QUALITY OF SERVICES AND USER EXPERIENCE IN WIRELESS VEHICULAR INTERNET ACCESS

CHISHENG ZHANG

Ph.D
The Hong Kong Polytechnic University
2015

THE HONG KONG POLYTECHNIC UNIVERSITY
DEPARTMENT OF COMPUTING

QUALITY OF SERVICES AND USER EXPERIENCE IN WIRELESS VEHICULAR INTERNET ACCESS

CHISHENG ZHANG

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

July 2014

CERTIFICATE OF ORIGINALITY

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

(Signed)

Chisheng Zhang (Name of Student)

Abstract

Vehicular Ad-hoc Networks (VANET) have attracted increasing attentions from both research and industry communities. By equipping vehicles with on-board wireless communication facilities, the vehicular networking technology provides a variety of novel and exciting applications to vehicular users to enhance their road safety and travel comfort, and increase experiences of their on-road life. In the development of vehicular applications and services, we witness the trend of complementary vehicular networks and the Internet. To support high quality applications and services, well-planned network infrastructure and vehicle-oriented application should be jointly considered in designing vehicular Internet access networks.

To improve the quality of vehicular Internet access from the above two perspectives, we use Quality of Service (QoS) and Quality of user Experience (QoE) respectively as the measures. Quality of service is a performance indicator for vehicular access networks and characterizes the attributes associated with the networks and their protocols. On the other hand, quality of user experience is a performance metric defined from the end-user's point of view and represents the perceived quality of applications and services.

Although with a bright future, the way to ubiquitous and efficient vehicular Internet access still faces with numerous research and engineering challenges. Due to the wireless spectrum crisis in licensed wireless communication channels, the existing well-planned cellular networks are not adequate to satisfy the increasing data demand from vehicular applications. It is a challenging issue to efficiently exploit Wi-Fi or other wireless communication technologies to establish or supplement city-wide or nation-wide road-side vehicular Internet access networks. Secondly, to be resilient to performance degradation caused by highly dynamic topology of vehicular networks, packet routing and data dissemination protocols

should be adaptive to network partition and link disconnection. Last but not the least, quality of conventional Internet applications and services suffer greatly from the vehicle mobility and intermittent Internet access. To improve the user experience, application-level service enhancement should be considered in designing vehicular applications and services.

In this research, we systematically analyze the above problems and propose effective and efficient solutions to address the afore-mentioned challenges. Our goal is to build a framework to facilitate the multi-user Internet access in wireless vehicular networks with QoS guaranteed network infrastructure and protocols, and QoE enhanced applications and services. Our contribution are summarized as follows.

To address the issue of infrastructure planning, we identify the problem how to deploy different types of network components, such as base stations and relays, in a cost-efficient way to maximize the network performance. First, we propose a detailed analytical model for infrastructure-based vehicular Internet access in the presence of different network components. Then, through extensive simulations based on the proposed model, we obtain some guidelines on choosing and deploying different network components in vehicular Internet access networks. Finally, as a special case study of infrastructure planning, we investigate the incremental and cost-efficient network planning (ICNP) problem in vehicular Internet access networks with the objective of improving network capacity for an existing network under limited deployment cost. After defining a *cost efficiency* metric, namely incremental upgrade utility (IUU), we develop a heuristic algorithm, called Greedy Link Deployment algorithm (GLiD) with approximation ratio $(1 - 1/\sqrt{e})$ for deploying additional wireless links to the existing networks to achieve the maximum performance improvement.

To address the issue of protocol adaption, we investigate the maximum network connectivity problem in vehicular Internet access networks, in which vehicles adopt both vehicle-to-infrastructure and vehicle-to-vehicle communications. We first analyze the characteristics of network connectivity in vehicular Internet access networks and derive a closed form to represent the connectivity probability under specific road-side units (RSU) deployment model and vehicle arrival rate. Then, we develop a connectivity-oriented data dissemination (CoDA) protocol to maximize the minimum network connectivity in vehicular Internet access systems based on the probabilistic analysis. Finally, we validate the CoDA protocol through extensive simulations and the results show that the CoDA protocol can achieve the maximum connectivity probability in vehicular communications in different network scenarios.

To address the issue of service enhancement, we take localization service, a fundamental service in vehicular applications, as an example. We study the vehicular localization problem in urban environments. It is a challenging task to achieve high localization accuracy due to the severe multi-path effects caused by buildings and interfering sources in urban area. First, we perform a series of experiments which shows that the vehicular mobility and multi-path effect impose severe degradation in localization performance of GPS receivers. We then identify the correlation between the similarity of signal-noise ratio (SNR) values of GPS signals from different GPS satellites and relative distance among those GPS receivers. Based on the observation, we design the Networking-GPS algorithm, which constructs atomic redundantly rigid graphs according to the signal similarity and expands atomic triangles to reach anchor points along the road side. Therefore, the local accuracy can be transited to the global accuracy through the redundantly rigid transformation. Our evaluation based on real GPS traces shows that Networking-GPS can achieve higher accuracy against multi-path effect in urban area.

Publications

Book

1. Jiannong Cao, **Chisheng Zhang**, "*Seamless and Secure Communications over Heterogeneous Wireless Networks*". Springer Briefs in Computer Science, Springer, Feb. 2014. ISBN: 978-1493904150.

Journal Papers

1. **Chisheng Zhang**, Jiannong Cao, Gang Yao: "*Networking-GPS: Cooperative Vehicle Localization Using Commodity GPS in Urban Area*". ZTE Communication. Vol.12, No.1, pp.24-33, Mar., 2014.
2. Wei Feng, Jiannong Cao, **Chisheng Zhang**, Jun Zhang, and Qin Xin. *Coordination of multi-link spectrum handoff in multi-radio multi-hop cognitive networks.* Journal of Parallel and Distributed Computing, 72(4):613-625, 2012.

Conference Papers

1. **Chisheng Zhang**, Jiannong Cao, Gang Yao: "*CoDA: Connectivity-Oriented Data Dissemination Algorithm for Vehicular Internet Access Networks*". In Mobile Ad-hoc and Sensor Networks (MSN), 2015 11th International Conference on, December 2015.
2. Jiannong Cao, **Chisheng Zhang**, et al. *SHAWK: Platform for secure integration of heterogeneous advanced wireless networks.* In Advanced Information Networking and Applications Workshops (WAINA), 2012 26th International Conference on, pages 13-18, March 2012.
3. Jie Zhou, Jiannong Cao, Jun Zhang, **Chisheng Zhang**, and Yao Yu. *Analysis and countermeasure for wormhole attacks in wireless mesh networks on a real testbed.* In

Advanced Information Networking and Applications (AINA), 2012 IEEE 26th International Conference on, pages 59-66, March 2012.

4. **Chisheng Zhang**, Jiannong Cao, Jun Zhang, and Jie Zhou. *Cost-constrained incremental network planning in multi-hop wireless networks.* In Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE, pages 1-5, Dec 2011.

5. Wei Feng, Jiannong Cao, **Chisheng Zhang**, and C. Liu. *Joint optimization of spectrum handoff scheduling and routing in multi-hop multi-radio cognitive networks.* In Distributed Computing Systems, 2009. ICDCS 09. 29th IEEE International Conference on, pages 85-92, June 2009.

6. Jiannong Cao, Kun Xie, Weigang Wu, C. Liu, Gang Yao, Wei Feng, Yang Zou, Jigang Wen, **Chisheng Zhang**, et al. *HAWK: Real-world implementation of high-performance heterogeneous wireless network for internet access.* In Distributed Computing Systems Workshops, 2009. ICDCS Workshops 09. 29th IEEE International Conference on, pages 214-220, June 2009.

Demo / Video

1. Jiannong Cao, Xuefeng Liu, Steven Lai, Yang Zou, Jun Zhang, Yang Liu, and **Chisheng Zhang**. *A ubiquitous wireless video surveillance system based on pub/sub.* In Proceedings of the 13th International Conference on Ubiquitous Computing, UbiComp'11, pages 611-612, New York, NY, USA, 2011. ACM.

Patents

1. Jun Zhang, Yang Zou, Jiannong Cao, **Chisheng Zhang**, "*A packet scheduling method and system*", C.N. Patent. CN102739500A, Oct, 17, 2013

2. Gang Yao, Jiannong Cao, **Chisheng Zhang** Chuda Liu: "*A scheduling method for fast handoff using dual radio interfaces*", C.N. Patent. CN101873673B, Aug, 15, 2012

3. Gang Yao, Jiannong Cao, **Chisheng Zhang**: "*A smart handoff decision method and system using fingerprint based on fuzzy logic*", C.N. Patent. CN101998381B, May, 15, 2013

Acknowledgements

Foremost, I would like to express my deepest and sincere gratitude to my supervisor, Prof. Jiannong Cao, for his devoted guidance and systematical training of my Ph.D study. His board knowledge in many research areas and great enthusiasm to pursue impacts in research really impressed me. He unremittingly train me to be a good researcher, including finding research topics, solving research problems and presenting research works, etc. He is not only an outstanding researcher, but also an excellent leader and a good person. It is a precious and beneficial experience working with him during the period of my Ph.D study in my life.

I would like to thank Dr. Bin Xiao, serving as the BoE Chair, for his kind help and good ordination in my research degree examination. I would also like thank Prof. Xiaohua Jia and Dr. Xue Liu, who serve as members of the BoE, for their valuable comments to my this thesis.

I want to thank my colleagues, including Gang Yao, Wei Feng, Jie Zhou, Miao Xiong, Xin Xiao, Xuefeng Liu, Joanna Siebert, Weiping Zhu, Tao Li, Lei Yang, Yang Liu, Jingjing Li, Chao Ma, Jun Ma, Yin Yuan, Zongjian He, Guanqing Liang, Peng Guo and all other members of our research group at Hong Kong PolyU that I could not enumerate here. Thanks for their kind help in my research and project works.

Last but not least, a special gratitude and love goes to my family for their continuous love, support, trust, and encouragement. I thank my lovely little son, because his his sweet smile always give me great encouragement and stimulate me to further efforts. Finally, I would like to express my deepest love and thanks to my wife, Tao Tao for her support. During the most difficult time of my Ph.D study, she always stay with me and encourage me to pursue my dream. Without them, this thesis would not have been possible.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACK: Acknowledgement

BSS: Basic Service Set

C2C-CC: Car-to-Car Communication Consortium

CSMA/CA: Carrier Sense Multiple Access with Collision Avoidance

DCF: Distributed Coordination Function

DIFS: Distributed IFS

DSRC: Dedicated Short Range Communication

EDCA: Enhanced Distributed Channel Access

E-DCH: Enhanced Dedicated Channel

eNodeB: LTE Base Station

FCC: Federal Communications Commission

FDD: Frequency Division Duplex

GPRS: General Packet Radio Service

GPS: Global Positioning System

GSM: Global System for Mobile Communications

HSDPA: High Speed Downlink Packet Access

HSPA: High Speed Packet Access

HS-PDSCH: High Speed Physical Downlink Shared Channel

IBSS: Independent Basic Service Set

IEEE: Institute of Electrical and Electronics Engineers

IP: Internet Protocol

ITS: Intelligent Transportation System

LOS: Line-Of-Sight

LTE: Long Term Evolution

MAC: Medium Access Control

MBMS: Multimedia Broadcast Multicast Service
MIMO: Multiple Input Multiple Output
NLOS: Non-Line-Of-Sight
NodeB: UMTS Base Station
OFDM: Orthogonal Frequency Division Multiplexing
PCF: Point Coordination Function
PDF: Probability Density Function
PHY: Physical Layer
QoE: Quality of Experience
QoS: Quality of Service
QPSK: Quadrature Phase-Shift Keying
RAN: Radio Access Network
RNC: Radio Network Controller
RSU: Road Side Unit
SIM: Subscriber Identity Module
SINR: Signal to Interference-plus-Noise Ratio
SMS: Short Message Service
UE: User Equipment
UMTS: Universal Mobile Communication System
V2V: Vehicle-to-Vehicle
VIA: Vehicular Internet Access
WCDMA: Wideband Code Division Multiple Access
WLAN: Wireless Local Area Network

Chapter 1

Introduction

This research investigates the requirements and presents the modeling, optimization and evaluation of mobile Internet access in vehicular communication networks. In this chapter, we first describe the background of vehicular Internet access network in Section 1.1. After that, we explain the motivation of our work in Section 1.2. In Section 1.3, we summarize the main contributions of this thesis. Finally, we outline the organization of this thesis in Section 1.4.

1.1 Motivation and Background

Vehicular networking is regarded as one of the most important fundamental techniques to enable smart transportation system. For example, in 1980s Japan and North America made plans and attempts to adopt mobile communication techniques, such as microwave, mobile data terminal, and cellular radio, in vehicles to facilitate driver information systems [Kaw90] [Che90]. In recent decade, we have witnessed a great development in research and engineering in this area. There are several factors leading to this progress, which we summarize into two main categories as follows.

The first one is the increasingly advance in wireless communication technologies. In last 20 years, cellular communications have evolved from 2G (GSM/CDMA) to 4G (LTE) networks, in which network bandwidth is largely increased and mobility support have gradually met the requirements for vehicular environments. Especially, the penetration of IEEE 802.11 (Wi-Fi) technologies in home and office around the world is a huge driving force for ubiquitous Internet access. Based on IEEE 802.11 standards, the protocol suits for vehicular networking have been standardized by different organizations and nations. For example, the IEEE 802.11p standard specifies the MAC and physical layers and the IEEE 1609 protocol suite specifies the upper-layer operations. In addition, more and more vehicle manufactures tend to adopt networking technologies to ensure the safety, environmental, and comfort issues of their vehicles.

The other main factor is the ever-changing life style. With the explosive popularity of mobile devices, such as the *iPhone*®[1], *iPad*®[2], and *Android*™[3] phones, a great increase in demand for mobile data services is being experienced. Research reports released by *Ericsson*[4] in 2011 have found the demand for data service in mobile Internet has actually doubled in the past 12 months [Mal11] and predicted that mobile data traffic will multiply over 10 times by the year 2016 [WIL11], as shown in Figure 1.1 (reprinted from [Eri11]). The amount of data a consumer uses or requires per day is highly dependent on the user's device.

Although, much more attention has been paid to vehicular networking from research to engineering fields, ubiquitous and efficient information exchange between vehicles and the Internet are still not available or very expensive in most regions of the world. As reported in [PBN03], Americans spend 15 hours on average in a car each week. This is to say, in around 8.9% of a day, people would have very intermittent or no Internet access at all. The surge demand of ubiquitous access thus drives the vehicular communication ever more

---

[1] *iPhone* is a trademark of Apple Inc.

[2] *iPad* is a trademark of Apple Inc.

[3] *Android* is a trademark of Google Inc.

[4] *Ericsson* is the trademark or registered trademark of Telefonaktiebolaget LM Ericsson.

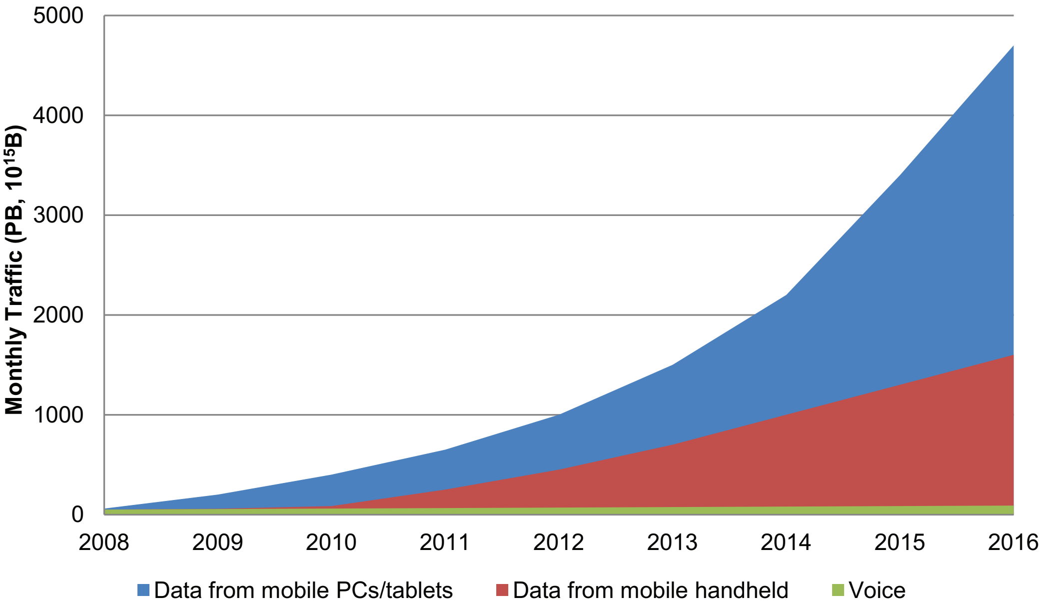


Figure 1.1: Status and prediction of mobile data demand by Ericsson

important.

To satisfy the ever-growing communication requirements from vehicular environments, the vehicular ad hoc networks (VANET) have been proposed. As illustrated in Figure 1.2, there are two networking entities and three communication paradigms. Vehicles have the communication capability to others via on-board modules, namely on-board units (OBU), and base stations (BS) or access points (AP) are deployed along the road side, namely roadside units (RSU). It is worth noting that some of the RSUs are directly connected to the Internet, but others may only serve as dedicated forwarders. Therefore, three basic communication modes exist in such a vehicular network: the vehicle-to-vehicle (V2V) communication for data message exchange among peer vehicles, the vehicle-to-RSU (V2R) communication for data forwarding between vehicular and RSU, and the vehicle-to-Internet (V2I) communication for data access between vehicular networks and Internet. The three modes can be combined together to form transmission paths among the vehicular communication networks.

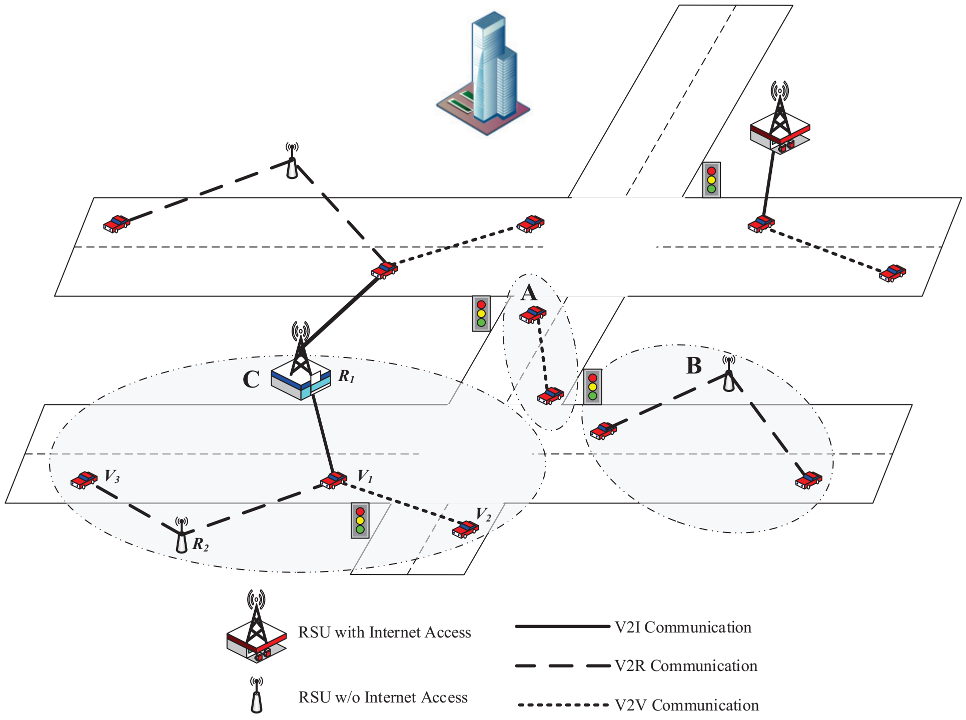


Figure 1.2: Communication paradigms in vehicular networking

In such communication network, vehicular applications can be classified in three main categories according to different objectives.

The first one is road safety application. There are a significant number of vehicle accidents in the world and most of the them are attributed to the collisions between vehicles. To predict and avoid collisions, the road safety application shares transportation information between vehicles and infrastructure entities. It is responsible for providing drivers preliminary assistance to detect and avoid collisions with other vehicles. The information includes geographic positions of vehicles and intersections, the relative distance between vehicles and intersections, the velocity and directions of vehicles, etc.

The second one is related to transportation efficiency and management. This kind of application focuses on improving the efficiency of the transportation system and providing up-to-date traffic information and navigation guidance. After gathering and processing raw traffic information from road segments, transportation efficiency and management applications dispatch the refined information to the vehicles and road-side units.

The third category is information and entertainment applications. More and more conventional Internet applications are brought in vehicular environments to enrich the journey. For example, information or entertainment services can be obtained from local wireless services running on mobile vehicles and stationary road-side units, or directly retrieved from the Internet. In this thesis, we mainly focus on the third category of VANET applications, especially on vehicular Internet access. In particular, we investigate the practical infrastructure planning, efficient network protocols and enhanced vehicular services, and develop such a vehicular Internet access system to provision the guaranteed quality-of-service (QoS) and the preserved quality-of-user-experience (QoE).

1.2 Research Challenges

Although, with the significant commercial and social benefits, the status of vehicular Internet access system still stays in the initial stage. This mainly attributes to the following reasons of current vehicular networks.

1.2.1 Lack of Infrastructure Support

Most of conventional vehicles are only equipped with limited wireless networking interfaces to fulfill the in-vehicle communications, such as infrared interfaces for remote control and Bluetooth interfaces for interaction with in-car multimedia system. Vehicles can be connected to the Internet via the cellular communication networks. However, this kind of large cell communication technologies are inefficient and low-bandwidth for inter-vehicle networking. In addition, due to the wireless spectrum crisis in licensed wireless communication channels, the existing well-planned cellular networks are not adequate to satisfy the increasing data demand from vehicular applications. Therefore, IEEE 802.11 technologies were introduced into vehicular communication networks. IEEE 802.11-based WLAN which has get a high get market penetration supports high throughput data transmission in short communication range. For example, the expected single-link throughput in its latest version 802.11ac is at least 500 Mbps.

Derived from IEEE 802.11a with mobility support, IEEE 802.11p, also referred to the dedicated short range communication (DSRC) standard, is a new communication standard which is proposed for wireless access in vehicular environments (WAVE) to support intelligent transportation systems. Drafts were developed from 2005 through 2009 and the approved amendment was published July 15, 2010. However, the infrastructure planning for vehicular networks has severely fallen behind the development of the enabling technologies. The on-board communication modules are always not listed in the compulsory components of vehicles and the establishment of road side units must introduce huge spending on infrastructure planning. *Therefore, a practical and cost-efficient solution to infrastructure planning for vehicular Internet access systems is in high demand.*

1.2.2 Highly Dynamic Network Topology

In VANETs, each communication node is a fast-moving vehicle. As a result, the V2V and V2R communications are highly violative and susceptible to frequent interruptions with the transient contact durations among nodes. For example, [BHM$^+$06] investigates on the download bandwidth of vehicles from the unplanned open residential Wi-Fi access points in Boston. It is shown that mobile vehicles can establish intermittent communications with road-side access points when the Wi-Fi coverage is available based on the the dense deployment of access points in cities. The transient and intermittent download connectivity of vehicles inevitably renders significant impairments to the media access protocols and data disseminations algorithms.

