski, K. Wessel, D. Willkomm, P. T. K. Haneveld, T. E. V. Parker, O. W. Visser, H. S. Lichte, and S. Valentin, "Simulating wireless and mobile networks in OMNeT++ the MiXiM vision," in *Proc. Simutools*, Mar. 2008, pp. 71:1–71:8.

[94] "INET Framework for OMNeT++," p. 142, 2012. [Online]. Available: http://inet.omnetpp.org/doc/inet-manual-DRAFT.pdf

[95] C. Sommer and F. Dressler, "Using the Right Two-Ray Model? A Measurement based Evaluation of PHY Models in VANETs," in *Proc. ACM MobiCom*, Sept. 2011, p. 3.

[96] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research," *Wireless Communication and Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, vol. 2, no. 5, pp. 483–502, 2002.

[97] A. D. May, *Traffic Flow Fundamentals*. Englewood Cliffs, NJ: Prentice-Hall, 1990, p. 476.

[98] I. F. Akyildiz, J. McNair, J. Ho, H. Uzunalioglu, and W. Wang, "Mobility Management in Next-Generation Wireless Systems," *Proc. IEEE*, vol. 87, no. 8, pp. 1347–1384, Aug. 1999.

[99] R. Moskowitz and P. Nikander, "Host Identity Protocol (HIP) Architecture," *IETF Secretariat, RFC 4423*, May 2006.

[100] K.-W. Lee, W.-K. Seo, Y.-Z. Cho, J.-W. Kim, J.-S. Park, and B.-S. Moon, "Inter-Domain Handover Scheme Using an Intermediate Mobile Access Gateway for Seamless Service in Vehicular Networks," *Int. J. Commun. Syst.*, vol. 23, no. 9-10, pp. 1127–1144, Sept. 2009.

[101] J.-H. Lee, H.-J. Lim, and T.-M. Chung, "A Competent Global Mobility Support Scheme in NETLMM," *AEU - International Journal of Electronics and Communications*, vol. 63, no. 11, pp. 950–967, Nov. 2009.

[102] G. Iapichino, C. Bonnet, O. del Rio Herrero, C. Baudoin, and I. Buret, "Combining Mobility and Heterogeneous Networking for Emergency Management: a PMIPv6 and HIP-based Approach," in *Proc. IWCMC*, June 2009, pp. 603–607.

[103] S. Herborn, L. Haslett, R. Boreli, and A. Seneviratne, "HarMoNy - HIP Mobile Networks," in *Proc. IEEE 63rd VTC-Spring.*, vol. 2, May 2006, pp. 871–875.

[104] W. Song, H. Jiang, W. Zhuang, and X. Shen, "Resource management for QoS support in cellular/WLAN interworking," *IEEE Network*, vol. 19, no. 5, pp. 12–18, Sept. 2005.

[105] P. Nikander and J. Laganier, "Host Identity Protocol (HIP) Domain Name System (DNS) Extension," *IETF Secreatariat, RFC 5205*, Apr. 2008.

[106] R. Hinden and Q. Haberman, "Unique local IPv6 Unicast Addresses," *IETF Secretariat, RFC 4193*, Oct. 2005.

[107] S. Gundavelli, M. Townsley, O. Troan, and W. Dec, "Address Mapping of IPv6 Multicast Packets on Ethernet," *IETF Secretariat, RFC 6085*, Jan. 2011.

[108] R. Draves, "Default Address Selection for Internet Protocol version 6 (IPv6)," *IETF Secretariat, RFC 3484*, Feb. 2003.

[109] A. Gurtov, *Host Identity Protocol (HIP): Towards the Secure Mobile Internet*. Chippenham, England: John Wiley & Sons Ltd, 2008, pp. 274–277.

[110] P. Nikander, T. Henderson, C. Vogt, and J. Arkko, "End-Host Mobility and Multihoming with the Host Identity Protocol," *IETF Secretariat, RFC 5206*, Apr. 2008.

[111] E. Natalizio, A. Molinaro, and S. Marano, "The Effect of a Realistic Urban Scenario on the Performance of Algorithms for Handover and Call Management in Hierarchical Cellular Systems," in *Proc. Intl Conf. Telecomm. (ICT)*, Aug. 2004, pp. 1143–1150.

[112] M. Turcotte, "Commuting to work: Results of the 2010 General Social Survey," *Canadian Social Trends, no. 92*, Aug. 2011.

[113] N. Aschenbruck, R. Ernst, E. Gerhards-Padilla, and M. Schwamborn, "BonnMotion: A Mobility Scenario Generation and Analysis Tool," in *Proc. International Conference on Simulation Tools and Techniques (ICST)*, Mar. 2010, p. 51.

[114] R. Wakikawa, V. Devarapalli, G. Tsirtsis, T. Ernst, and K. Nagami, "Multiple Care-of Addresses Registration," *IETF Secretariat, RFC 5648*, May 2009.

[115] A. Ford, C. Raiciu, M. Handley, S. Barre, and J. Iyengar, "Architectural Guidelines for Multipath TCP Development," *IETF Secretariat, RFC 6182*, Mar. 2011.