Moreover, the dramatic changing connectivity and locations of vehicles lead to the dynamic network topology. This dictates any content distribution protocols to be resilient to the topology change and can self-heal quickly. Note that the mobility of vehicles are pertained to the road layout and presents specific patterns in different road environments. For example, the trajectory of vehicles on the highway is one-dimensional, and it is two-dimensional in the urban areas with dense street intersections. The heterogeneity and diversity of the vehicle mobility and topology patterns in VANETs thus indicate that there is no one-for-all solution to form the Internet access networks. Instead, to address the network dynamics in vehicular Internet access networks, the key is to exploit the specific mobility patterns of vehicles in different deploying environments, and accordingly determine the best forwarding peers and strategies to the particular topology patterns. *Therefore, vehicular algorithms and protocols should be adaptive to network partition and link disconnection caused by topology changes.*

1.2.3 Performance Degradation of Application and Services

An increasing number of applications and services are emerging in vehicular environments. Some of them are conventional Internet applications and services, which are originally designed in stationary or low speed scenarios. The performance of those applications and services may suffer a lot when network conditions change over the time.

With the capability of communication with neighbors, vehicles can share information around the network, which is a big difference in standlone cases. From the other perspective, vehicular networking may provide a potential opportunity to improve the performance of application and services, since the correlations between neighboring vehicles can be utilized to suppress the degradation. *Therefore, the vehicular applications and services should be redesigned by leveraging the characteristics of vehicular networking to improve the performance.*

1.3 Contributions of the Thesis

To measure the quality improvement of vehicular Internet access from the two perspectives, we use quality of service (QoS) and quality of user experience (QoE). Quality of service is an general performance indicator for vehicular access networks and characterizes the attributes associated with the networks and protocols objectively. Quality of user experience is a performance metric from the end-user side and represents the perceived quality of applications and services objectively or subjectively. In this thesis, we report our solutions to provision QoS guaranteed and QoE preserved vehicular Internet access system by addressing the aforementioned challenges. As shown in zone C of Figure 1.2, in a typical vehicular Internet access network, vehicles get access to the Internet in different manners. V_1 is covered by the nearby RSU with Internet access (R_1), and it can access the Internet via the V2I communication link. V_2 is only covered by V_1. So, V_1 has to first send the Internet request to V_1 via V2V communication link, then V_1 forwards the request to R_1. The leftmost vehicle V_3 cannot communicate with any other vehicles and RSUs with Internet access. V_3 has to follow a 3-hop path from R_2, a RSU without Internet access, to V_1, finally to R_1, which involves V2R communication to forward Internet traffic.

The contribution of the thesis lies in infrastructure planning, protocol adaption and service enhancement. As shown in Figure 1.3, our contributions includes three parts:

First, with respect to infrastructure planning, we design a practical and cost-efficient incremental network planning solution to achieve the better coverage and network performance. Second, with respect to protocol adaption, we develop a connectivity-oriented data dissemination protocol to maximize the minimum network connectivity in vehicular Internet access systems based on the probabilistic analysis. Third, with respect to service enhancement, we take the vehicular localization service as an example and develop a cooperative vehicle localization algorithm, named Networking-GPS, using commodity GPS in

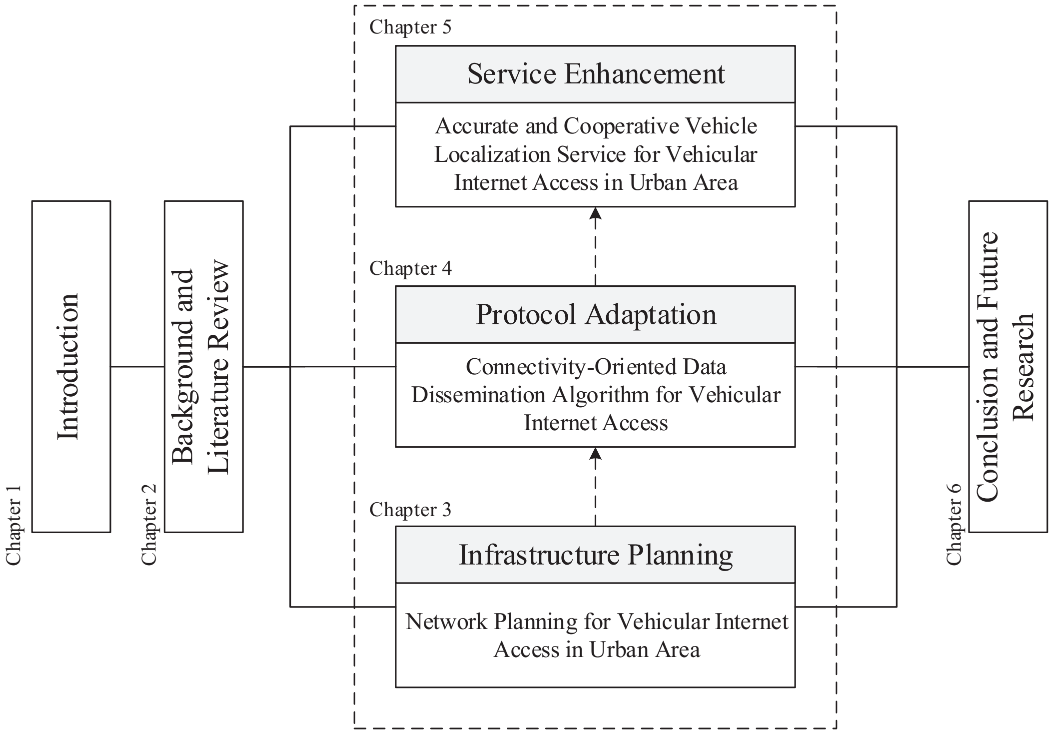


Figure 1.3: An outline of the contributions of this thesis

urban area to achieve better localization accuracy. We will illustrate these contributions in detail one by one:

1.3.1 Contribution to Infrastructure Planning

To address the issue of infrastructure planning, we identify the problem how to deploy different types of network components, such as base stations and relays, in a cost-efficient way to maximize the network performance. First, we propose a detailed analytical model for infrastructure-based vehicular Internet access in the presence of different network components. Then, through extensive simulations based on the proposed model, we obtain some guidelines on choosing and deploying different network components in vehicular Internet access networks. Finally, as a special case study of infrastructure planning, we investigate the incremental and cost-efficient network planning (ICNP) problem in vehicular Internet access networks with the objective of improving network capacity for an existing network under limited deployment cost. After defining a *cost efficiency* metric, namely incremental upgrade utility (IUU), we develop a heuristic algorithm, called Greedy Link Deployment algorithm (GLiD) with approximation ratio $(1 - 1/\sqrt{e})$ for deploying additional wireless links to the existing networks to achieve the maximum performance improvement.

1.3.2 Contribution to Protocol Adaption

For the protocol adaption issue, we focus on the data dissemination algorithms in the scenario of vehicular Internet access.

The flexibility of communications between vehicles and infrastructure networks are facilitated in the presence of vehicular communication networks. On the top of communication networks, diverse vehicular applications impose different requirements based on their unique characteristic. The capability of multi-hop data dissemination is one of the major advantages of vehicular Internet access systems. Furthermore, a various of vehicular applications emerge by leveraging multi-hop communications and also bring new reflection to the original design of data dissemination algorithms in vehicular environments. Two general requirements are needed in most of vehicular data dissemination algorithms, including location information of mobile vehicles and a method of selecting candidate node to forward packets. Since the vehicular network is always dynamic due to the movement of vehicles, the requirements are also changing with the variance of vehicular environments. Therefore, the design of data dissemination algorithm should be adaptive to the environment.

To satisfy the requirements aforementioned, we investigate the maximum network connectivity problem in vehicular Internet access networks, in which vehicles adopt both vehicle-to-infrastructure and vehicle-to-vehicle communications. We first analyze the characteristics of network connectivity in vehicular Internet access networks and derive a closed form to represent the connectivity probability under specific road-side units (RSU) deployment model and vehicle arrival rate. Then, we develop a connectivity-oriented data dissemination (CoDA) protocol to maximize the minimum network connectivity in vehicular Internet access systems based on the probabilistic analysis. Finally, we validate the CoDA protocol through extensive simulations and the results show that the CoDA protocol can achieve the maximum connectivity probability in vehicular communications in different network scenarios.

1.3.3 Contribution to Service Enhancement

For the service enhancement issue, we believe that online real-time vehicle localization is a fundamental service for vehicular Internet applications. In recent years, satellite-based positioning systems, such as the Global Positioning System (GPS), have been adopted in most of vehicular environments. People can use the localization services via commodity GPS modules, which are pre-deployed in vehicles or GPS software in their mobile phones. Although it is easy-using and cost less, the accuracy of localization using GPS sometimes degrades in urban areas due to obstacles and radio interference. Therefore, the accurate localization in vehicular Internet access systems remains an open issue.

To achieve an accurate localization solution, we develop a cooperative vehicular localization algorithm, named Networking-GPS, using only commodity GPS receivers in urban area. The Networking-GPS algorithm is developed based on some key observations, revealed by our trace data via data processing from our experiments. First, we show that multi-path effect imposes severe degradation in localization performance of GPS receivers. Then, we identified the correlation between the similarity of SNR values from different GPS satellites and relative distance among different GPS receivers. Based on the observation, we propose the Networking-GPS algorithm, which constructs atomic redundantly rigid graphs according to the signal similarity and expands atomic triangles to reach anchor points along the road side. Therefore, the local accuracy can be transited to the global accuracy through the redundantly rigid transformation. Our evaluation based on real GPS traces show that

Networking-GPS can achieve higher accuracy against multi-path effect.

1.4 Organization of the Thesis

The structure of this thesis is shown as follows. Chapter 1 is the introduction to this thesis. Chapter 2 introduces the background knowledge of vehicular Internet access networks and reviews related works in the literature.

The main body of this thesis is divided into three parts from Chapter 3 to Chapter 5. Chapter 3 presents a practical and cost-efficient incremental network planning solution to achieve the better coverage and network performance. Chapter 4 proposes a connectivity-oriented data dissemination protocol to maximize the minimum network connectivity in vehicular Internet access systems based on the probabilistic analysis. Chapter 5 develops a cooperative vehicle localization algorithm, named Networking-GPS, using commodity GPS in urban area to improvement localization accuracy in urban area.

Finally, Chapter 6 concludes the thesis, and points out our future research directions.

Chapter 2

Background and Literature Review

This chapter first introduces the basic concepts and background knowledge of the vehicular Internet access networks. After that, an overview and comparison of the existing works on three aspects, namely infrastructure planning, protocol adaption and service enhancement, are provided.

2.1 Vehicular Internet Access

Vehicular networking plays one of the most promising enabling technologies to promote various of vehicular applications, ranging from vehicle-based, driver-based, to passengers-based and pedestrians-based applications. In this section, we focus on the enabling wireless communication technologies and real deployment of vehicular applications based on the technologies.

2.1.1 Enabling Communication Technologies

To achieve networking in vehicles, wireless communication must be adopted to connect mobile entities and the infrastructure. In this section, we present a brief introduction to various wireless networking technologies available or dedicated to vehicles and vehicular networks, including DSRC/WAVE, Wi-Fi, cellular networks, etc.

DSRC/WAVE: DSRC is short for Dedicated Short-Range Communication and WAVE is for Wireless Access in Vehicular Environments. DSRC is a wireless technology for short range communications working in a dedicated spectrum of 5.9 GHz. DSRC supports vehicle speeds up to 120mph, media communication range of 300m, and data rate from 3Mbps to 27Mbps. DSRC has two operation modes: (1) ad hoc mode, which is characterized by wireless multi-hop networking via vehicle-to-vehicle communications in a distributed way, (2) infrastructure mode, which is characterized by a one-hop wireless network via vehicle-to-infrastructure communications in a centralized manner.[LCG09]

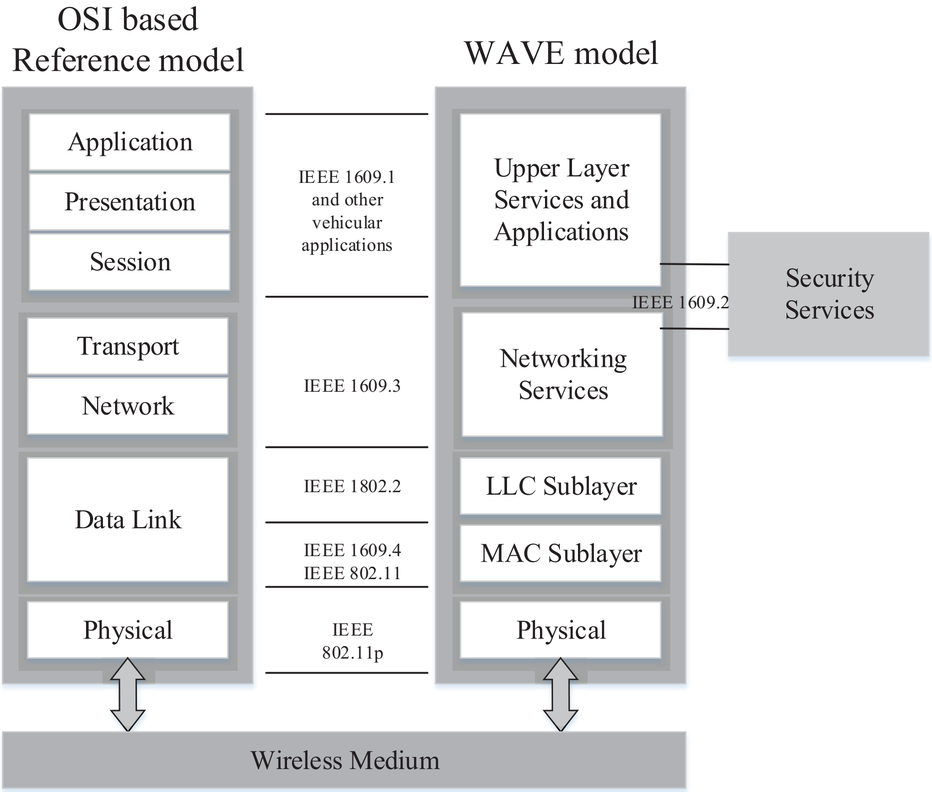


Figure 2.1: WAVE protocol suits

WLAN

IEEE 802.11 based WLAN is the most popular micro-cellular wireless communication system in the world. However, WLAN technology was not initially designed for short range communication. When the proprietary Direct Sequence Spread Spectrum (DSSS) technology over 900MHz was adopted, it could set up quite a long distant link providing data throughput of 860Kbps. In 1990s, the working frequency of WLAN was set to 2.4GHz and it could support a higher data throughput of 1Mbps or 2Mbps in a shorter distance. The year 1992 witnessed the start of drafting of the IEEE 802.11 standard for wireless LAN technologies, which had proposed 1Mbps as a standard data rate and 2Mbps as a Turbo mode. Since then, a series of IEEE 802.11 standards have been proposed to improve the performance of WLAN, some of which are listed as follows.

- IEEE 802.11a: IEEE 802.11a-1999 High-speed Physical Layer in the 5 GHz band
- IEEE 802.11b: IEEE 802.11b-1999 Higher Speed Physical Layer Extension in the 2.4 GHz band
- IEEE 802.11d: an approved amendment to the IEEE 802.11 specification that adds support for additional regulatory domains
- IEEE 802.11e: an approved amendment to the IEEE 802.11 standard that defines QoS Enhancements, including packet bursting
- IEEE 802.11g: IEEE 802.11g-2003: Further Higher Data Rate Extension in the 2.4 GHz Band
- IEEE 802.11h: an approved amendment to IEEE 802.11 standard that adds Spectrum and Transmit Power Management Extensions
- IEEE 802.11i: an approved amendment to the original IEEE 802.11 that specifies security mechanisms for wireless networks and implements WPA2
- IEEE 802.11n: an approved amendment to the IEEE 802.11 standard to improve network throughput using multiple input multiple output (MIMO) technology
- IEEE 802.11p: an approved amendment to the IEEE 802.11 standard to support Wireless Access for the Vehicular Environment
- IEEE 802.11r: an approved amendment to the IEEE 802.11 standard to permit continuous connectivity aboard wireless devices in motion, with fast and secure handovers from one access point to another, managed in a seamless manner
- IEEE 802.11s: an approved amendment to the IEEE 802.11 standard to support mesh networking

- IEEE 802.11u: an approved amendment to the IEEE 802.11 standard to add features that improve interworking with external networks.

The 802.11 standards define two types of communication modes: infrastructure mode and ad hoc mode. In infrastructure mode, the wireless network includes a wireless access point (AP) and multiple wireless clients. The AP acts as a base station and is also responsible for security management. All wireless clients communicate with external networks through the access point, which provides the connection from the wireless communication media to the hard-wired communication media. Since the wireless communication media is open, all wireless clients within the communication range can receive the packets. But in the default protocol operation, mobile clients with corresponding source and destination addresses will handle the packets.

In the ad hoc mode, the wireless network is only comprised of IEEE 802.11 wireless clients, which communicate directly with each other without forwarding packets to a access point. This communication mode is helpful to deploy a wireless network for an emergent purpose in some locations where wired infrastructure is not adequate to support all wireless clients.

Wireless Mesh Networks

A wireless mesh network (WMN) is a new wireless communications network consisting of mesh routers, mesh clients and gateways, which are self-organized in mesh topology. Wireless mesh networks can be implemented with various wireless technologies including IEEE 802.11, IEEE 802.15, IEEE 802.16, cellular technologies, or combinations of more than one type. Wireless mesh networking is a promising technology for next-generation wireless communication systems, with the capability of rapid deployment and flexible reconfigurability for disaster recovery, convention centers and hard-to-wire areas. WMN is also an alternative for long-term network infrastructure to extend wireless broadband access in dense urban areas and low-cost backbone networking to Internet gateway in remote rural areas.

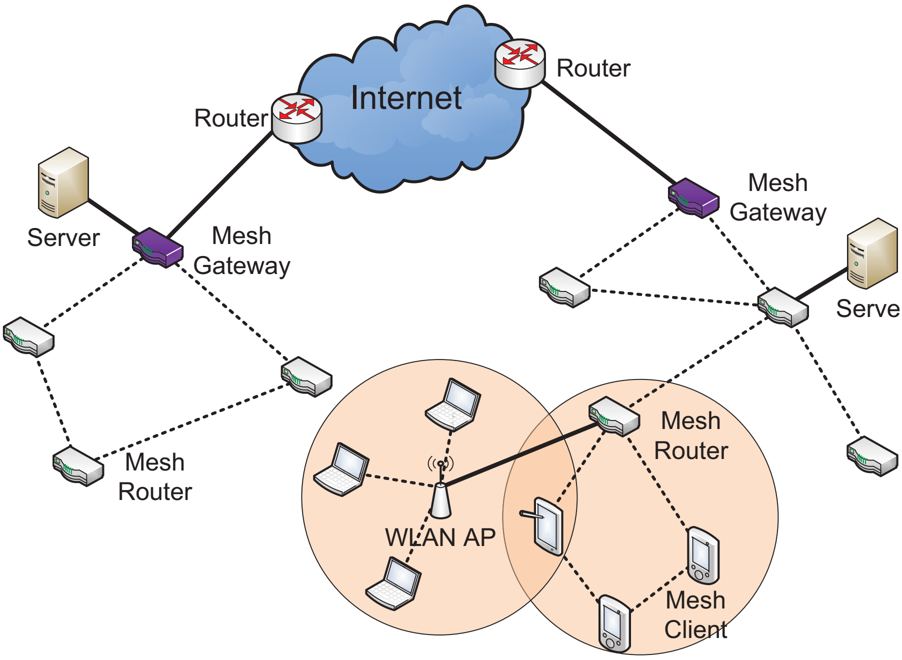


Figure 2.2: Wireless mesh networks

There are three kinds of network structures in wireless mesh networks, including infrastructure-based WMN, client-based WMN and hybrid WMN. In the first architecture, mesh routers establish a backbone network with the wireless technologies, and mesh clients obtain the network access through the core network. If some of the mesh routers have a connection to the Internet via wireless or wired links, they can be set to be the gateways, which provide the entire network Internet access service. In addition, mesh routers are usually stationary in the whole network. In client-based WMNs, it allows peer-to-peer networking between mesh clients, with no dedicated mesh routers. To a certain extent from network topology aspect, a client-based WMN is the same as a traditional ad hoc network. To meet diverse requirements of different application scenarios, a hybrid WMN is proposed to combine the above two networks.

The characteristics of WMNs [AWW05] are listed as follows:

- **Multi-hop and ad hoc wireless networking**.

  One of the most important features in wireless mesh networks is its multi-hop and ad hoc wireless networking technology to forward data packets, which can effectively extend the network coverage constrained by a single base station in WLAN, and achieve non-line-of-sight propagation, which is impossible for high radio frequency.

- **Capability of self-organization and self-healing**.

  Due to the distributed nature of multi-hop and ad hoc networking, WMN can simplify network design and deployment, thus reduce network deployment and maintenance costs. WMNs can also upgrade network performance in terms of load balance and fault tolerance. Because of these features, WMNs have low initial deployment costs, and can be deployed in an incremental fashion.

- **Various mobility and access patterns of mesh nodes**.

  Different mobility patterns exist in WMNs, including the stationary mesh routers and moving mesh clients. In addition, there are also different network access patterns, including only peer-to-peer communication between mesh routers, and hybrid communications among mesh clients. All these differences will require different protocol design and strategy adoption.

- **Different power-consumption constraints for mesh nodes**.

  Mesh routers are almost stationary, and usually connected to the power infrastructure. However, mobile mesh clients are powered by battery and may meet the power efficiency requirements. Thus, the optimization objectives of networking protocols for mesh routers and clients may be different.

- **Compatibility and interoperability with existing wireless networks**.

Since WMN can be implemented in multiple wireless communication technologies, compatibility and interoperability will be solved in the radio access layer. The researchers may focus on the MAC layer and upper layers to design suitable protocols and applications.

Cellular Networks:

In the 1980s, the 1G wireless communication system came to the mobile communication environment, which provided a data speed of 2.4 Kbps to support data communication with mobile phones. An example is Nordic Mobile Telephone (NMT). However, this generation still worked in analog system and there were tight limitations in terms of the system capacity and data rate.

In the last decades of the last century, 2G wireless communication systems with increased capacity and higher speed gradually replaced the previous generation through technology development and performance enhancement. It was worth noting that 2G wireless systems supported digital communication, such as in Global System for Mobile Communication (GSM). After the transition to the 2G systems, some protocols were developed to increase the data speed that produced the 2.5G wireless systems. The General Packet Radio Service (GPRS), the most common one of those protocols, provided a speed up to 144Kbps. Later on, the 2.75G wireless system came with a higher data rate than previous generations and provided more enhanced performance in terms of high speed in data service. For example, by adopting advanced coding schemes and transmission mechanisms, Enhanced Data rates for GSM Evolution (EDGE) achieved high data-rate in term of wireless spectrum, which obtains a three-time increase in the user capacity and system performance than an traditional GSM/GPRS connection.

In the late 1990s, the 3G wireless communication system emerged with better multimedia capability and greater networking speed to meet the ever-increasing demand on data services. There are several widely used protocols and standards in the 3G systems, including UMTS, WCDMA, CDMA2000/EVDO, CDMA2000/EVDV, CDMA2000/EVDO-Rev A. In addition, for the economic concern, telecommunication operators usually adopted the evolution by efficiently integrating both 3G and previous generations to reduce the upgrade cost. Beyond 3G system was developed to improve the performance the conventional 3G system and provide a higher speed of up to 14.4 Mbps. As one of the technologies in the 3.5G wireless communication system, High Speed Downlink Packet Access (HSDPA) provides a dramatic performance improvement, based on the bandwidth's substantial increase and support more applications, such as graphics-intensive web browsing, on-demand video playing and multi-user video conferencing.

In the early twenty-first century, some telecommunication operators began to upgrade the 3G to 4G wireless communication systems, which provide a more comprehensive commercialized communication solution with a much higher data rate and better system performance, in terms of high reception ratio, low packet loss ratio and low packet delivery latency. International Telecommunication Union (ITU) has stated that 4G technologies require a data transmission rate of at least 100 Mbps while a user moves at high speed and a much high data rate up to 1 Gbps in a fixed location to support better multimedia applications, such as video-on-demand services. There are several challenges to achieve such performance criteria in mobile scenarios, including routing optimization technique, fast handover technique, integration technique between Mobile IP and cellular IP, multi-path technique, mobility management technique for all-IP networks.

In October 2010, ITU Radiocommunication Sector (ITU-R) completed the assessment of six candidate proposals for the future 4G mobile wireless broadband technology, called IMT-Advanced. Among these candidates, two technologies were accorded as the official designation of IMT-Advanced, including WirelessMAN-Advanced and LTE-Advanced.

WiMAX (Worldwide Interoperability for Microwave Access) is a telecommunication protocol [PLMD12], which is designed to provide high bandwidth mobile voice and data services and to serve as the potential wireless technology for the *last-mile* backhaul networks. The first version of IEEE 802.16 standard was released in 2001, and defined the basic air interface specification for wireless metropolitan area networks (MANs). Subsequently, physical layer technologies, such as orthogonal frequency division multiplexing (OFDM) and orthogonal frequency division multiple access (OFDMA), and new features, such as power-saving, idle mode, handover and an improved OFDMA physical layer, were amended to the original draft. The WirelessMAN-Advanced specification was incorporated in IEEE 802.16 standard beginning with approval of IEEE 802.16m standard.

In the evolution tree of mobile communication technologies, LTE (Long Term Evolution) is considered to be the latest standard. The first release, referred as LTE Release 8 [3GP09], was published in March 2009. The LTE-Advanced specification was developed by 3GPP as LTE Release 10 and Beyond. More details on enabling wireless technologies and protocols can be referred to our book [CZ14].

2.1.2 Practical Deployment and Applications

As VANET is still in the infancy stage, a rich body of research has been developed to validate the effectiveness of wireless communications in the presence of high vehicular mobility. Content distribution is one of the most frequently used services in vehicular Internet access, such as delivering road condition information, downloading multi-media resources. In [DNP04], SPAWN, a swarming protocol for vehicular ad-hoc wireless networks, is proposed to delivery content efficiently, which follows the design of the Bit-Torrent protocol. In SPAWN, a file with a unique ID is stored in the Internet and is divided into several pieces with unique sequence numbers. Vehicles download some pieces of the file when entering the communication range of APs due to the short contact time. Once leaving out

Table 2.1: Generations of Cellular Networks, 1

Generations	Protocols	Data Rate	Features
1G	AMPS, NMT, TACS	N/A	- Only support voice service - No data service
2G	CDMA, GSM, PHS	Up to 20 Kbps	- Digital voice service - Push-to-Talk (PTT) - Short message service (SMS) - Conference calling - Simple data applications such as email and Internet surfing
2.5G	GPRS, CDMA2000 1xRTT	114 Kbps(30-40 Kbps) 144 Kbps(60-80 Kbps)	All 2G features plus: - MMS(multimedia message service) - Internet surfing - Real-time location-based services, such as directions - Basic multimedia, including support for short audio and video clips, games and images
2.75G	EDGE, EGPRS, IMT-SC	384 Kbps 473 Kbps(ul) / 1.2 Mbps(dl) 600 Kbps	- Better performance for all 2/2.5G service
3G	WCDMA, CDMA2000 /EVDO, CDMA2000 /EVDO-RevA, TD-SCDMA	2 Mbps 2.4 Mbps 3.1 Mbps 2.8 Mbps	Support for all prior 2G and 2.5G features plus: - Fast web browsing - Streaming music - Full motion video - 3D gaming

Table 2.2: Generations of Cellular Networks, 2

Generations	**Protocols**	**Data Rate**	**Features**
3.5G	HSDPA /HSUPA, CDMA2000 /EVDO-RevB	14.4 Mbps 46 Mbps	Support for all prior 2/2.5/3G features plus: • Faster Internet surfing (especially graphics intensive sites) • On-demand video • Video conferencing
4G	WiMAX, UMB, LTE	100 Mbps 35 Mbps 100 Mps	Support for all prior 2G/3G features plus: • High quality streaming video • High quality Video conferencing • High quality Voice-over-IP (VoIP)

of the range, they exchange information cooperatively to obtain the all pieces of the file. In [LsPY$^+$06], Code-Torrent is proposed to improve the efficiency of the Bit-Torrent-like swarming protocol, suffering from collection problem of redundant pieces. By adopting network coding technique in content delivery, Code-Torrent is greatly helpful to the efficiency and effectiveness of content delivery in VANETs and shorten the file downloading time.

Wireless communication and networking between vehicles enable pervasive computing in vehicular environments, where seamless access to the Internet and other vehicles is supported. Thus, ubiquitous Internet access in vehicular environments is realized in the initial stage. When more and more vehicles are connected together in vehicular Internet access networks, the network is not only a traditional communication network, but it is also a mobile social network connecting more and more people in vehicles. For example, they can exchange safety related information opportunistically and share personal interests/goals purposefully. In the framework, FleaNet [LLPG10], a virtual market place on vehicular networks, is proposed to provide a virtual flea market in urban area. FleaNet develops a quality channel for people to communicate with each other and match the common interests to achieve potential transactions. In the networks, vehicles and road-side Ad-stations (Advertisement Stations, RSAS) can publish and forward information and queries.

2.2 Related Work on Infrastructure Planning

It has become a trend to deploy and manage city-wide or nation-wide Wi-Fi networks to promote the informatization in recent decades. It is common for the cities to adopt IEEE 802.11 standard as the mainstream wireless technology and provide their citizens and visitors high quality Internet access. Since IEEE 802.11 works in unlicensed wireless channels and spectrums, Wi-Fi operators do not need to purchase the operating radio-frequency resources. In addition, Wi-Fi is a pervasive wireless technology, which is widely installed in most of mobile terminals nowadays. Furthermore, the deployment of Wi-Fi networks is much convenient and less expensive than other networks. However, it is not a easy task to satisfy quality of service and user experience in the network. In the process of deployment, network upgrade is a continuous work to meet the increasing and varying demands.

Therefore, the time-consuming and costly infrastructure planning for vehicular Internet access system have to involve serval parities, such as government and private companies. To accelerate the procedure, an incremental and cost-efficient solution to vehicular infrastructure planning is in highly demand.

2.3 Related Work on Protocol Adaption

In dynamic vehicular networking environments, adaption should be taken into consideration in protocol design. There is much effort which has been devoted for improving the system performance of vehicular Internet access, ranging from MAC layer protocols [ZZC09] [Sik10], routing protocols [ZKL$^+$07], transport [OK05] and application layer [HGNH09]. In [ZZC09], VC-MAC is proposed to improve the throughput and system coverage of vehicular Internet access networks by leveraging cooperative communications among vehicles and road-side units. In [Sik10], the authors propose a reservation-based MAC protocol for vehicular communications. In this work, APs in the network will broadcast the modified beacon packets with the available transmission information. The in-coming vehicle will check the information and then associate the appropriate APs with the available time slots reserved for following transmissions. In [LLS12], the authors analyze the impact of the legacy IEEE 802.11 DCF to vehicular environments and target to some insights for effective utilization for IEEE 802.11 in vehicular Internet access networks, since DCF protocol is very important to Wi-Fi networks and emerging vehicular networks.

On the other hand, it is also important to understand the impacts of vehicles' mobility patten on the system performance in vehicular Internet access. In [TLYH11], the authors devise an analytical model to evaluate the downloading performance of vehicles in Internet access networks.

2.4 Related Work on Service Enhancement

In vehicular applications and services, vehicles play different roles according to their functionalities. We categorize them into four groups, including content suppliers, content consumers, both content suppliers and consumers, or intermediary.

For the first role, vehicles with Internet connectivity is a high quality mobile platform for mobile data collection in urban environments [LSK$^+$09][LZG$^+$06] [LMG$^+$09][HBZ$^+$06]. Vehicles equipped with all kinds of sensors, such as light, temperature, humidity sensors and cameras, can form a comprehensive powerful vehicular sensor nodes, which can not only collect raw sensing data, but also conduct pre-process on the raw data. Because there is no computation or energy constraint on vehicular sensor node, they can generate and handle much more data than the traditional wireless sensor nodes. The collected data can be forwarded to other vehicles or stored in local storage for later retrieval. In such category, the location information is required to label the raw data and find the potential consumers.

For the second role, vehicles also accounts for a significant majority of consumers of contents. Most of the data traffic is related to entertainment, such as music listening, movie watching and file downloading. In the scenario, the high throughput of vehicular networks is desirable to shorten the buffering delay and downloading time.

For the combination role of both suppliers and consumers of content, vehicles conduct information exchange between the Internet and other vehicles in the network. For example, the traditional vehicular applications for road safety can report their own road conditions and also receive other road conditions from neighboring vehicles. In addition, for some interactive applications, such as VoIP application and on-line gaming, vehicles need to maintain the real-time connections to the Internet and other peers.

There are two distinguished requirements for the above applications from other non-vehicular applications.

1. Location awareness: As a typical scenario of mobile networks, vehicular Internet access networks are highly location dependent not only for data collection but also for data delivery. This characteristic of vehicular Internet access networks provide us some insight on service enhancement. For example, the performance of vehicular applications and services can be upgraded by considering location-based information. In addition, the security and privacy issues related to the location information should be addressed in the design of vehicular applications and services.

2. Time sensitivity: There are two types of time sensitivity in vehicular Internet access

networks. One is caused by the short contact between vehicles and infrastructure networks. If vehicles miss the contact period, the delay of packet delivery is prolonged. The other one is caused by some time-sensitive applications and services. The data packet should be delivered within a given delay constrained by upper-layer applications and services.

From this perspective, vehicular localization service is a fundamental application in vehicular Internet access systems. In this thesis, we take vehicular localization service as an example to show how to enhance applications by leveraging the features of vehicular networking.

Chapter 3

Cost-efficient Infrastructure Planning for Vehicular Internet Access in Urban Area

In this chapter, we study the infrastructure planning in vehicular Internet access system. We develop a comprehensive analytical model for vehicular Internet access with the infrastructure support. This chapter is organized as follows. Section 3.1 gives the overview of the work. After that, we present the system model and analytic model for the infrastructure planning, respectively in Section 3.2 and Section 3.3. In Section 3.4, though extensive simulations based on the proposed model, we obtain some results on infrastructure planning in vehicular Internet access networks. In Section 3.5, as a special case of infrastructure planning, we present a incremental infrastructure planning problem by using wireless long distance links to reduce the upgrade cost. Finally, Section 3.6 concludes the whole chapter.

3.1 Introduction

In recent years, mobile Internet has changed people's daily life and always-on-the-Internet has become a lifestyle for teenagers. It is expected to be connected to the Internet everywhere and anytime, even in the moving vehicles for drivers and passengers. With the development of ubiquitous Internet access, people can improve their in-car experiences, making journey safer, efficiently and comfortable. Nowadays, most of Internet access, initiated by peoples' mobile devices, is loose-coupled with the vehicles. In the near future, more and more Internet-integrated vehicles will come out and achieve the complete Internet coverage in the vehicular environment.

Constrained by the current wireless access technologies in vehicles, cellular networks, such as GPRS, 3G and LTE, dominate the ubiquitous vehicular Internet access, as the infrastructure is well planned and widely available. However, wireless bandwidth crisis is always on the way to development of wireless communications for cellular operators, which results from the limited licensed spectrums. It is reported in Cisco's report that the number of mobile terminal which are connected to the Internet have exceeded the world's population by the end of 2014, and the monthly global mobile data traffic will reach 24.3 exabytes by 2019. As the ever-increasing demand for mobile broadband communication services, the availability of licensed wireless spectrums for cellular networks may become critical to guarantee service quality. Therefore, the service performance in vehicular Internet access simply using cellular networks may be degraded; even the performance of other non-vehicle users would also suffer due to the chain effect.

More and more wireless access technologies are adopted in the vehicular Internet access networks. IEEE 802.11 is an important alternative technology to achieve pervasive access. It is also predicted in the report that more data traffic in cellular networks will be offloaded to Wi-Fi networks by 2016. To supplement the wireless Internet access in vehicular environments, several types of infrastructure components are developed to satisfying different requirements. Among the components, In vehicular Internet access system, there are two major infrastructure components involved in the Infrastructure planning, namely base station and relays. Base stations are connected to the Internet and play the role of Internet gateways for vehicular access networks. Relays do not need direct connections to the infrastructure and can be installed where it is needed. In terms of the capital expenditure, the cost of base stations are much higher than relays, not only for the hardware cost, but also for the deployment and operation cost. In real practice, deployment and operation cost usually dominates the whole budget. In addition, each component also exists different operation modes. For instance, base stations can be located in a fixed position, or deployed in moving vehicles.

However, the performance improvement of infrastructure planning using different network entities is not fully investigated. There are several open questions to guide the network deployment. For example, if Internet access to base stations is not available, how many base stations or relays should be deployed to achieve the low delay communications? How to deal with the trade-offs between the deployment cost and network performance? To answer these questions, we derive a analytic model for the infrastructure planning in vehicular Internet access network. Based on the model, we analyze the benefits of each infrastructure component in the hybrid mobile networks.

The rest of this chapter is organized as follows. We present the system model and analytic model for the infrastructure planning, respectively in Section 3.2 and Section 3.3. In Section 3.4 validates the proposed analytic model via trace-based simulations. In Section 3.5, as a special case of infrastructure planning, we present a incremental infrastructure planning problem by using wireless long distance links to reduce the upgrade cost. Finally, Section 3.6 draws a conclusion to the entire chapter.

3.2 System Model

In this section, we present the system model adopted in the following analysis, including the network model, placement strategies for stationary nodes, mobility pattern of mobile nodes, traffic model and data forwarding model.

3.2.1 Network Model

In our model of vehicular Internet access networks, there are M mobile nodes and N stationary nodes. To facilitate the following analysis, the entire area covered by the considered networks is divided into k disjoint regions, as shown in Figure 3.1 when we choose $k = 2$. The stationary nodes can be placed in everywhere in each region. The placement strategy either follow the uniform distribution or non-uniform distribution, which will be detailed described in the following section. The mobile nodes can freely enter or leave any region in the area. When mobile nodes meet the stationary nodes or meet other mobile node on the move, there exist the pairwise contact times between them. We assume the pairwise contact times are represented by exponentially distributed random variables. Due to the instinct deployment nature, the mean contact times are different from the node movement. We assume the mean pairwise contact time between mobile nodes is $1/\beta_i$ and the mean time between stationary nodes and mobile nodes is $1/\gamma_i$. The mean rate of the time at which mobile nodes move between two regions is μ_{ij}, when the movement is from region i to j.

It is worth to note that in the above model we do not distinguish the mobile vehicles, mobile base stations and mobile relays in mobile nodes. And also for the stationary nodes, there are a faction of nodes working as base stations and the other part working as stationary relays. To improve the network performance, we assign different network functionalities to those nodes. Intuitively, base station is the best option to support the increasing application requirements. However, when the deployment and operation cost is taken into consideration as mentioned before, optimizing the network composition among different network devices is more important to achieve a cost-efficient high quality network deployment. We will analyze the incremental performance improvement by upgrading networks with more relays

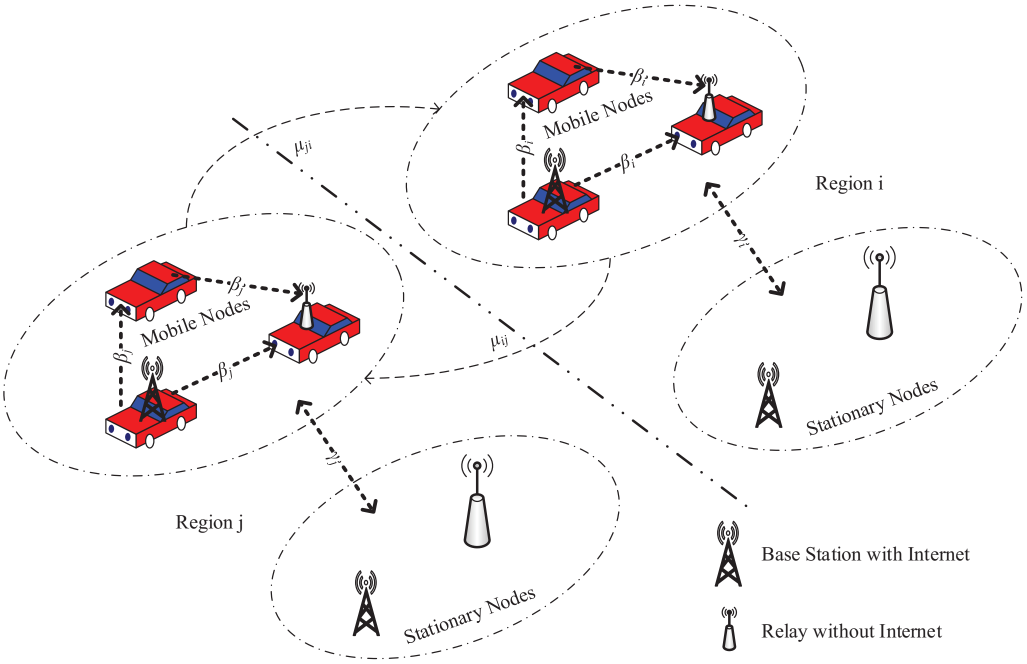


Figure 3.1: Network model with divided regions

and base stations.

In the following analysis, we adopt a simple yet general traffic model. All the traffic in the network is uni-cast and initiated by mobile nodes. The traffic initiators are randomly selected following the uniform distribution and all the traffic go to the Internet via the base stations. There are no correlation between different traffic flows.

3.2.2 Placement Strategies of Stationary Nodes

In this model, we adopt two strategies to place stationary nodes on top of the regions in the vehicular Internet access networks.

- **Uniform placement** In this strategy, we uniformly place the nodes across the entire geographic area and for each region the node distribution is also uniform. The only placement constraint is the geographic limitation, such as some buildings and obstacles.

- **Hotspot-based placement** In the entire geographic area, we first identify several regions with more mobile nodes and traffic flows, called hotspot areas. Based on these hotspot areas, we place more stationary nodes by adopting this strategy. We use a simple algorithm to generate the node distribution. The number of stationary nodes in each region is proportional to the square of the sum of duration when mobile nodes stay in the region. In each region, the placement still the uniform placement strategy.

3.2.3 Mobility Patterns of Mobile Nodes

In this model, we adopt two mobility patterns to change the position of mobile nodes on top of the regions in the vehicular Internet access networks.

- **Vehicular Random Waypoint Mobility Model** Random waypoint mobility model is a common random mobility model used in mobile networks. However, this model can not directly apply to vehicular Internet access network, due to the limitation by the road constraints. It has the similar parameters and operations as the original random waypoint mobility model, such as a uniformly distributed random pause time between a minimal pausetime (min_pausetime) and a maximal pausetime (max_pausetime) , and a speed between a minimal speed (min_speed) and a maximal speed (max_speed) with a uniform distribution. The difference between two models is on the change of direction. In vehicular random waypoint mobility model, mobile nodes do not always change their directions at when the pause times expire, until the next move is beyond the intersections.

- **Fixed Route Mobility Model** This mobility model extracted from the public transportation system. In this model, each mobile node has its own schedule and route to follow, just like the bus. The starting point, end point, starting time and direction of mobile node in each round is pre-defined. Once the mobility is started, the mobile node at the starting point first chooses a pause time between (min_pausetime, max_pausetime) with a uniform distribution. After the pause time expires, the node chooses a uniformly distributed random speed value between (min_speed, max_speed). And the iteration continues, until the end point is reached.

3.2.4 Data Forwarding Model

In our analytical model, we considered a general data forwarding model in which the packet exchanges when two nodes enter into the communication range. After packet exchange, base stations forward packets to the Internet and remove the packets. Other nodes, such as mobile vehicles and relays, will keep storing all the received packets until the packets are forwarded to base stations. As the traffic flows have not correlation, each packet is independent to all other packets.

This data forwarding model includes several non-linear differential equations for each disjoint region according to the considered network scenario with the following variables: $x_i^{mV}(t)$, $x_i^{mR}(t)$, $x_i^{mB}(t)$, $y_i^{sR}(t)$ and $y_i^{sB}(t)$. $x_i^{mV}(t)$ is the mobile vehicles receiving a packet within region i at time t. At the same period, the number of mobile relays receiving a packet within region i is $x_i^{mR}(t)$, and the number of mobile base stations receiving a packet within region i is $x_i^{mB}(t)$, respectively. Similarly, the number of stationary relays receiving a packet within region i is $x_i^{sR}(t)$, and the number of stationary base stations receiving a packet within region i is $x_i^{sB}(t)$.

The network successfully finishes one packet delivery once a mobile base station or stationary base station receives a packet. The objective is to determine the probability that the time when a packet is delivered to base stations is less than t, which is denoted by $P(t)$. Many network performance indicators can be derived from this probability. For example, the expected normalized delivery throughput is given by $\int_0^\infty \frac{1}{t(1-P(t))}dt$.

To simplify the derivation in the following analysis, we assume that all duplicated copies of a packet are dropped when the packet is successfully forwarded to a stationary base station or mobile base station in the following analysis.

3.3 Analytic Model

3.3.1 Preliminary Knowledge

We assume that the number of stationary and mobile nodes in region i is denoted by $n_i(t)$, which can be broken down as the following form.

$$n_i(t) = n_i^{mV}(t) + n_i^{mR}(t) + n_i^{mB}(t) + n_i^{sR}(t) + n_i^{sB}(t) \tag{3.1}$$

where $n_i^{mV}(t)$ is the number of mobile vehicles in region i at time t. At the same time, the number of mobile relays within region i is $n_i^{mR}(t)$, and the number of mobile base stations within region i is $n_i^{mB}(t)$, respectively. Similarly, the number of stationary relays within region i at time t is $n_i^{sR}(t)$, and the number of stationary base stations within region i at time t is $n_i^{sB}(t)$.

According to the network model, we can easily obtain the following equations.

$$\begin{aligned} M &= \sum_{i=1}^{k} n_i^{mV}(t) + n_i^{mR}(t) + n_i^{mB}(t) \\ N &= \sum_{i=1}^{k} n_i^{sR}(t) + n_i^{sB}(t) \end{aligned} \tag{3.2}$$

We assume that a fraction f_1 of mobile nodes are assigned as mobile base stations and another faction f_2 of mobile nodes are mobile relays. Similarly, a fraction f_3 of stationary nodes are assigned as stationary base stations and another faction f_4 of stationary nodes are stationary relays.

$$\frac{\sum_{i=1}^{k} n_i^{mB}(t)}{f_1} = \frac{\sum_{i=1}^{k} n_i^{mR}(t)}{f_2}$$

$$\frac{\sum_{i=1}^{k} n_i^{sB}(t)}{f_3} = \frac{\sum_{i=1}^{k} n_i^{sR}(t)}{f_4} \tag{3.3}$$

$$f_3 + f_4 = 1$$

In the following analysis, we will assign different values to those fractions to generate several special cases of infrastructure planning. For example, if we set f_1=1, f_2=0 and f_3=1, there is no relay in the network. However, at any instant of time t, the fractions of mobile nodes may vary with the vehicles' mobility. We assume that a fraction $f_1^{'i}(t)$ of mobile nodes are assigned as mobile base stations and another faction $f_2^{'i}(t)$ of mobile nodes are mobile relays at time t in region i respectively, which are varying with the node mobility. In addition, we assume $f_3^{'i}$ and $f_4^{'i}$ are the factions of stationary base stations and station relays in region i respective, which are constant with the time and only dependent of initial deployment of stationary nodes.

$$\frac{n_i^{mB}(t)}{f_1^{'i}(t)} = \frac{n_i^{mR}(t)}{f_2^{'i}(t)}$$

$$\frac{n_i^{sB}(t)}{f_3^{'i}} = \frac{n_i^{sR}(t)}{f_4^{'i}} \tag{3.4}$$

$$f_3^{'i} + f_4^{'i} = 1$$

3.3.2 Data Forwarding Analysis

In the following analysis, we consider one region to analyze the the data forwarding behaviors. The data forwarding operations are broken down, as shown in Figure 3.1. Recall that at any time t, the pairwise contact times between different nodes follows an exponential random distribution. And the mean pairwise contact time between mobile nodes is $1/\beta_i$ and the mean time between stationary nodes and mobile nodes is $1/\gamma_i$. The mean rate of the time at which mobile nodes move between different regions is μ_{ij}, when the movement is from region i to j.

Since we consider the vehicular Internet access, a packet is delivered when a stationary base station or stationary base station receives the packet. We highlight the changing rate of packet forwarding in base stations. The following differential equations show that such forwarding process in region i.

$$
\begin{aligned}
x'^{mB}_i(t) &= ((1 - f_1'(t))f_3'\gamma + \delta_i) \cdot n_i^{mB}(t) \\
&\quad + \sum_{i=1}^{k} \mu_{ji} x_j^{mB}(t) - \sum_{i=1}^{k} \mu_{ij} x_i^{mB}(t), \\
\\
y'^{sB}_i(t) &= ((1 - f_2'(t))f_1'(t)\beta + f_1'(t)f_2'(t)\beta \\
&\quad + f_1' f_4' \gamma / (f_1'(t) + f_2'(t))) \cdot n_i^{sB}(t), \\
\\
n'^{\,m}_i(t) &= \sum_{i=1}^{k} \mu_{ji}(n_j^{mV}(t) + n_j^{mR}(t) + n_j^{mB}(t)) \\
&\quad - \sum_{i=1}^{k} \mu_{ij}(n_i^{mV}(t) + n_i^{mR}(t) + n_i^{mB}(t)), \\
\\
x'^{mB}_i(t) &= 0, x'^{mB}_i(t) = 0, \\
\\
f_1'(0) &= f_1, f_2'(0) = f_2
\end{aligned}
\tag{3.5}
$$

Here, $x'^{mB}_i(t)$ presents the number of mobile base station newly receiving packets in region i. Mobile base stations have three ways to receive the packets, including from the mobile vehicles, from mobile relays and from stationary relays. In addition, the mobility also contributes to the change. Hence, $x'^{mB}_i(t)$ is the sum of the newly mobile base stations receiving data packets in region i and the rate of mobile base stations entering region i, minus the rate mobile base station leaving region i at time t. It is noting that δ_i, which means the rate of communication capability between stationary base stations and stationary relays in region i, is a constant value when the deployment of stationary nodes is fixed.

${y'}_i^{sB}(t)$ presents the number of stationary base station newly receiving packets in region i at time t. It is equal to the newly stationary base stations receiving data packets in region i from the mobile vehicles, from mobile relays and from stationary relays.

${n'}_i^{m}(t)$ presents the sum of all the mobile nodes changes in region i, which equals to the changing rate of the number of mobile stations entering region i, minus the changing rate of the number of mobile stations leaving region i at time t.

$$\begin{aligned} & P'(t) = \frac{\sum_{i=0}^{k} {x'}_i^{mB}(t)}{(1-f_1-f_2)M} + \frac{\sum_{i=0}^{k} {y'}_i^{fB}(t)}{f_3 N} \\ & P(0) = 0 \end{aligned} \tag{3.6}$$

$P'(t)$ is the changing rate of $P(t)$, presenting the changes in the number of both stationary base stations and mobile base stations receiving a packets respectively divided by the number of stationary base stations and mobile base stations around the network at time t, as all the traffic flows in the networks go to the Internet. The last set of equations shows the initial state of the network, where there is no traffic flows in the entire network.

3.3.3 Model Validation

To achieve the model validation, we conduct simulations based on the above model and compare the simulation results with the data trace from real vehicular Internet access networks.

The real trace comes from the DieselNet test-bed. In the test-bed, there are 40 buses and each bus is equipped dual Wi-Fi radios. One is for access point and the other is for channel scanning, which can shorten the channel/AP switching delays. Once the association between a station and an access point is completed, a data transmission is initiated and continues until the wireless link is disconnected as mobile vehicles runs out of communication range. After the transmission stops, the transmission activity is recorded with the receiver logs, including the node ID of the transmitter, the reception time, duration, and the number of bytes received.

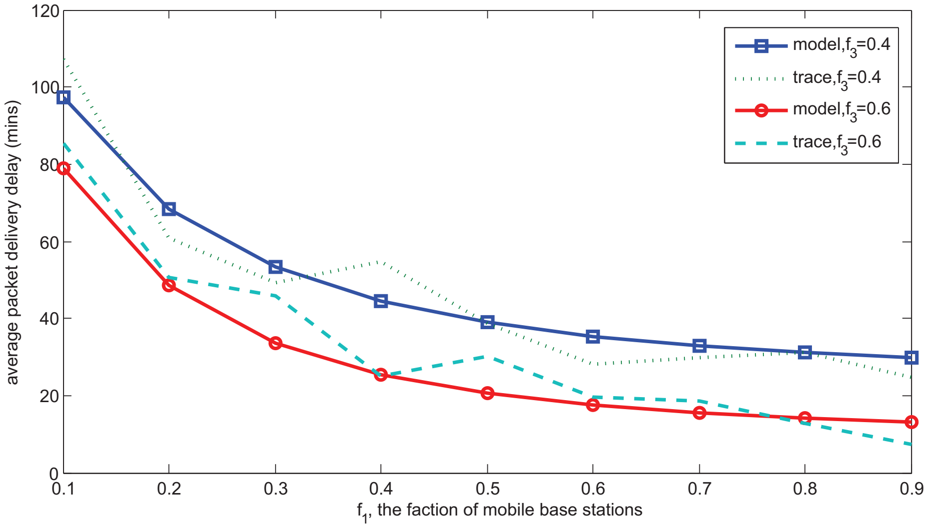


Figure 3.2: The average delay of packet delivery v.s. the faction of mobile base stations with 40 mobile nodes and 25 stationary nodes.

In the original data trace, there are only three kinds of network components considered in our model, including mobile vehicles, stationary base stations and stationary relays. To generate two more entities, some data processing on the original data trace should be adopted before the comparison. If one mobile node is selected as a mobile base station, all the out-going data flows will be removed from the data set. If one mobile node is selected as a mobile relay, all the data flow initiated by the mobile node should be removed. In addition, to extend the network scenarios, we should change the role of stationary nodes. If one stationary base station is changed to a stationary relay, all the data flow initiated by the base station should be removed.

In the simulation, we deploy 40 mobile nodes and 25 stationary nodes in the network.

In different simulation scenarios, we change the factions of mobile base stations and station base stations in each node group. In the first experiments, f_3 is set to 0.4 and 0.6, and the average delay of packet delivery is represented as a function of f_3, the faction of stationary base stations. In the second experiments, f_1 is set to 0.3 and 0.7, and the average delay of packet delivery is represented as a function of f_1, the faction of mobile base stations.

The two sets of simulations results are illustrated in Figure 3.2 and Figure 3.3, respectively. From the above figures, it is easy to find that the results from our proposed model confirm to the trace-based simulations from real deployment. In addition, the factions of stationary and mobile base stations contribute to different impacts of delay of packet delivery, which will be detailed in following parts.

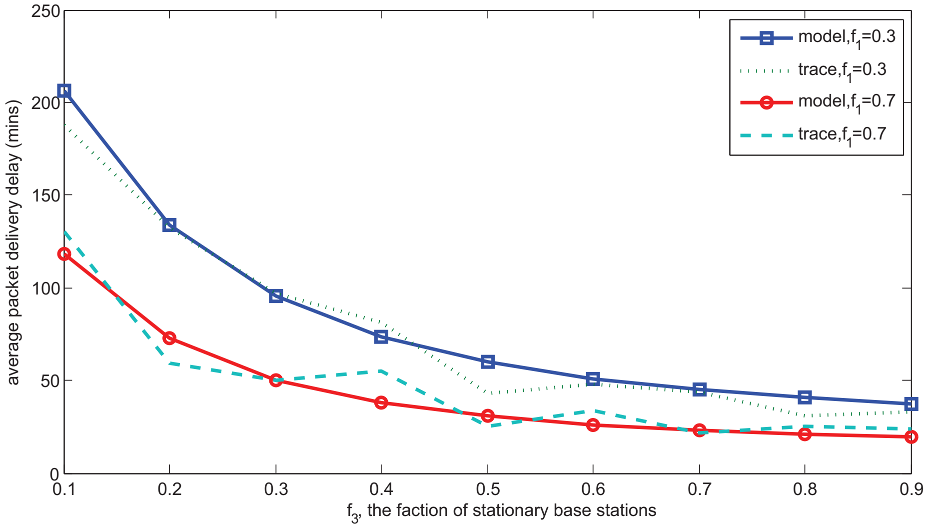


Figure 3.3: The average delay of packet delivery v.s. the faction of stationary base stations with 40 mobile nodes and 25 stationary nodes.

3.4 More Results from the Analytic Model

In this section, we conduct extensive simulations based on the above analytic model and investigate the effects on network performance imposed by the different number of infrastructure components and different types of infrastructure planning.

The number of stationary nodes

In this experiments, we set the total number of mobile nodes to 40 and the combinations of (f_1, f_3) to (0.3,0.4), (0.3,0.6), (0.7,0.4) and (0.7,0.6). The average delay of packet delivery as a function of the number of stationary nodes is shown in Figure 3.4. We can find that the number of the stationary nodes dominates the asymptotic average package delivery delay when the number is greater than some threshold determined by network scales.

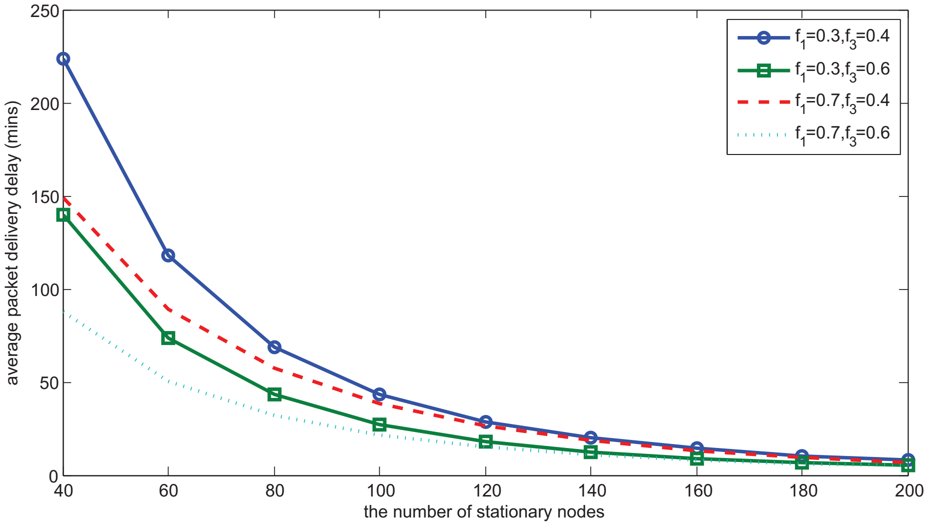


Figure 3.4: The average delay of packet delivery v.s. the number of stationary nodes with 40 mobile nodes.

The number of mobile nodes

In this experiments, we set the total number of stationary nodes to 25 and the combinations of (f_1, f_3) to (0.3,0.4), (0.3,0.6), (0.7,0.4) and (0.7,0.6). The average delay of packet delivery as a function of the number of mobile nodes is shown in Figure 3.5. It shows that the asymptotic average package delivery delay is determined by the number of mobile base stations. The delivery delays can approach different threshold values.

The faction of base stations in stationary nodes

In this experiments, we set the faction of mobile base stations f_3 to 0.5, and the combinations of the number of mobile and stationary nodes (M, N) to (60, 60), (60, 120), (120,

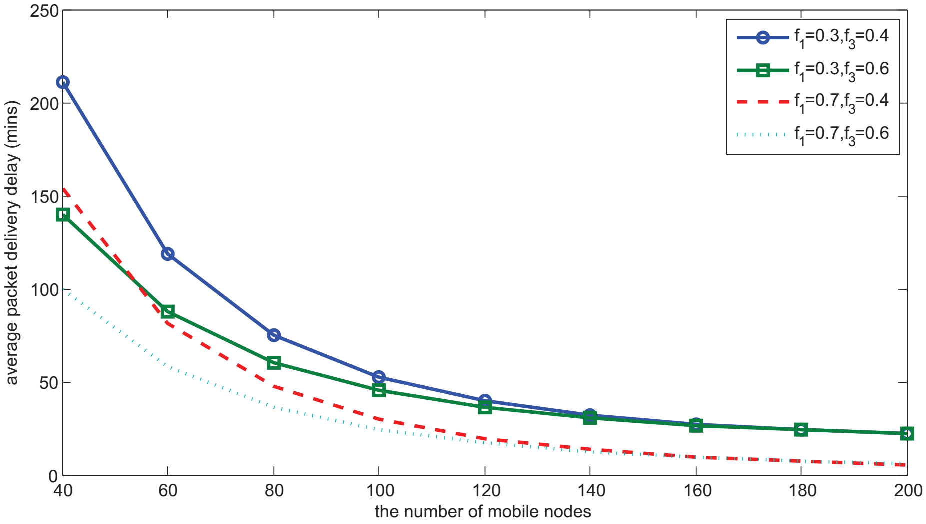


Figure 3.5: The average delay of packet delivery v.s. the number of mobile nodes with 25 stationary nodes.

60) and (120, 120), as illustrated in Figure 3.6. It shows that the average package delivery delay is determined by the faction of base stations in stationary nodes when the number of stationary nodes is fixed. The delivery delays can approach different threshold values according to different number of stationary nodes.

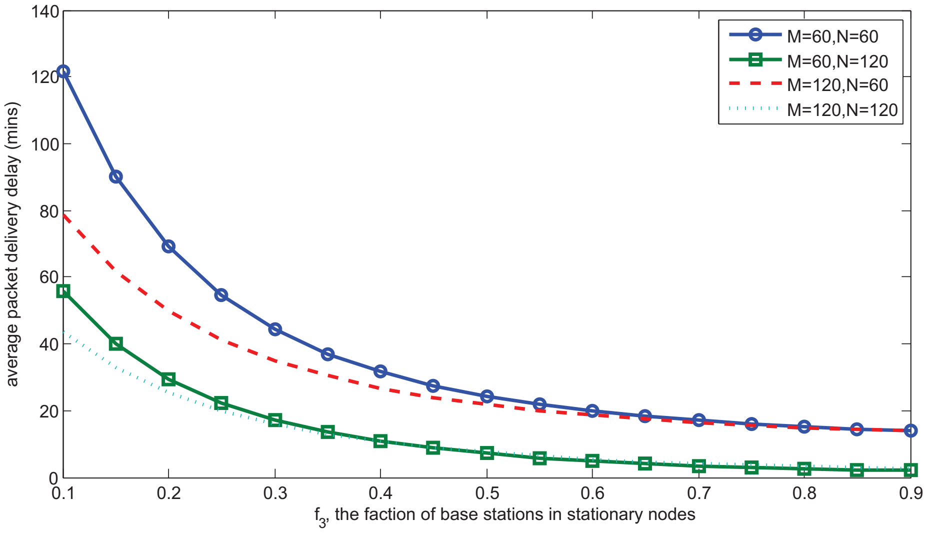


Figure 3.6: The average delay of packet delivery v.s. the faction of base stations in stationary nodes.

The factions of base stations in mobile nodes

In this experiments, we set the faction of stationary base stations f_1 to 0.5, and the combinations of the number of mobile and stationary nodes (M, N) to (60, 60), (60, 120), (120, 60) and (120, 120). The average delay of packet delivery is represented as a function of the faction of base stations in mobile nodes is shown in Figure 3.7.

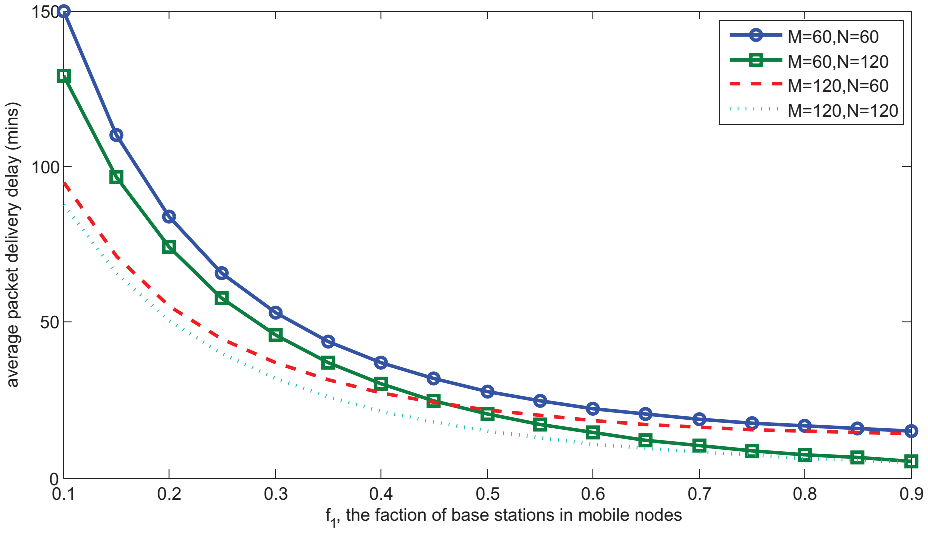
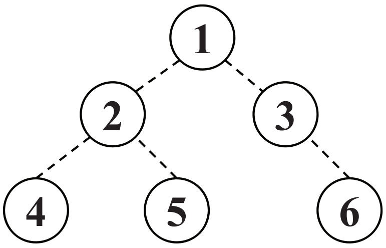


Figure 3.7: The average delay of packet delivery v.s. the faction of base stations in mobile nodes.

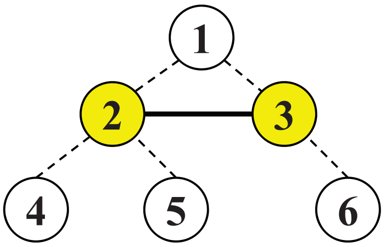
3.5 Incremental Cost-efficient Network Planning using Wireless Long Distance Links

As mentioned before, most of existing deployment of vehicular Internet access networks have been ad-hoc in nature and the planning was almost done manually without the well-tested planning[SR07]. As data traffic tends to increase continuously, network operators regularly face the necessity to upgrade their network to match the growing demands. Due to the economic and practical concerns on the deployment cost, it is impossible to upgrade and replace all the network devices. So, a practical way is to carry out the network upgrade in an incremental manner. We believe that the problem is becoming crucial for the existing wireless multi-hop networks.

In this section, we study the cost-constrained incremental network upgrade problem in multi-hop wireless networks. When the number of users or their traffic requirements increase and exceeds the capacity of an existing networks, an update is needed. The distribution of data traffic in wireless networks is not always uniform. The throughput of multi-hop wireless networks can be improved in a manner of deploying some additional quality links in heavy traffic areas to upgrade the network. For example in Fig.3.8a, the original network, formed as a tree topology rooted at node 1, is limited in performance by the bottleneck at node 1. When the traffic between two subtrees of node 1 increase, the performance will deteriorate. If a link is established between node 2 and node 3 by adding directional antennas as shown in 3.8b, the network performance will be improved.



(a) Original network



(b) Upgraded Network

Figure 3.8: An example to illustrate benefit from incremental upgrade.

As illustrated by the above idea, this section considers the problem of how to improve the network performance by adding long distance directional wireless links under the constraint of the deployment cost for an existing network. We first identify the components of deployment cost in the incremental network upgrade, and analyze the relationship between the network throughput and the network upgrade. The deployment of links between the different node pairs causes different effects on the network throughput. Based on the analysis, we formulate the Incremental Cost-efficient Network Planning ($ICNP$) problem, which is an NP-hard problem. Then, we propose a heuristic algorithm to solve the problem. The algorithm uses a performance metric that consider both performance improvement and deployment cost. Finally, we conduct the numerical and network simulations to evaluate the performance of our algorithm from different aspects.

Recent years have witnessed an increasing number of deployments [Ash][Ara][Nep]and applications [Dju][Ram] adopting IEEE 802.11 into the large area wireless networks, especially in rural areas and some developing countries, as backbone networks for the access to the Internet. It is mainly because the frequency spectrum of IEEE 802.11 is free of charge and the cost of IEEE 802.11 devices is much lower than that of other technologies. This kind of networks is also called the Wi-Fi over Long-Distance Wireless Networks (WiLDNets).

There are some characteristics of these deployments and applications in WiLDNets. First, the distant neighboring nodes are almost equipped with high-power wireless interface and high-gain directional antennas to construct the long-distance links, as shown in Figure.3.9, which results in the different interference model and concurrent communications conflict relationship between links. For example, it is observed that the inter-radio interference is much severer due to the high-power wireless interface cards and radio frequency leakage from the side lobes of directional antennas. Therefore, Simultaneous Synchronized Operation (SynOP) is proposed to avoid the such interference, where two adjacent directional links that either transmit simultaneously (SynTx), or receive simultaneously (SynRx).

Second, the long propagation delay of the long-distance links is another salient characteristic from the common wireless networks. As we know, the round-trip delay is very critical

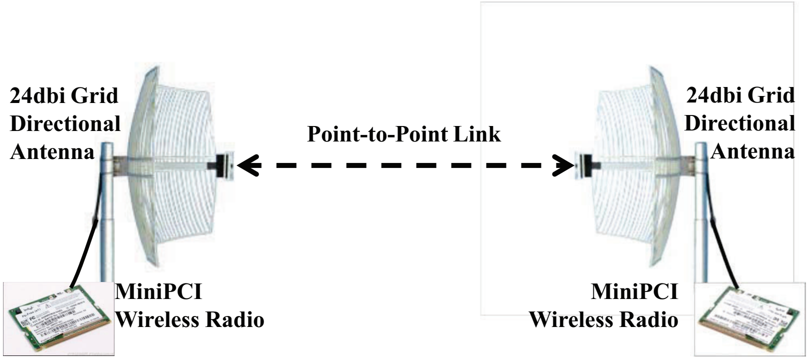


Figure 3.9: A example of long-distance link formation via the high-gain grid directional antennas.

to some ACK-based protocols and applications, such as MAC layer Data-ACK mechanism. Few work on link scheduling in multi-hop wireless networks considered the round-trip delay due to the short air time of ACK frame. However, in WiLDNets, the long propagation delay makes link scheduling more complicated when considering link conflicts during the ACK phase.

Furthermore, in long-distance networks, the transmission power of the radios is always set to the high level to achieve the large coverage, which results in the severe interference and electromagnetic pollution to the environment. On the other hand, the higher transmission power allows the higher data rate in the wireless links. So, the power control mechanism is quite essential to the WiLDNets to minimize the energy consumption while maintaining the network performance.

Network Model

A multi-hop wireless network is modeled by a connected undirected graph $G = (V, E)$, where $V = \{v_1, ..., v_n\}$ is the set of n nodes and E is the set of possible wireless links. We assume that all links are symmetric and the routing paths between each pair of nodes are determined by a pre-installed routing protocol based on routing metrics, which are very critical to the network performance of the network[DPZ04, YW08]. These network-specific routing metrics can capture the characteristics of the target networks and as well as measurement of the network performance [RS07]. In this work, we use routing metrics as the basis of defining performance metrics for the wireless network. We define the path routing utility function (PRU) in the network as follows.

Definition 1. *(PRU) Given a routing path between node v_i and v_j, denoted as $p_{i,j} = \{v_i, ..., v_k, ..., v_j\}$, and the routing metric function $H(\cdot)$ adopted in the routing protocol, the* path routing utility (PRU) *of $p_{i,j}$ is defined as*

$$PRU(p_{i,j}) = \sum\nolimits_{i \leq a \leq j-1} \frac{H(v_{a,a+1})}{H(v_{i,j})} \cdot \ln \frac{1}{H(v_{a,a+1})} \tag{3.7}$$

where $H(v_{i,j})$ is the routing cost between source and destination nodes i and j, and $H(v_{a,a+1})$ is the routing cost between two neighboring nodes a and $a+1$ along the path. Based on PRU, we can easily obtain the definition of network routing utility (NRU) in the whole network.

Definition 2. *(NRU) Given a network with a pre-installed routing protocol, the* network routing utility (NRU) *of $G(V, E)$ is defined as*

$$NRU(G(V, E)) = \sum\nolimits_{v_i, v_j \in V, i \neq j} PRU(p_{i,j}) \tag{3.8}$$

Especially, when deploying a set of directional links S to an existing network $G(V, E)$, the whole network routing utility will be improved. Therefore, we can define the incremental upgrade utility (IUU) in $G(V, E)$ due to the extra links in S as follows.

Definition 3. *(IUU) Given a network $G(V, E)$ and a set of links S, the* incremental upgrade utility (IUU) *of link set S in $G(V, E)$ is defined as*

$$IUU(S) = NRU(G(V, E \cup S)) - NRU(G(V, E)) \tag{3.9}$$

Deployment Cost Function

An important and practical consideration in network planning is the deployment cost during the network deployment and maintenance. In this work, we categorize the deployment cost into two major classes, including one-time hardware cost used to deploy extra network interfaces and wireless antennas, and long-time operation cost to maintain these extra links. When deploying long distance wireless links, the adoption of different antennas depends on the distance between the two endpoints, which requires different device costs in hardware part. When operating these links, the maintenance cost stems from the length of the links, which involves the failure diagnosis and device replacement. Some research works have already showed that the cost is usually proportional to the square of the length of the links [Ltd, Sta]. Thus, the cost function $C(i,j)$ when deploying a directional link $l_{i,j}$ from node i to node j can be formally defined as

$$C(l_{i,j}) = K_1 * d_{i,j} + K_2 * d_{i,j}^2 \tag{3.10}$$

where K_1 and K_2 are constants, which can be obtained by some empirical data. In the performance evaluation part, we will refer to the prices of some off-the-shelf wireless devices and some technical reports to approximate the values of K_1 and K_2.

Problem Statement and Hardness

Given a multi-hop wireless ad hoc network $G(V,E)$, where a set D of nodes have the capability to set up a high quality link with some deployment cost, the problem is to find a subset W from $D \times D$ to set up such kind of links to maximize the network routing utility between all node pairs under the constraint of a bounded deployment budget. Therefore, we define the incremental cost-constrained network planning ($ICNP$) as follows.

Definition 4. *($ICNP$) Given an undirected connected graph $G(V,E)$ and $D \subseteq V$, find*

$W \subseteq D \times D$, where $W \cap E = \emptyset$, such that

$$NRU(G(V, E \cup W)) \tag{3.11}$$

is maximized subject to the following constraint

$$\sum\nolimits_{w_i \in W} C(w_i^1, w_i^2) \leq L \tag{3.12}$$

where $w_i(w_i^1, w_i^2)$ is the link between node w_i^1 and w_i^2.

We can prove that the special instance of the budgeted maximum coverage problem [KMN99], which is a classical NP-hard problem, can be reduced to $ICNP$. Due to limitations of space, we omit the derivation here.

3.5.1 Heuristic for $ICNP$

Instead of searching through all the possible combinations for link deployment, we introduce a new performance metric, *cost efficiency*, to represent the link deployment preference, and propose a greedy link deployment ($GLiD$) algorithm.

Cost Efficiency

We first introduce some notations to be used in the following algorithm description of the work. For a connected undirected $G(V, E)$, a set D of nodes has the capability to set up a high quality link with some deployment cost. To simplify the presentation below, we denote the deployable links in the network by $L_D = D \times D$, the weight function for any link $l \in L_D$ as $w(l) = IUU(l)$ and the deployment cost function for $l \in L_D$ as $c(l) = C(l)$.

To improve the network routing utility while reducing the total deployment cost, we define a new metric called *cost efficiency* as

$$\varphi(l) = IUU(l)/c(l). \tag{3.13}$$

Generally, the *cost efficiency* metric strikes a balance between the network performance and the deployment cost. Increasing incremental upgrade utility helps improve the efficiency of the routing protocol and reduce the path length, which will help improve network throughput while keeping the cost below some threshold. In the following sections, we will experimentally validate the correctness of the metric.

Greedy Link Deployment Algorithm

In this section, we give the $GLiD$ algorithm to the $ICNP$ problem. Generally speaking, $GLiD$ is based on the limited enumeration method and greedy searching technique. Limited enumeration methods rely on the observation that enumeration methods often find the better solutions at the initial stages and widely used in the practical and engineering fields.

$GLiD$ uses a subroutine called $GLiDSearch$, which is listed in Algorithm 1. $GLiDSearch$ follows the best-first search algorithm, which finds the solution by expanding the most promising node chosen according to a specified rules. The rules here are the link with high *cost efficiency* $\varphi(l)$, as shown at line 2 of Algorithm1, and the bounded total deployment budget L at line 3. After checking all links in the unselected link set, $GLiDSearch$ finds all satisfied links and adds them to the selected sets constructing the candidate solutions.

$GLiD$ proceeds in an iterative manner. At the beginning of the k_{th} iteration round, the algorithm first selects the link with maximum weight under the bounded total deployment budget into the selected set. Then, from the unselected link set, the algorithm enumerates all available links and apply the $GLiDSearch$ algorithm to find the greedy solutions. Until there are no available links in the candidate link set, the algorithm repeats the above steps and maintains the best solution from all possible solutions. The pseudocode of the $GLiD$ algorithm is shown in Algorithm 2.

When $GLiD$ is executed into the iterations, it involves the $GLiDSearch$ algorithm at line 8 in Algorithm 1. As proven in [KMN99], the approximation ratio of the greedy

```
Algorithm 1: GLiDSearch Algorithm
Input  : A connected graph $G(V, E)$, a set of unselected links $U$, a set of selected links $S$,
         the weight function $w(\cdot)$, the cost function $c(\cdot)$, and the deployment budget $L$;
Output: The selected subset $W \in S$
The selected subset $W \in S$; while $U \neq \emptyset$ do
    select $l = argmax\{\varphi(l)|l \in U\}$, if  $c(S) + c(l) \leq L$ then
        $S \leftarrow G \cup \{l\}$,
    end
    $U \leftarrow U \setminus \{l\}$.
end
return $W \leftarrow S$;
```

searching algorithm is $(1 - 1/\sqrt{e})$ to the optimal solution, where e is the base of natural logarithms. In the *GLiD* algorithm, after enumerating the all possible candidate combinations, it adopts the *GLiDSearch* subroutine to get the suboptimal solutions, so Algorithm2 can achieve the approximation ratio of $(1 - 1/\sqrt{e})$ in a loose way.

```
Algorithm 2: GLiD Algorithm for ICNP
Input  : A connected graph $G(V, E)$, a set of possible links $L_D$, the weight function
         $w(\cdot)$, the cost function $c(\cdot)$, and the deployment budget $L$;
Output: The selected subset $W \in L_D$;
$W_1^0 \leftarrow \emptyset; r \leftarrow 1$; while $|W|! = |L_D|$ and $\exists l \in L_D \setminus W, w(W) + w(l) \leq L$ do
    $l = argmax\{w(l)|l \in L_D \setminus\ W_1^{r-1}$ and $c(W_1^{r-1} \cup \{l\}) \leq L\}$, $W_1^r \leftarrow W_1^{r-1}$; $W_2^r \leftarrow \emptyset$
    $W \leftarrow W_1^{r-1} \cup \{l\}$ for $\forall l_i \in \{l_i|l_i \in L_D \setminus W_1^r$, and $c(W_1^r \cup \{l_i\}) \leq L\}$ do
        $S \leftarrow W_1^r \cup \{l_i\}$, $U \leftarrow L_D \setminus S$ $GLiDSearch(U, S)$ if $w(S) > w(W_2^r)$ then
            $W_2^r \leftarrow S$
        end
    end
    if  $w(W) < w(W_2^r)$ then
        $W \leftarrow W_2^r$
    end
    $r \leftarrow r + 1$.
end
return $W$;
```

3.5.2 Performance Evaluation

In this section, we evaluate the performance of our proposed algorithm via simulations. To validate the correctness of the *GLiD* algorithm to solve the *ICNP* problem, we use Matlab to conduct the numerical simulations and compare the performance improvement with the random selection algorithm and purely greedy selection algorithm. To validate the effectiveness of the *cost efficiency* metric to represent the tradeoff between network performance and cost, we use Qualnet simulator to evaluate the performance of those algorithms.

Numerical Simulations

As shown in Fig.3.10, a connected wireless ad hoc network with 100 stationary nodes are randomly generated in a $1500m \times 1500m$ open area, where the regular communication range is 250 meters. In the simulation, we assume the candidate nodes are labeled from 1 to 10, which are square vertexes in the figure. So, there are $10 * (10 - 1)/2 = 45$ candidate directional links in the network. The transmission range of regular nodes is 250 meters, while the range of the to-be-deploy directional links is 2500 meters.

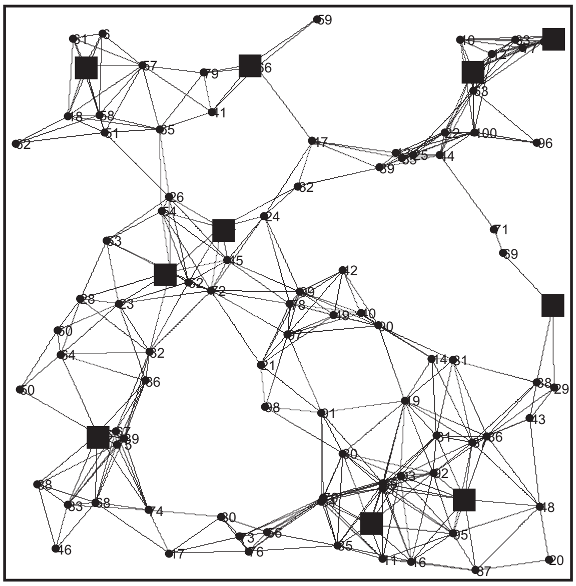


Figure 3.10: Network topology in the simulations.

For the parameters used in calculating the deployment cost, we refer to the prices from off-the-shelf commodity wireless devices and the technical reports [Ltd, Sta] from industry. As shown in Table 3.1, to form one directional links between two endpoints with multiple interface cards and directional antennas, we assume $K_1 = 400$, which is the estimated hardware cost of two candidate nodes (each one needs one motherboard with multiple slots, two different network cards and two different antennas). For maintenance cost, we assume K_2 is equal to an empirical value 0.001.

Table 3.1: Cost for generic wireless devices.

Devices	Cost(USD)
Motherboard with 1 miniPCI slot	40
Motherboard with 3 miniPCI slots	80
Unidirectional antenna(5dbi)	3
Grid directional antenna(24dbi)	30
Wireless network interface card(100mw)	20
Wireless network interface card(400mw)	50

We compare our algorithm to other two heuristic algorithms, the random selection algorithm (RSA) and IUU-Priority selection algorithm($IPSA$). The former one selects the candidate links randomly from the available link sets, while the latter one selects the links according to the best incremental upgrade utility in link sets. We consider two performance metrics in the simulations. The first one is the incremental upgrade utility, which is the direct performance indicator for different algorithms. The other one is the cumulative distribution function (CDF) of routing cost in terms of different routing metrics between any two nodes in the network after using different algorithms to conduct upgrade.

First, we take hop count as routing metric and form the wireless multi-hop network. In Fig.3.11, we set different budget constraints to the problem and compare the distribution of IUUs under different algorithms. The budget and the IUUs in the figure are both normalized. As we can see from the results, the $GLiD$ algorithm outperforms the RSA and $IPSA$ almost over the whole budget span. It is worthy to notice that although the performance of the RSA is better than the $GLiD$ algorithm occasionally in the budget interval from 0.7 to 0.9, the heavy fluctuations show that the RSA cannot provide any performance guarantee to the problem.

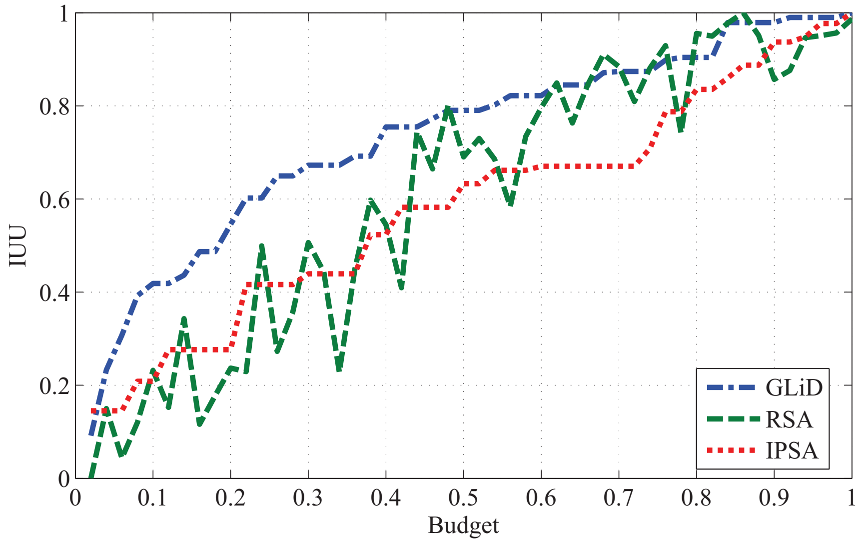


Figure 3.11: Numerical simulation result - IUUs under the different budget constraints.

In Fig.3.12, we show the CDF of hop counts between any two nodes in the network after using different algorithms, when we set the budget constraint to 0.3. It shows that our *GLiD* algorithm can reduce 80% node pairs within 5 hops, while other two algorithms only have about 70% node pairs.

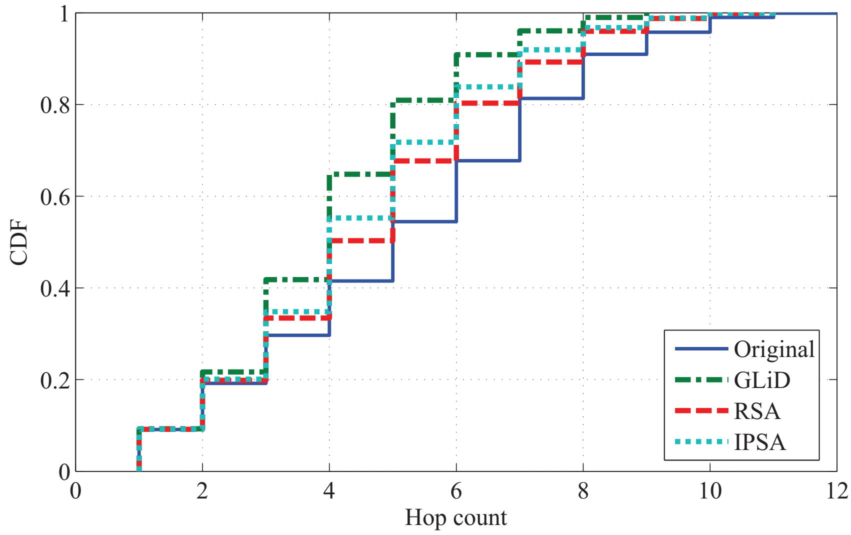


Figure 3.12: Numerical simulation result - CDF of hop counts.

Then, we take *ETX* as routing metric and form the wireless multi-hop network and we show the CDF of ETX between any two nodes in the network after using different algorithms, when the budget constraint is set to 0.3. It shows the similar results in Fig.3.13, and $GLiD$ can reduce the maximum ETX to about 1.75, which is much less than other two algorithms.

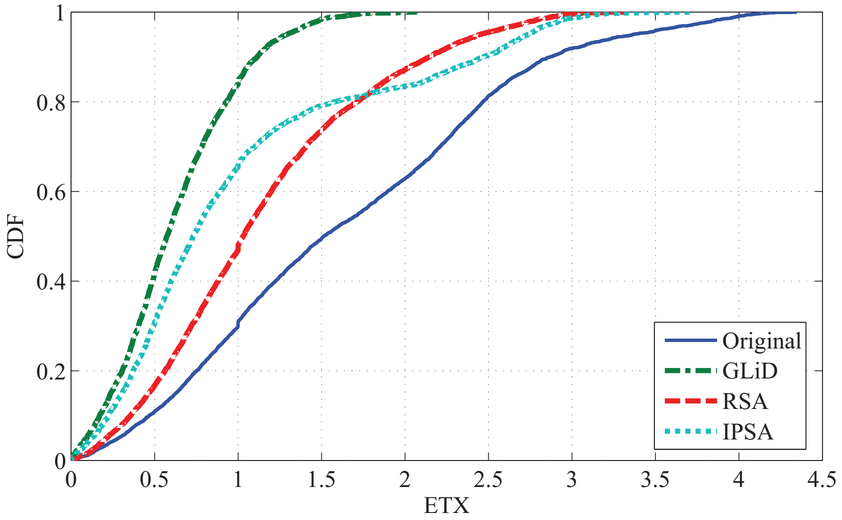


Figure 3.13: Numerical simulation result - CDF of ETX metrics.

Qualnet Simulations

To evaluate the network performance gains from the incremental upgrade, we have implemented our algorithm in Qualnet, which is a discrete event simulator used in the simulation of wireless networks. In the simulations, IEEE 802.11b MAC layer and physical layer models are adopted for regular nodes, and the transmission range is set at 250 meters. OLSR is used as the routing protocol, because we can configure different types of routing metrics for the routing protocol, *hop count* and ETX.

For those deployable point to point wireless links, IEEE 802.11a MAC layer and physical layer models are used, and the transmission range is set to 2500 meters. In Qualnet, the default transmission of the IEEE 802.11 is limited to about 500 meters, because the simulator hardcodes the signal propagation delay to 1ms. Due to the ACK and timeout mechanisms in IEEE 802.11, the senders never receive the ACK until the timeout expires. In our simulation, we modify the source code of Qualnet to support the long distance wireless links.

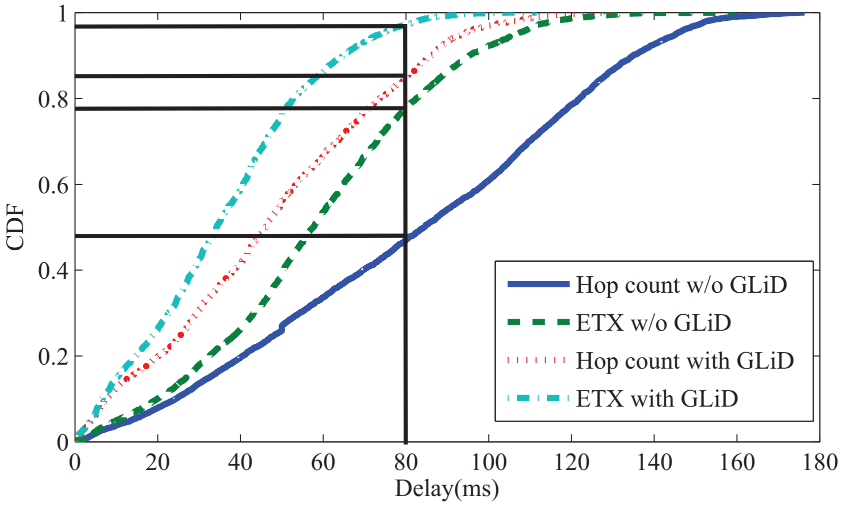


Figure 3.14: Qualnet simulation result - CDF of delay of packet delivery.

The network topology is the same as the aforementioned numerical simulations, and we export the node locations from the locate matrix in Matlab to the node configuration file in Qualnet. And, we randomly initialize 100 CBR (constant bit rate) sessions between different node pairs among the total 100 nodes. We consider four scenarios, including original network topology with hop count and ETX routing metrics, and upgraded network by $GLiD$ with hop count and ETX routing metrics.

In Fig.3.14, we compare the PDD among the 100 random CBR sessions, and find that the upgraded network by the $GLiD$ algorithm with ETX routing metric has the best performance, where the most PDD among all the sessions are less than 80 milliseconds.

In Fig.3.15, we compare the ST among the 100 random CBR sessions, and some similar results are revealed from the results. The upgraded network by the $GLiD$ algorithm with ETX routing metric outperforms the other network scenarios, where 80% sessions achieve the throughput below 500Kbps.

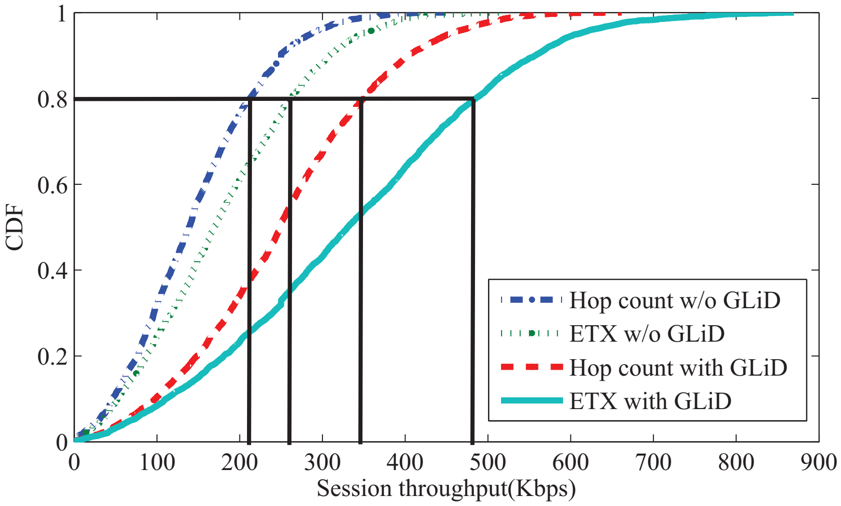


Figure 3.15: Qualnet simulation results - CDF of session throughput.

3.6 Summary

In this chapter, we investigate the problem how to deploy different types of network components, such as base stations and relays, in a cost-efficient way to maximize the network performance. First, we propose a detailed analytical model for infrastructure-based vehicular Internet access in the presence of different network components. Then, through extensive simulations based on the proposed model, we obtain some guidelines on choosing and deploying different network components in vehicular Internet access networks. Finally, as a special case study of infrastructure planning, we investigate the incremental and cost-efficient network planning (ICNP) problem in vehicular Internet access networks with the objective of improving network capacity for an existing network under limited deployment cost. After defining a *cost efficiency* metric, namely incremental upgrade utility (IUU), we develop a heuristic algorithm, called Greedy Link Deployment algorithm (GLiD) with approximation ratio $(1-1/\sqrt{e})$ for deploying additional wireless links to the existing networks to achieve the maximum performance improvement. Through extensive simulations, the results show that our algorithm can effectively exploit the available resources and perform much better than other heuristic algorithms.

Chapter 4

Connectivity-Oriented Data Dissemination Algorithm for Vehicular Internet Access

In this chapter, we study the maximum network connectivity problem in vehicular Internet access networks. We first analyze the characteristics of network connectivity in vehicular Internet access networks and derive a closed form to represent the connectivity probability under specific road-side units (RSU) deployment model and vehicle arrival rate. Then, we develop a connectivity-oriented data dissemination (CoDA) protocol to maximize the minimum network connectivity in vehicular Internet access systems based on the probabilistic analysis. Finally, we validate the CoDA protocol through extensive simulations and the results.

4.1 Introduction

Mobile Internet has already changed a lot of things in our daily life, including the way to communication and information retrieval. As a typical scenario of mobile Internet, vehicular Internet access network plays an important role in the total Internet usage. For example, some traditional vehicular applications can be transplanted to vehicular Internet networks, including emergence applications on road safety, applications on transportation efficiency and management, and information and entertainment applications. Although most of them can be done with vehicle-to-vehicle communications, the information can be shared to city-wide people over the Internet in assistance of infrastructure networks. In recent years, some innovative vehicular applications exploiting the capability of Internet access have provided value-added functionalities to both drivers and passengers, such as Uber and Bus Arrival Time. This is a strong motivation for people using Internet access in vehicular environments nowadays.

In traditional VANETs, vehicles can exchange data packets with each other within vehicle-to-vehicle communication range. To improve data dissemination across the entire vehicular networks, node-pair based connectivity between vehicles should be greatly satisfied and the inter-vehicle contact time should be prolonged to increase the information exchange. However, in the presence of infrastructure support in vehicular Internet access networks, the objective of network connectivity becomes the connectivity between vehicles and the infrastructure, which means we can improve the network connectivity by leveraging the infrastructure networks. For example, even though two vehicles in a vehicular Internet access network can not communicate directly with each other or indirectly by other vehicles, they could be still connected via the backend network if they are both connected to the infrastructure networks.

In this paper, we investigate the maximum network connectivity problem in vehicular Internet access networks, where vehicles adopt both vehicle-to-infrastructure and vehicle-to-vehicle communications. First, we analyze the characteristics of network connectivity in vehicular Internet access networks and derive a closed form equation to represent the connectivity probability under specific RSU deployment model and vehicle arrival rate. Then, we develop a connectivity-oriented data dissemination (CoDA) protocol to maximize the network connectivity in vehicular Internet access networks based on the probabilistic analysis. Finally, we validate the CoDA algorithm through extensive simulations, driven by empirical vehicular traffic traces.

The following contributions have been made in this paper.

- A closed form equation to represent the connectivity probability under specific RSU deployment model and vehicle arrival rate is derived. From the expression, the impact of some critical factors, such as vehicles' velocity, RSUs' distribution and communication ranges, are analyzed based on the model.
- A connectivity-oriented data dissemination (CoDA) algorithm has been developed to maximize the minimum network connectivity in vehicular Internet access networks.

The rest of this paper is organized as follows. We firstly introduce related work mainly on vehicular Internet access and connectivity analysis in Section 4.2. In Section 4.3, we present the connectivity analysis based on the system model in our paper. In Section 4.4, we formulate the problem for maximum connectivity probability and present the design of CoDA. After that, we validate our algorithm by trace-based simulation in Section 4.5 and draw conclusion in 4.6.

4.2 Related Work

4.2.1 Existing Vehicular Internet Access

We classify vehicular Internet access into two categories: drive-thru based vehicular Internet access and delay tolerant networks (DTN) based vehicular Internet access.

Drive-thru based Vehicular Internet Access

In drive-thru based VIA systems, moving vehicles connect to the road-side APs and access to the Internet via intermittent connections. In [OK04], a testbed including a single vehicle and a road-side AP in an empty road segment was set up. Based on the testbed, the evaluations of TCP and UDP transmission between vehicle and AP were conducted under different vehicle's velocities. After that, more works have been taken to evaluate the performance of vehicular communications in various network environments [OK04, CC09, ZLSK10]. These works further confirm that vehicular Internet access based on road-side Wi-Fi APs is feasible and capable to non-interactive applications. However, due to the unplanned deployment nature and limited coverage of road-side Wi-Fi networks, most of the works assume that vehicles adopt opportunistic access to the Internet and do not consider the connectivity requirements for the upper layer applications.

DTN based Vehicular Internet Access

Vehicular networks can be considered as a special type of DTNs and and focuses on leveraging vehicle to vehicle (V2V) communication to deliver content. In [WFGH04], a vehicular data dissemination algorithm (MDDV) was proposed to leverage geographical information and forward data to the destinations, which focuses on the vehicular mobility. A vehicle-assisted data delivery system (VADD) [ZC08] studied the problem how to find the best relay nodes based on traffic information. In TBD [JGG+11] and TSF [JGG+12], trajectory information of vehicles was exploited to forward data in uplink and downlink communications. In addition, Maxprop [BGJL06] considered the limited relay buffer in vehicular networks and determine how to forward or drop packets in communication period. However, this category of the work focused on the data replication and forwarding policies in vehicular networks and required modifications on network entities and protocols on both on-board units and road-side APs.

4.2.2 Connectivity Analysis in Vehicular Networks

Connectivity is an important property in wireless networks and much effort has been taken in vehicular networks to model the connectivity in high dynamic network and analyze impacts caused by connectivity. In [FL04], the authors considered a one-dimensional

VANET with vehicles in uniform distribution and derived a connectivity formula under the protocol communication model. In [MA06], a physical communication model was adopted in connectivity analysis by taking channel fading into consideration. In the presence of infrastructure support, network connectivity was addressed in [DTH02] and [KPD$^+$09] by considering the connection between vehicles and road-side APs. In [NZZ$^+$11], 2-hop connectivity probability for infrastructure based vehicular networks was derived under Poisson traffic model and a generic connection model. Similarly in [AZ11], the authors proposed an analytical framework to characterize the probabilistic delay of packet delivery under the random vehicular traffic and disrupted wireless channels.

However, most of the works above focus on the node-pair connectivity, which is quite important in MANETs or infrastructure-less VANETs. In infrastructure-based VANETs, especially in vehicular Internet access networks, we focus on the connectivity between vehicles and the Internet rather than the connectivity between vehicles.

4.3 System Model

In this section, we first present the system model considered in this paper, including the deployment model of road side units and the mobility model of moving vehicles. Then, a closed form of network connectivity probability is derived and some theoretical results are also analyzed. Finally, we will introduce the problem formulation of the connectivity-oriented data dissemination problem.

4.3.1 System Model

In this paper, we consider the vehicular Internet access system which consists of moving vehicles in the streets and stationary road-side units (RSU) along the streets. The communication ranges of vehicles and RSUs are the same according to the IEEE and industrial standard. Let R be the communication range of wireless communications. Vehicles equipped with one wireless network interface card (NIC) are capable of both vehicle-to-vehicle and vehicle-to-infrastructure communications, and they can communicate with RSU directly or through multi-hop forwarding among neighboring vehicles. They can also set to the two modes simultaneously, since most drivers of wireless NICs provide the Virtual-AP (VAP) functionality, which creates multiple virtual interfaces with different modes, including access point and ad hoc modes. In addition, all possible wireless interferences are ignored. In this part, we only consider the connectivity probability in terms of the parameters related to road layout and vehicle density. In the next section, we take the interference into consideration to design the routing protocol.

Figure.4.1 illustrates the system model with vehicles and RSUs in the road layout. RSUs are deployed along the road and cover different ranges on the road due to different distances between mount-points and roads. For example, the radiuses of coverage areas by RSU_1 and RSU_2 are r_1 and r_2, where $r_1 < R$ and $r_2 < R$. Let d be the distance between RSU_1 and RSU_2. We can easily find that if $d =< r_1 + r_2$ the network is always fully connected no matter how many vehicles drive in the road. Therefore, we only consider the scenario with $d > r_1 + r_2$ in the following analysis.

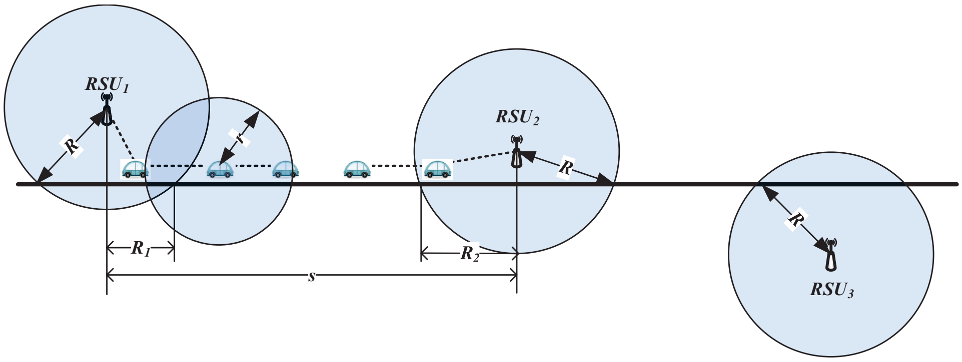
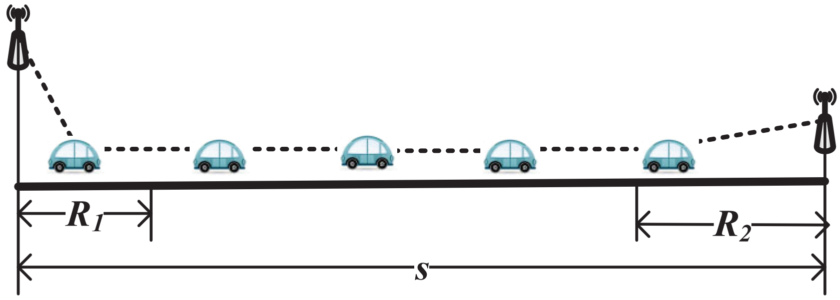
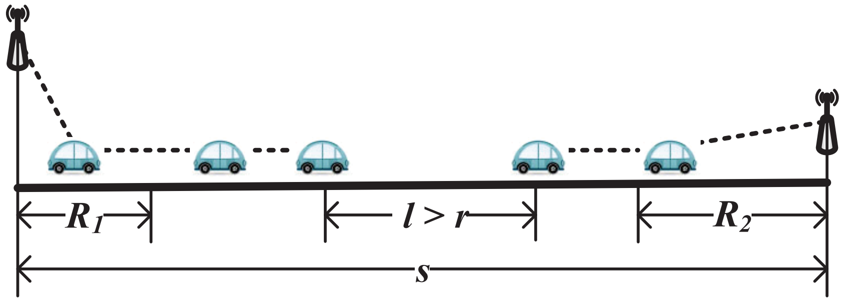


Figure 4.1: The system model of vehicular Internet access system

We only investigate the region bounded by two adjacent RSUs from the original road layout, and call any section in the region as a road segment. According to the communication paradigm of vehicles in our system model, we can identify the two possible network topologies where all vehicles can communicate with RSUs due to different vehicles' distribution, as shown in Figure.4.2. If the vehicles are distributed evenly, all vehicles in the segment may be connected to the road-side infrastructure, while the whole network may be partitioned into two parts when there exists at least one road segment where the length is greater than R having no vehicles. In the following part, we will analyze the corespondent network connectivity probabilities respectively.



(a) Fully connected group between RSU_1 and RSU_2



(b) Two fully connected networks associated with RSU_1 and RSU_2

Figure 4.2: Two possible cases of fully connected network.

4.3.2 Connectivity Analysis without Infrastructure Support

In transport engineering, a macroscopic traffic model can be used to predict the asymptotic state of a road system, such as the capacity and results from flow regulation or speed limits [GP15]. The model involves serval elements related to the road system, including traffic flux, traffic density and vehicle velocity. There are four main statements derived by the model, listed as follows.

- Relation between vehicle density and velocity: more vehicles moving on a road slow down their velocity.
- Congestion prevention: the number of incoming vehicles to a congestion area should be less than or equal to the number of outgoing vehicles in the same time.
- Critical traffic density: at this critical point, the flow state of the road traffic will become unstable from previous stable state.
- Traffic flow collapse: in the unstable traffic flow area, the flow would collapse if any vehicle brakes, which incurs much severe congestion.

In our analysis, the length of the considered road segment is s and we assume that the vehicles are Poissonly distributed with density ρ. So, the probability that there exist n vehicles on a road segment with length s is given as

$$p_1(n,s) = e^{-\rho s}\frac{(\rho s)^n}{n!}, n \geq 0. \tag{4.1}$$

We first consider the problem at what probability all the n vehicles on a road segment of s are fully connected. To tackle this problem, we use the nearest-neighbor methods to analyze the connectivity between vehicles. In nearest-neighbor methods, the nearest neighbor function is defined as the probability distribution of the distance from a specific point to the nearest neighbouring point under a certain point process. As assumed in a homogeneous Poisson point process N on a d-dimensional Euclidean space R^d, the nearest neighbour function is defined as $D_x(r) = 1 - e^{-\lambda|b(x,r)|}$, where λ is the intensity measure and $|b(x,r)|$ represents the Lebesgue measure of the space.

In our scenario, the length of the road segment is largely greater than the width. So, we can reduce the space to one dimension and λ is set to ρs according to the above traffic model. Therefore, the nearest neighbour function of a specific vehicle can be derived as $D_x(s) = 1 - e^{-\rho s x}$. After taking derivative with respective of x, we can obtain the nearest neighbour function of any vehicle in the road segment is given as

$$D(x) = \rho s \cdot e^{-\rho s x}. \tag{4.2}$$

Thus, the probability that the nearest neighbour distance of any vehicle in the network is not greater than the communication range of vehicles r can be derived as

$$P(x|x \leq r) = \int_{x=0}^{r} D(x)dx = 1 - e^{-\rho s r}. \tag{4.3}$$

From the above equation, we can obtain the probability that in road segment with length s no vehicle can be covered in the communication range, which is equal to the he probability that the nearest neighbour distance of any vehicle in the network is greater than the communication range.

$$p(0,s) = P(x|x > r) = 1 - P(x \leq r) = e^{-\rho s r}. \tag{4.4}$$

When we consider the network scenario where n vehicles denoted by $V_1, V_2, ..., V_n$ sequentially are deployed in the considered segment s. Therefore, $p_2(n,s)$ the probability that n vehicles located are fully connected in road segment with length s is given as

$$p_2(n,s) = (1 - p(0,s))^n = (1 - e^{-\rho s r})^n. \tag{4.5}$$

Combining Equation (4.1) and (4.5), we can derive the probability that all vehicles are fully connected in road segment with length s is given as

$$p_3(n,s) = p_1(n,s)p_2(n,s) = \frac{(\rho s)^n}{n!}(1 - e^{-\rho sr})^n e^{-\rho s}. \tag{4.6}$$

To provide a sense of the effects of different parameters on the network connectivity p_3, we conduct some numerical simulations based on the above equation. We choose the length of the considered road segment is 2000 meters. The node density and communication range of vehicles are set to 1/200 (vehicle/meter) and 100 meters respectively for baseline case. In the first simulation, we change the parameter ρ and check the impact of node density on network connectivity, as shown in Figure 4.3. We can easily find that the distribution of connectivity probability follows the similar manner and the probability value always reaches a peak at a certain length of the selected road segment. If the the parameter ρ is decreased, the peak value of connectivity probability drops, but the length value at this peak increases.

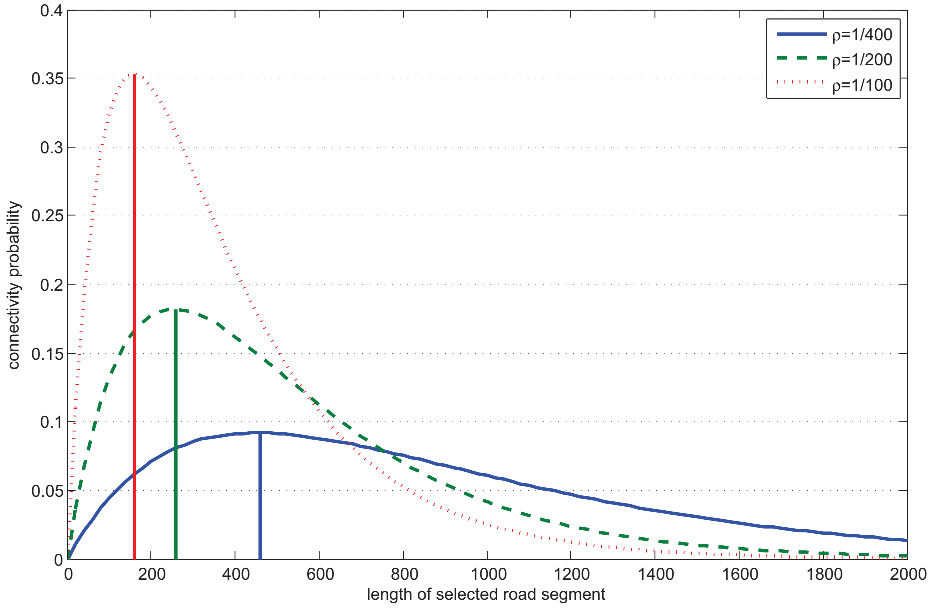


Figure 4.3: The impact of parameter ρ on connectivity probability p_3.

In the second simulation, we change the parameter r and check the impact of communication range on network connectivity, as shown in Figure 4.4. Similar to the previous one, if the the parameter r is increased, the peak value of connectivity probability and the length value at this peak are both grow up.

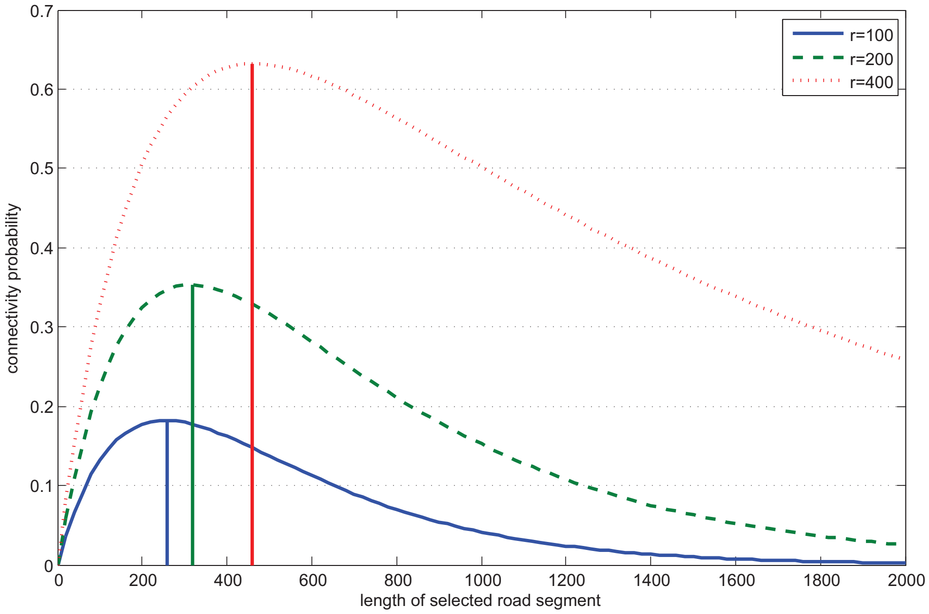


Figure 4.4: The impact of parameter r on connectivity probability p_3.

It is worth noting that Equation (4.6) only considers the network connectivity among vehicles without infrastructure support. In the following parts, we will analyze the connectivity probability of vehicular access networks within road-side units.

4.3.3 Connectivity Analysis with Infrastructure Support

As mentioned before, if all vehicles are connected to the Infrastructure, there exist two possible network topologies, as shown in Figure 4.2. We set the considered region at $[0, s]$ for ease of description, where the two RSUs are located at the both ends with the coverage ranges of R_1 and R_2. We assume that R_1 and R_2 are greater that the communication range of vehicles, r, since road-side units usually have more powerful antenna system to enlarge the communication range.

For the first case in Figure 4.2a, all vehicles are connected when the following conditions hold, as shown in Figure 4.5, including 1) there is at least one vehicle in the segment $[0, R_1]$, 2)there is at least one vehicle in the segment $[s - R_2, s]$, and 3) the vehicles in the segment $[0, s]$ are fully connected. For the prior two conditions, we can calculate the probability according to the $p(0, R_1)$ and $p(0, R_2)$. For the last condition, the probability is given in Equation 4.6. Thus, the probability of full connectivity in the first case is derived as

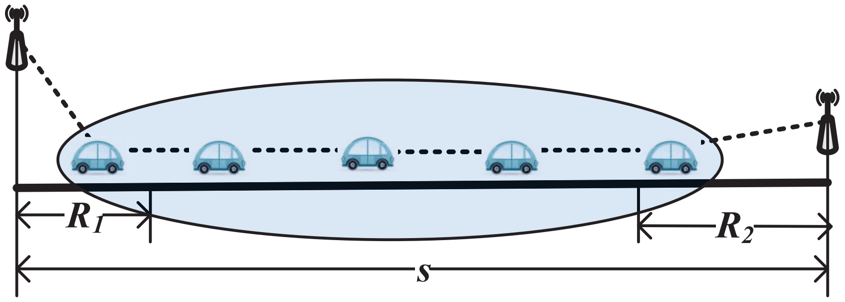


Figure 4.5: Fully connected group between RSU_1 and RSU_2.

To provide a sense of the effects of different parameters on the network connectivity p_4, we also conduct some numerical simulations based on the above equation with the similar simulation settings.

In the first simulation, we change the parameter ρ and check the impact of node density on network connectivity, as shown in Figure 4.6. We can easily find that the distribution of connectivity probability follows the similar manner and the probability value always reaches a peak at a certain length of the selected road segment. If the the parameter ρ is increased, the peak value of connectivity probability is increasing and the length value of the selected road segment at this peak is smaller. Compared with Figure 4.3, the peak values is less than the values in connectivity probability p_3, and the length of the selected road segment at peaks is smaller than the value at peak in connectivity probability p_3, as well. It is because

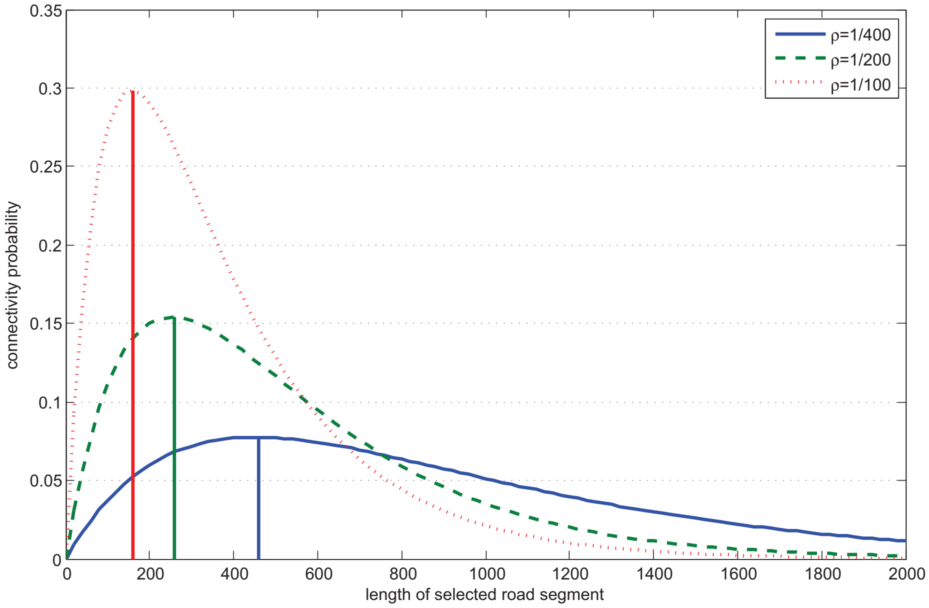


Figure 4.6: The impact of parameter ρ on connectivity probability p_4.

the probability is lower when both condition 1) and 2) hold.

In the second simulation, we change the parameter r and check the impact of communication range on network connectivity, as shown in Figure 4.7. Similar to the previous one, if the the parameter r is increased, the peak value of connectivity probability is increasing, but the length at this peak is greater. Compared with Figure 4.3, the peak values is less than the values in connectivity probability p_3, and the length values of the selected road segment at peaks are closer than that in connectivity probability p_3.

For the second case in Figure 4.2b, the network is actually partitioned into two networks around the two RSUs respectively. All vehicles are connected to the Internet when the following conditions hold, as shown in Figure 4.8, including 1) all vehicles are connected to the Internet in the segment $[0, d]$, 2) all vehicles are connected to the Internet in the segment $[s - d - R, 0]$ and 3) there is no vehicle in the segment $[d_1, d_1 + r]$.

To calculate the probability of network connectivity in the prior two conditions, we divide the entire road segment into three parts, namely $[0, R_1 - r]$,$[R_1 - r, s - R_2 + r]$ and

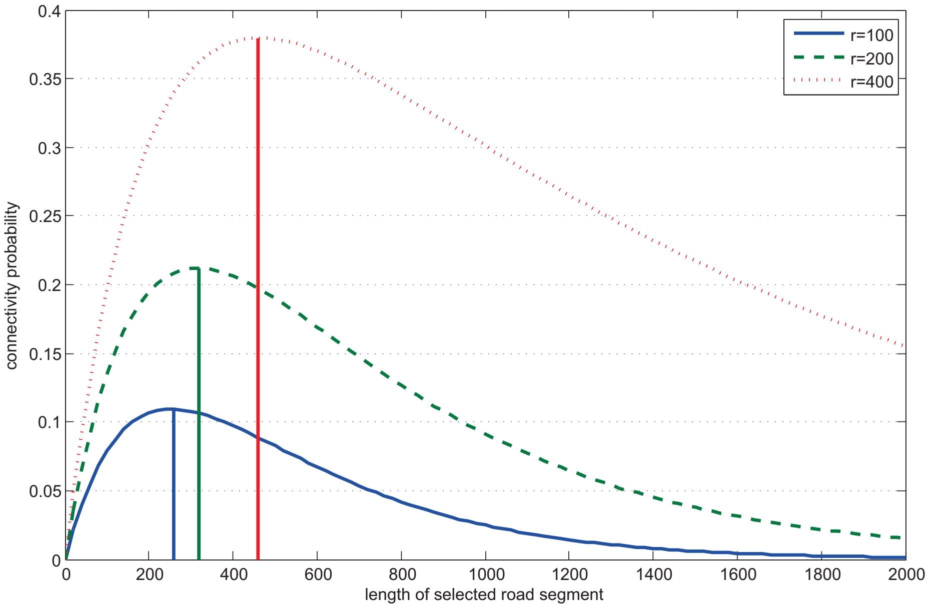


Figure 4.7: The impact of parameter r on connectivity probability p_4.

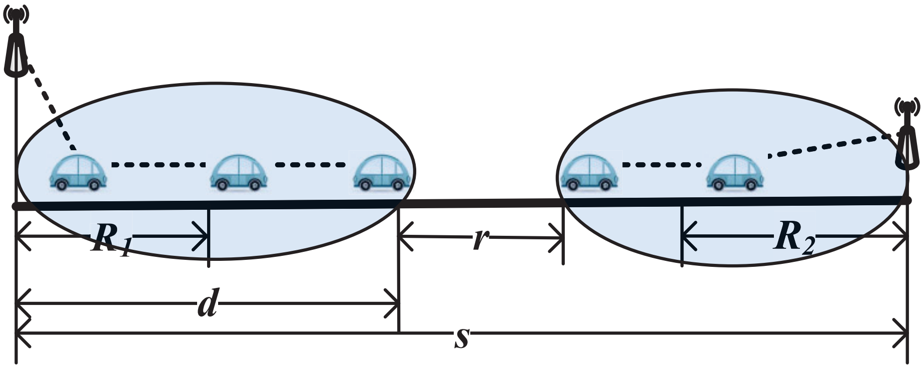


Figure 4.8: Fully connected group between RSU_1 and RSU_2.

$[s-R_2+r,0]$. In the road segment $[R_1-r,s-R_2+r]$, we can obtain the connectivity probability using the similar method employed in derivation of $p_4(n,s)$.

$$\begin{aligned} p_4(n,s) &= p_3(s)(1-(1-p(0,R_1))(1-p(0,R_2))) \\ &= e^{-\rho s}\frac{(\rho s)^n}{n!}(1-e^{-\rho sr})^n(1-(1-e^{-\rho sR_1})(1-e^{-\rho sR_2})) \end{aligned} \tag{4.7}$$

The probabilities that condition 1 and condition 2 hold are given as $p_3(m,d)(1-p(0,d))$ and $p_3(n-m,s-r-d)(1-p(0,s-r-d))$, respectively, wherr. Thus, the probability of full connectivity in this segment is derived as

$$p_5^1(n,s) = \int_{d=R_1-r}^{s-R_2+r}\int_{m=0}^{n} p_3(m,d)(1-p(0,d))p_3(n-m,s-r-d)(1-p(0,s-r-d))p(0,r) \tag{4.8}$$

In the road segment $[0,R_1-r]$, as shown in Figure 4.9, all vehicles in segment $[0,d]$ are obviously connected to the Internet since they are in the communication coverage of RSU1. If there exists any vehicle in road segment $[d,R1]$, the two adjacent networks will emerge into one network since the vehicle in $[d,R1]$ can directly communicate with RSU1. We exclude the probability that there exists any vehicle in road segment $[d,R1]$ in calculation. Thus, the probability of full connectivity in the segment $[d,R1]$ is given as

$$p_5^2(n,s) = \int_{d=0}^{R_1-r}\int_{m=0}^{n} p_3(n-m,s-r-d)(1-p(0,s-r-d))p(0,R_1-d) \tag{4.9}$$

Similarly, we can get the probability of full connectivity in the raod segment $[s-R_2+r,s]$, as denoted as $p_5^3(n,s)$,

$$p_5^3(n,s) = \int_{d=s-R_2+r}^{s}\int_{m=0}^{n} p_3(m,d)(1-p(0,d))p(0,r)p(0,R_2-d) \tag{4.10}$$

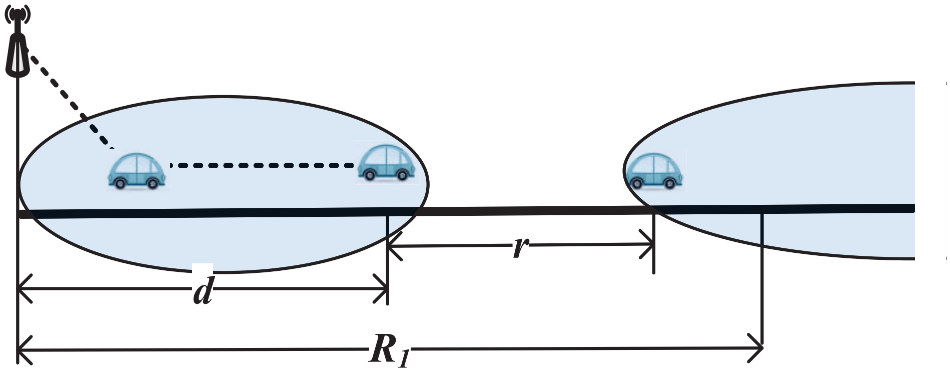


Figure 4.9: Fully connected group between RSU_1 and RSU_2.

According to Equation (4.8), Equation (4.9) and Equation (4.10), we can obtain the closed form expression of network connectivity probability for the second case to achieve full connectivity, as dented as $p_5(n,s)$,

$$p_5(n,s) = p_5^1(n,s) + p_5^2(n,s) + p_5^3(n,s) \tag{4.11}$$

In the numerical simulations, we follow the similar network settings and change the parameter ρ to check the impact of the node density on network connectivity, as shown in Figure 4.10. We can find that the decrease of ρ can also suppress the peak value of connectivity probability and the length values at peaks decrease as well. The relation between the communication range r and network connectivity is shown in Figure 4.11.

In the above derivation of connectivity probability p_4 and p_5, any network setting is independent to the others among the all random network settings. Thus, we can sum up p_4 and p_5 to obtain the connectivity probability for all cases as follows.

$$p_6(n,s) = p_4(n,s) + p_5(n,s) \tag{4.12}$$

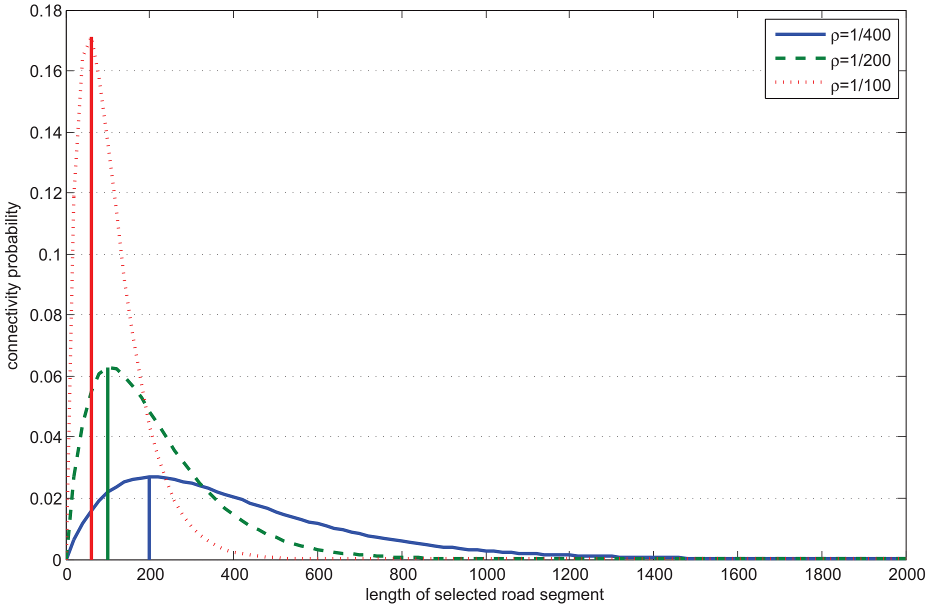


Figure 4.10: The impact of parameter ρ on connectivity probability p_5.

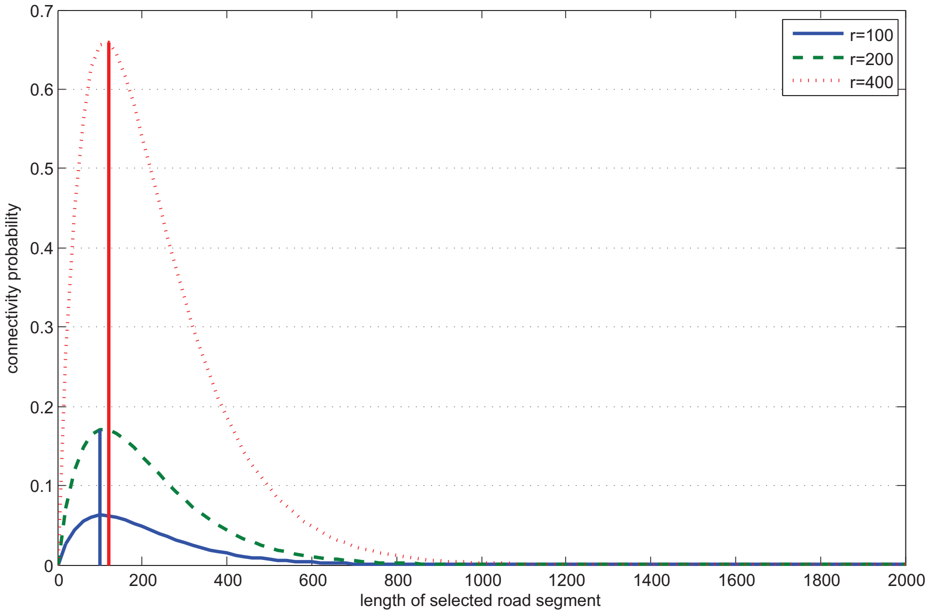


Figure 4.11: The impact of parameter r on connectivity probability p_5.

Similarly, to evaluate the effect of impact factor ρ on the connectivity probability p_6, we run the numerical simulations by changing the value of ρ, as shown in Figure 4.12 and the value of r, as shown in Figure 4.13.

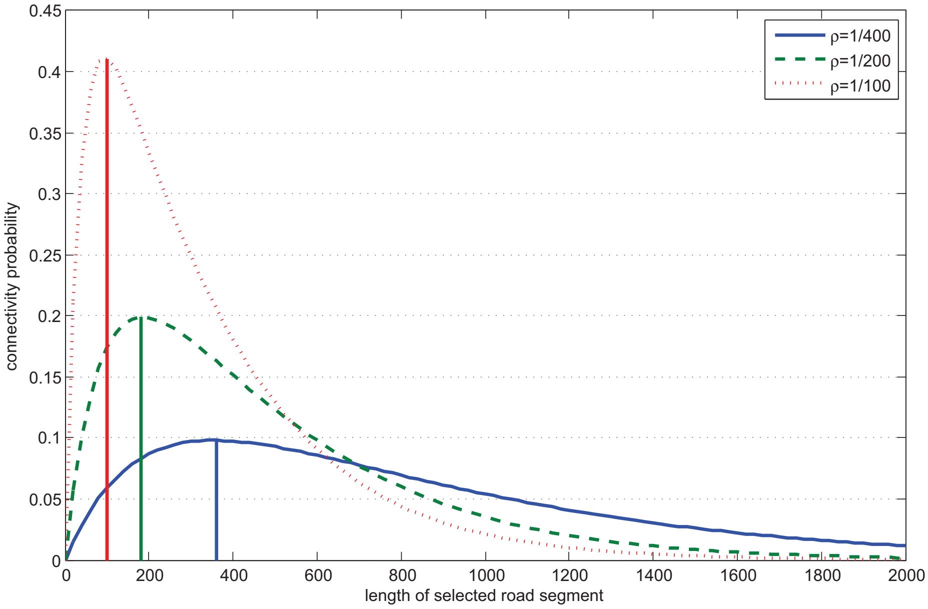


Figure 4.12: The impact of parameter ρ on connectivity probability p_6.

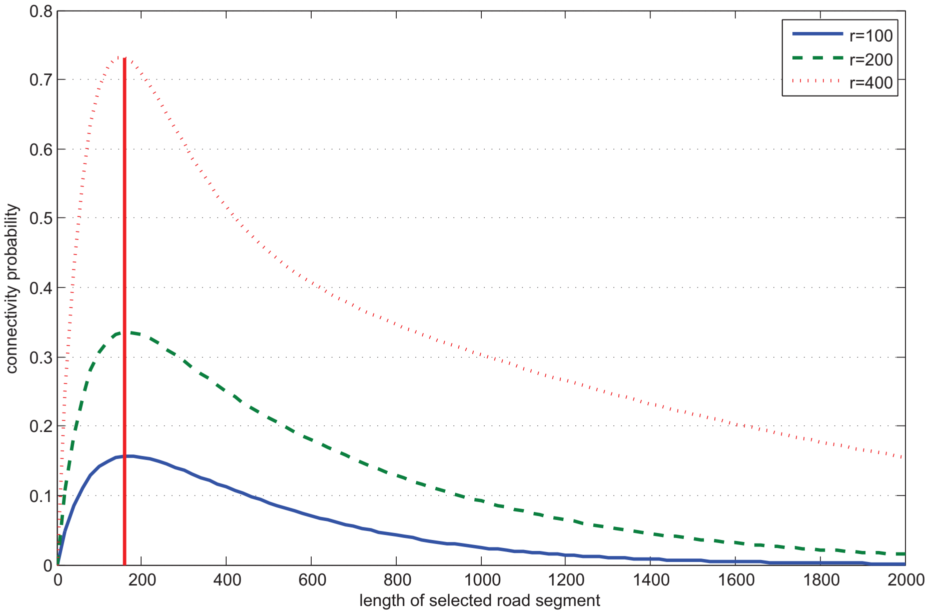


Figure 4.13: The impact of parameter r on connectivity probability p_6.

4.4 Connectivity-oriented Data Dissemination

4.4.1 Problem Formulation

In this work, we investigate the problem how to maximize the network connectivity in vehicular Internet access systems, where vehicles adopt both V2V and V2I communications to access the Internet. Given a road segment with vehicles and road-side APs, we assume that the vehicles' arrival rate follows the Poisson distribution and the deployment of APs follows the normal distribution. Each vehicle has both V2V and V2I capability and bandwidth demand to the Internet; each Wi-Fi AP has the bandwidth constraint to the Internet. In addition, according to the network conditions, we can obtain the connectivity probability between each vehicle and each AP. We try to find the optimal connection matchings between vehicles and APs such that the whole network connectivity probability is maximized, where the whole network connectivity probability is defined as the sum of connectivity probability of each connection matching. There are two constraints in finding the optimal solution. One is that for each vehicle, there is only one AP to be selected as Internet gateway. The other one is for each AP, the total bandwidth requirements can not exceed the bandwidth constraint. The formal mathematic formulation is listed as follows.

In this work, there some given conditions. For example, the length of the considered road segment and the number of vehicles and road-side APs are known in the beginning. There some some requirements and constraints on vehicles and APs. According to the derived formulations in previous section, we can obtain the connectivity probability between each vehicle and AP pair. We summarize the given conditions in the following point form.

given:

1) a road segment with length of s

2) n vehicles V and m road-side APs U in the considered area

3) $D = \{d_i,\ i \in (1, n)\}$: bandwidth demand of mobile vehicles

4) $B = \{b_i,\ i \in (1,m)\}$: bandwidth constraint of APs

5) $CP = \{cp_{ij},\ i \in (1,n)\ and\ j \in (1,m)\}$: connectivity probability between vehicle i and AP j

To simplify the analysis, we introduce some assumptions in the considered scenarios. In our work, each vehicles has both vehicle-to-vehicle and vehicle-to-infrastructure communication capability. The arrival rate of vehicles follows the Poisson distribution. The road-side APs are normally distributed along the road segment. The vehicles adopt multi-hop single-path routing protocol to communication with other parties in the networks. Our objective is to find the optimal connection matchings between vehicles and road-side APs to maximize the network connectivity probability with some constraints.

objective:

Find a connection matchings between vehicles and road-side APs such that the network connectivity probability is maximized,

$max \sum\limits_{(i,j)} cp(i,j) \cdot a(i,j),$

where $A = \{a_{ij},\ i \in (1,n)\ and\ j \in (1,m)\}$, if $a_{ij} = 1$, vehicle i connects to the Internet through AP j

subject to:

1) $\sum\limits_{j=0}^{m} a_{ij} = 1$: path constraint of single-path routing

2) $\sum\limits_{i=0}^{n} a_{ij} * d_i \leq b_i$: bandwidth constraint of APs

Obviously, this problem can be regarded as an application of Constrained Weighted Bipartite Matching problem, which is already proven as NP-hard.

4.4.2 Design of CoDA

In this section, we design a heuristic algorithm to solve the proposed connectivity maximization problem. As mentioned before, the problem can be regarded as a constrained bipartite matching problem. The input is a complete bipartite graph, which has weighted edges between each vertex pair and the same number of vertices on both sides. At the initial stage, vertices on both sides are connected through disjoint directed paths. The optimization objective is to find a perfect matching with the maximum or minimum weight.

By negating the weights of the edges we can state the problem as the following integer linear programming formulation for the minimum weight perfect bipartite matching problem:

$$\begin{aligned} min \sum_{(i,j)} cp'(i,j) \cdot a(i,j) \\ \sum_{j=0}^{m} a_{ij} = 1 \\ \sum_{i=0}^{n} a_{ij} * d_i \le b_i \\ a_{ij} \in \{0,1\} \end{aligned} \tag{4.13}$$

According to linear programming, we give a simple algorithm, as shown in Algorithm 3 to solve the above problem. In the beginning of the execution, the initialization for potential of U and V are conducted. According to the default values, the tight edges are selected if $cp(v,u) = y(v) + y(u)$. After that, the initial graph matching for $G' = (U \bigcup V, E')$ can be determined by selecting the max cardinality matching.

From E' and M, we can obtain an extended directed graph $G'' = (U \bigcup V, E'')$, from which the possible vertexes will be selected as candidates for the next iteration according to the connectivity probability and the bandwidth constraint on one single AP. It is proven that Algorithm 3 can terminates in $O((|U| + |V|)^2)$ iterations.

4.5 Performance Evaluation

To evaluate the performance of our proposed algorithm, we carry out extensive simulations in vehicular Internet access systems, of which the system model is similar to the system model in previous section. For example, the length of the considered road segment is

Algorithm 3: CoDA for maximization of connectivity probability

```
$\forall u \in U$, $y(u) \leftarrow 0$
$\forall v \in V$, $y(v) \leftarrow min\ \{cp(v,u) | b \in U\}$
$E' \leftarrow$ set of tight edges
$A \leftarrow$ max cardinality matching for graph $G' = (U \bigcup V, E')$
while $A$ is not a perfect matching do
    let $E'' \leftarrow \{$ e directed from V to U $| e \in E', e \notin M \}$
    $E'' \leftarrow \{$ e directed from U to V $| e \in E', e \in M \}$
    let $G'' = (U \bigcup V, E'')$ be a directed graph
    let $C \leftarrow \{$ v $| \exists v' \in V, v'v \in E'' \}$
    $\epsilon = min\ (cp'(v,u) - y(v) - y(u))$, where $v \in (U \bigcup V \bigcap C), u \in (U \bigcup V \backslash C)$,
        $\sum_{i=1,n} m_{v_i,u} * d_i \leq b_i, v_i \in V$
    $\forall v \in G''$, $y(v) = y(v) + \epsilon$
    $\forall u \in G''$, $y(u) = y(u) - \epsilon$
    update $E', A$
end
return A
```

2000 meters and the density of mobile vehicles is 1/200 (vehicle/meter). The transmission ranges of V2V and V2I communications are set to 100 meters and 300 meters, respectively.

We compare CoDA with other two heuristic algorithms, namely closest-BS forwarding and random forwarding. Different from CoDA, closest-BS forwarding algorithm only chooses the nearest base station as the candidate and random forwarding algorithm selects the candidate from its neighboring vehicles randomly.

In the first series of simulations, the communication traffic is set to quite light, such that all the bases stations are always available for vehicles' access. Thus, the delay of packet delivery is mainly attributed to the delay caused by link disconnection. As shown in Figure 4.14, the cumulative distribution function of delay of packet delivery is illustrated on three algorithm. Within 120s, CoDA can achieve a packet delivery ratio of 80%, which is much higher than the other two algorithms. When the delay is increased, the performance gain falls down. Within 160s, CoDA and closest-BS forwarding obtain the similar packet delivery ratio, which is still higher than random forwarding algorithm.

In the second series of simulations, the communication traffic follows normal distribution

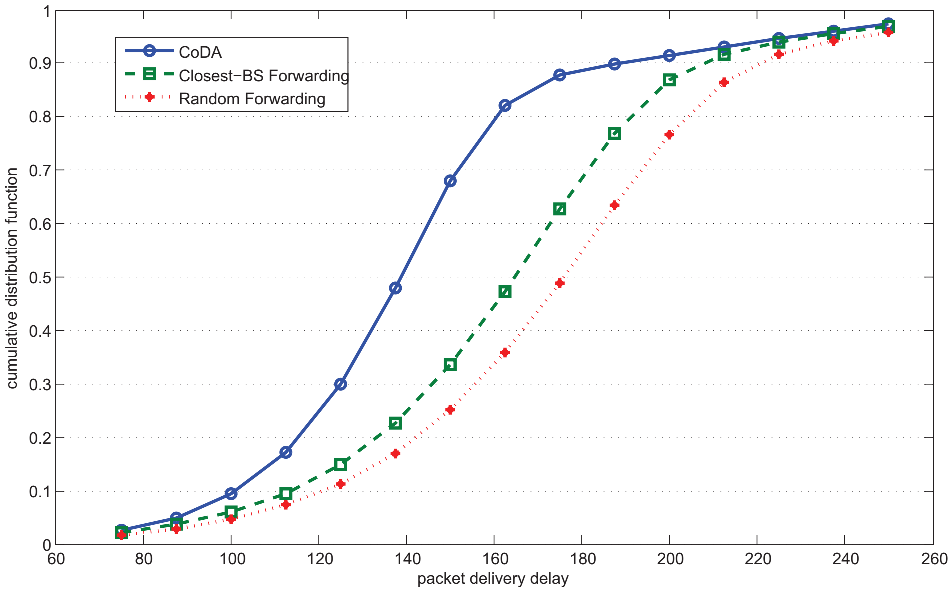


Figure 4.14: The cumulative distribution function of packet delivery on three algorithms.

with the mean bandwidth requirement d among the vehicles and the bandwidth support at APs are equal to b. Thus, the delay of packet delivery is jointly affected by the disconnection delay and waiting delay caused by bandwidth bottleneck at APs.The cumulative distribution function of delay of packet delivery on different ratios of (d/b) is shown in Figure 4.15. Obviously, the packet delay delay is prolonged with the increase of the ratios (d/b). It is worth noting that the CDF value can not reach to 1, because the packet transmission in this simulation settings may failed and the packet may lost in the forwarding path due the timeout of transmission sessions.

4.6 Summary

In this paper, we investigate the maximum network connectivity problem in vehicular Internet access networks, in which vehicles adopt both vehicle-to-infrastructure and vehicle-to-vehicle communications. We first analyze the characteristics of network connectivity in vehicular Internet access networks and derive a closed form to represent the connectivity

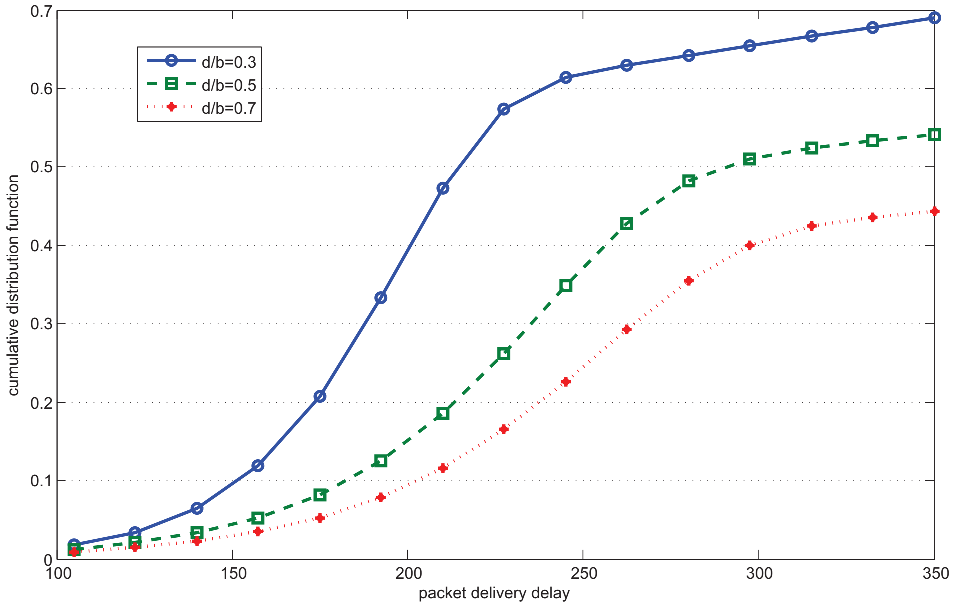


Figure 4.15: the cumulative distribution function of delay of packet delivery on the bandwidth ratio of (d/b)

probability under specific road-side units (RSU) deployment model and vehicle arrival rate. Then, we develop a connectivity-oriented data dissemination (CoDA) algorithm to maximize the network connectivity in vehicular Internet access networks based on the probabilistic analysis. Finally, we validate the CoDA algorithm through extensive simulations and the results show that the CoDA algorithm can achieve the maximum connectivity probability in vehicular communications in different network scenarios.

Chapter 5

An Accurate and Cooperative Vehicle Localization Service for Vehicular Internet Access in Urban Area

In this chapter, we study a QoE-related application in vehicular Internet access system. We proposed an accurate and cooperative vehicle localization service for vehicular Internet access in urban area. This chapter is organized as follows. In Section 5.1, we give the overview of the work and show some related work in Section 5.2. Section 5.3 analyzes the results from field testing and illustrates the opportunities to improve the localization accuracy through neighboring GPS data. Section 5.4 presents the design of the *Networking-GPS* and Section 5.5 reports the evaluation reports. Finally, Section 5.6 concludes the whole chapter.

5.1 Introduction

As an important part of Intelligent transportation system (ITS), vehicular applications have been proposed over the several decades. Most of the conventional vehicular applications operate in stand-alone mode and mainly focus on the in-vehicle status maintenance and safety improvement. In the recent years, with the popularity of mobile and smart devices, an increasing number of vehicular applications have emerged by enabling the communication capability and invoking all kinds of existing services. As a fundamental service, online real-time vehicle localization service is critical to the normal execution of vehicular applications. In addition to obtaining benefits to the administration, the accurate localization system can also provide better user experience for car drivers and passengers. For example, vehicle-related location based services (LBS) need to acquire the accurate location information of vehicles to provide the better services to users.

A real-time vehicle localization system in urban area is a challenging issue due to several rigorous constraints and requirements. First, the localization accuracy should be satisfied by the users' applications. For example, the navigation system in urban area requires the precision of the lane level in order to differentiate the different directions of lanes. GPS is the most suitable solution for metropolitan-scale localization system in terms of the feasibility and robustness. However, GPS devices suffer the inaccuracy from the severe multi-path effect in urban area. Second, most of the positioning queries are attached with a time constraint, within which the requestor need to obtain the results. Otherwise, the answer may be useless for vehicles due to the high speed of moving vehicles. Third, the system should be scalable with the number of vehicles and compatible with all kinds of GPS devices.

Aiming to significantly improve the accuracy localization information from GPS measurements using commodity devices, we propose an accurate and cooperative localization algorithm, called *Networking-GPS*, for public transportation system in urban area. *Networking-GPS* is based on the observation that the relative distances between neighboring vehicles with similar GPS signal data in terms of satellite sets and signal strength are much more stable. Running on each vehicle, *Networking-GPS* constructs the *atom triangle* with neighboring vehicles according to similarity of GPS data. Then through multi-hop multilaretation, road-side anchor nodes which broadcast their GPS observation data and physical locations periodically are included to compute the absolute locations of the considered *atom triangle*. We have designed *Networking-GPS* using the commodity GPS devices without any modification of built-in GPS algorithms. The contributions of the work are summarized as follows.

1. We have carried out comprehensive experiments on several commercial off-the-shelf GPS receivers to investigate the relationship between the multi-path effect and the GPS measurements. We find that there exist the correlations in neighboring GPS measurements under the similar impact of multi-path effect. Based on the limited outputs from the commodity GPS receivers, we identified that the satellite sets and signal strength of GPS signals are critical to the relative distances between neighboring nodes. Based on the observation, we can identify the *atom redundant rigid graph* among neighboring vehicles.
2. To improve the accuracy of noisy GPS measurements against the multi-path effect, we develop a rigid expansion algorithm by exploiting the signal similarity and ranging information from GPS measurements.
3. Based on the above two steps, we propose an accurate and cooperative localization algorithm, called *Networking-GPS*, for public transportation system in urban area. Through extensive simulations based on the real road network and trace data of vehicle mobility, *Networking-GPS* can meet the real-time constraint associated with each vehicle and improve the normalized accuracy of the entire system.

The remainder of the chapter is organized as follows. We show some related work in Section 5.2. Section 5.3 analyzes the results from field testing and illustrates the opportunities to improve the localization accuracy through neighboring GPS data. Section 5.4 presents the design of the *Networking-GPS* and Section 5.5 reports the evaluation reports.

Finally, Section 5.6 concludes the whole chapter.

5.2 Related Work

In this section, we introduce some related works in this area, including Satellite-based positioning technology and other metropolitan-scale positioning technologies.

5.2.1 Satellite-based Positioning Technologies

Satellite-based positioning system is a system consisting of a set of satellites, which can provide necessary positioning signal information for the global area. To obtain the location information on the ground, small electronic devices can receive the signals transmitted by several satellites and calculate locations in 3D space based on the information decoded from satellite signals. GPS-based positioning can be regarded as the first global positioning technology in the world, which was initially developed by the U.S. Department of Defense in 1960s. In GPS system, each satellite continuously transmits signals including the following information:

- the transmission time of the signal
- the orbital information of the satellite generating the signal
- the global information of all the satellites in the satellite systems

Once receiving signals from satellites, GPS receivers can calculate distances to the corresponding satellite based on the time difference between signal transmission and reception. According to the embedded GPS algorithms in receivers, position information can be computed by the locations of satellites and theirs distances to receivers. Then, the position information along with other derived information, such as signal strength, velocity and acceleration, are serialized to other devices via communication interfaces.

In standard positioning procedure, at least four GPS satellites are required and the localization accuracy will be improved if more correlated information is involved in the positioning calculation. For example, the elevation value is available in some application scenarios and the localization calculation can utilize the known condition as input to the GPS algorithms. In addition, there are several variations of GPS systems, including Assisted GPS (A-GPS) and Differential GPS (DGPS), to provide high positioning accuracy and shorten the localization time.

Mitigation of Multi-path Effect

It is very common to adopt GPS receivers as the localization solution in vehicular environments. However in the urban area, the positioning information may not be accurate since the GPS signals may be severely distorted by the environmental factors. Among them, the multi-path effect plays a dominant role in accuracy degradation of localization system. To mitigate the multi-path effect in localization system, a number of approaches have been proposed by decreasing the effects of sensing and ranging biases caused by multi-path on the localization computation procedure. We summarize the improvement solutions to combat the multi-path effect into two categories.

One category of approaches is based on the consistency analysis and peak separation. It discards GPS signals with great inconsistency of pseudo-range estimates and conduct the peak separation procedure to identify the GPS signal. However, it assumes that the multi-path model in the approach is quite simple, which can to apply to the real urban area. Second one is based on the statistical analysis. Based on the real GPS data, it first develops a statistical model to capture the characteristics of pseudo-range error distributions. Then, the localization information can be determined by feeding the real-time GPS data based on the statistical model. However, establishing such statistical model is a time-consuming task and the model has to be adjusted if the environments change.

5.2.2 Vehicular Cooperative Localization System

In the traditional vehicular environments without communications between vehicles and infrastructure, the localization system has to work in a stand-alone way. Once the inter-vehicle communication is enabled, the localization information can be exchanged among neighboring vehicles, even reported to remote servers via Internet gateways. By fusing the positioning information from different vehicles, the localization accuracy can be improved in a cooperative way. The cooperative localization is initially designated for wireless sensor networks, in which sensor nodes are lack of global localization information. Sensor nodes have to be localized by local information exchanged among neighboring nodes in communication range.

In [Ben05], a cooperative vehicular localization approach was proposed to the vehicles without GPS positioning information. The vehicles can be equipped with no GPS modules or can not receive the required GPS signals to fix the location. In the method, the positioning information is spread over the network via the inter-vehicle communication. Vehicles can estimate the relative distances between the other vehicles in the communication range by leveraging the characteristics of radio frequency. The vehicle without GPS positioning information can determine the absolute location once receiving the relative location information from at least three neighboring vehicles. However, this work does not consider the multi-path effect on the GPS signals, which is the domain factor in the degradation of positioning accuracy in urban vehicular environments.

To mitigate the degradation of the multi-path effect, [DB10] proposed a new localization approach, called InterVehicle-Communication-Assisted Localization (IVCAL), to improve the accuracy and reliability in vehicle localization. In the approach, a Kalman filter and an inter-vehicular networking system are jointly considered to detect the multi-path effect and improve the localization performance. By sharing the inaccurate location information, the relative distances between vehicles can be used to achieve robust and accurate localization.

5.3 Motivation and Approach

In this section, we will demonstrate and analyzes the experiment results of GPS measurements using different commodity devices. Then, we illustrate the opportunities to improve the localization accuracy through neighboring GPS data.

5.3.1 Preliminary Experiments

In practical GPS localization system, GPS receivers calculate the positions after receiving enough valid GPS signal. If the results are ready, GPS chips will write the values to its interfaces, such as serial or USB ports, in the form of ASCII codes. NMEA standard, developed by National Marine Electronics Association to present the information interface between different electronic devices, is a common format for GPS result output. The data generated by the GPS receiver includes some information related to localization, such as latitude, longitude, altitude, time and speed. For example, GPGGA is short for Global Positioning System Fix Data and GPRMC stands for Recommended minimum specific GPS/Transit data.

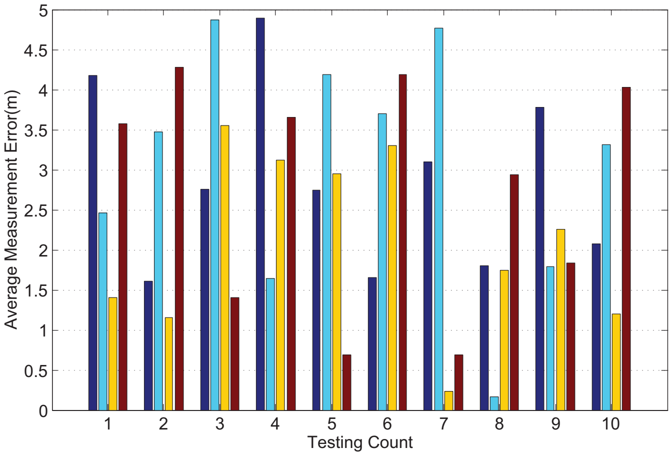
To examine the multi-path effect on GPS performance, we used 4 GPS receivers and deployed them in different scenarios. Among the 4 GPS receivers, two are built in SiRFstar-III GPS chips and the other two are equipped with MTK GPS chips. After analyzing the GPS output, the SiRFstar-III based GPS receivers follows the standard NMEA format. However, the MTK based GPS receivers includes some proprietary sentences. To ease of system design, we first reconfigured the MTK based GPS receivers and converted the output format from the proprietary messages to standard NMEA sentences. Then, we fixed the GPS receivers on four corners of a square board. The relative distances between each GPS receivers can be ignored to compare with its dimensions. We set the sampling rate of GPS receivers to 1 second, which means every one second, the devices report its calculation results to the interfaces.

Both stationary and moving scenarios are considered in our experiments. In the stationary scenario, we placed the board in two different locations. First, we carried out the experiment in an open-sky environment, and then we placed the board in an outdoor location near our office building. We collected all the GPS raw results from serial ports for further analysis. In the moving scenario, we put the board on the top of a taxi and obtained the GPS output along a round trip in downtown.

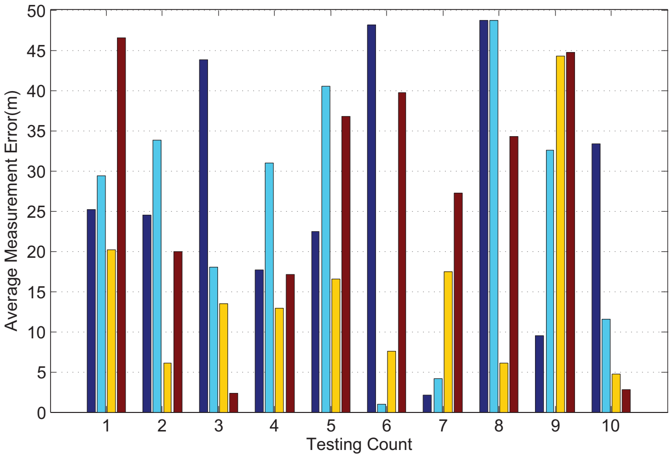
The final results are shown in Figure 5.1, in which the average GPS measurement errors are plotted against the experiment times. From the three figures, we can find that the GPS localization performance is relatively stable in open-sky environment. However, the multi-path effect imposes severe degradation in localization performance in near building environment and moving scenario in urban area. The localization errors can be up to hundreds of meters.

5.3.2 Similarity Between GPS Signals

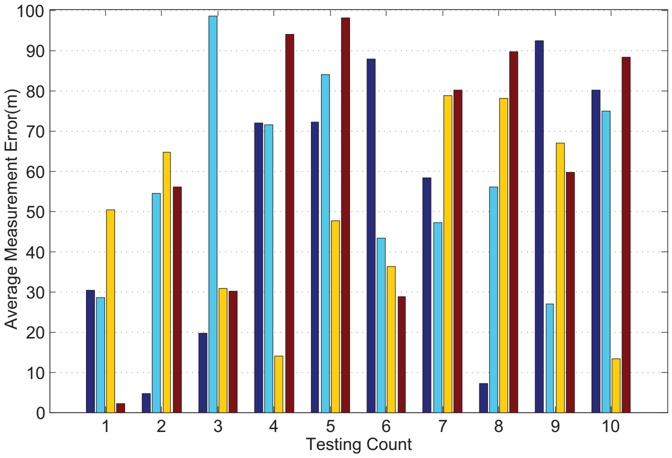
In the raw GPS result output, there is one type of sentences called GPGSV, which means GNSS satellite in view. It represents data related to the satellites which the GPS device could detect based on its configurations and decode the information from. Based on the standard, each GPGSV sentence only includes data from up to 4 satellites. Thus, we need at least 3 different GPGSV sentence to collect full data from in-view satellites. The message format of GPGSV is shown as $\$GPGSV,m,n,vv,i_1, e_1, a_1, s_1,,{*}CS{<}CR{>}{<}LF{>}$, in which T stands for total number of messages and M is message number (1 to 3). In every GPGSV message, there exist up to 4 field group, as i_1, e_1, a_1, s_1, which stands for PRN (PseudoRandom Noise) code, Elevation in degrees, Azimuth in degrees and SNR from the respective satellites.



(a) Stationary scenario 1



(b) Stationary scenario 2



(c) Moving scenario

Figure 5.1: Localization variation with multi-path effect in two scenarios.

The SNR filed in a GPGSV sentence is also called as signal strength. According to the NMEA standard, SNR value is a useful metric to measure the signal strength of raw GPS signals and the variation of SNR value is from 0 to 99 with units of dB. Noting that value of 0 is a special case when satellites are detected in the view but can not be tracked. There are much efforts to be made to reveal the correlation between SNR value of GPS signals and multi-path effect [BLA06, BL07]. However, most of the work still tried to design the GPS algorithm to improve the performance of standalone GPS receivers.

In this work, we will utilize the GPS raw data from nearby GPS receivers, especially using the SNR data and localization information, to improve the overall localization performance of multiple GPS receivers. Through the GPS raw data, we can obtain the corresponding SNR values of GPS signals from different satellites. To capture the spatio-temporal similarity in SNR variation from GPS signals, we develop a detailed expression of the correlation in form of a correlation matrix, $S(m, n)$. Each individual column corresponds to a SNR observation from GPS signal in a particular time, including the SNR values from m GPS satellites sensed by receivers. Each row is an n-element correlation vector, which means the signal variation from one particular GPS satellite in time domain. We captured the dominant variation patterns by adopting Singular Value Decomposition (SVD) of the correlation matrix. In practice, SVD is usually beneficial to achieve the following three objective. First, it can be used to transform correlated variables to a set of uncorrelated variables. Second, it can be adopted to identify the dimensions with most variation. Third, SVD can be also exploited to reduce the raw data set in fewer dimensions.

In our analysis, the correlation matrix, $S(m, n)$ can be transformed by SVD to an expression of a product of three matrices, written as $S = U \cdot A \cdot V^T$, where $U^T \cdot U = I$ and $V \cdot V^T = I$. The singular values of $A = \{a_1, a_2, ..., a_r\}$ and the faction of power in form of each eigen vector is calculated by

$$w_i = \sum_{i=1}^{k} a_i^2 / \sum_{i=1}^{rank(S)} a_i^2 \tag{5.1}$$

The eigen vectors of correlation matrix S is used to measure the similarity between SNR values from different GPS receivers quantitatively. For example, GPS raw data at two different locations, with respective eigen vectors as $X = \{x_1, x_2, ..., x_{r_x}\}$ and $Y = \{y_1, y_2, ..., y_{r_y}\}$. The signal similarity can be derived in the form of their eigen vectors, as follows

$$Sim(X,Y) = \sum_{i=1}^{rank(X)} \sum_{j=1}^{rank(Y)} w_{xi} \cdot w_{yi} \cdot |x_i \cdot y_i| \tag{5.2}$$

$Sim(X,Y)$ is a quantitative measurement metric which describe the closeness of two GPS observation data in spatio-temporal dimension. The similarity value is between $0 \leqslant Sim(X,Y) \leqslant 1$.

According to Equation 5.2, we re-visited the experiment results and checked the relationship between the similarity of SNR values and the relative distance, as shown in Figure 5.5.

We can easily find that the relative distance has a strong correlation with the signal similarity. In Figure 5.2 and Figure 5.3, relative distance between different static GPS receivers still varied due to the severe multi-path effect near the tall building. However, when the signal similarity increases, the relative distance between the considered GPS receivers decreases correspondingly.

In Figure 5.4, the cumulative distribution function (CDF) between GPS signal similarity and relative distance are shown through pairwise comparison from our experiment results. In calculation, we set the accuracy threshold for relative distance to $0.5m$ and we claim that the relative distance is zero when the calculated relative distance is not greater than $0.5m$.

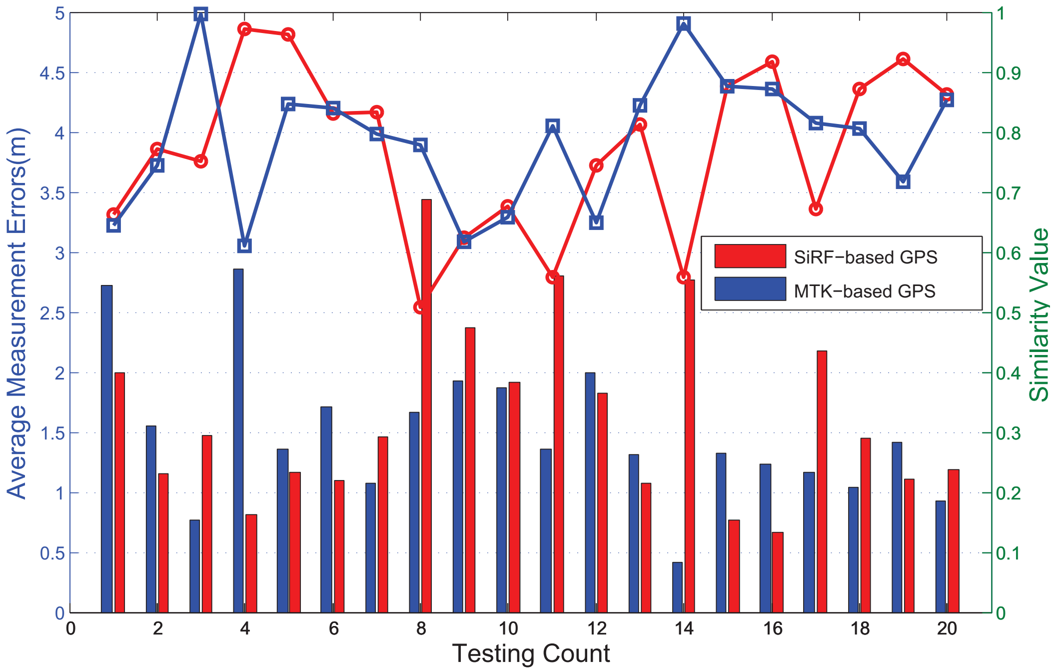


Figure 5.2: Average measurement error in stationary scenario 2

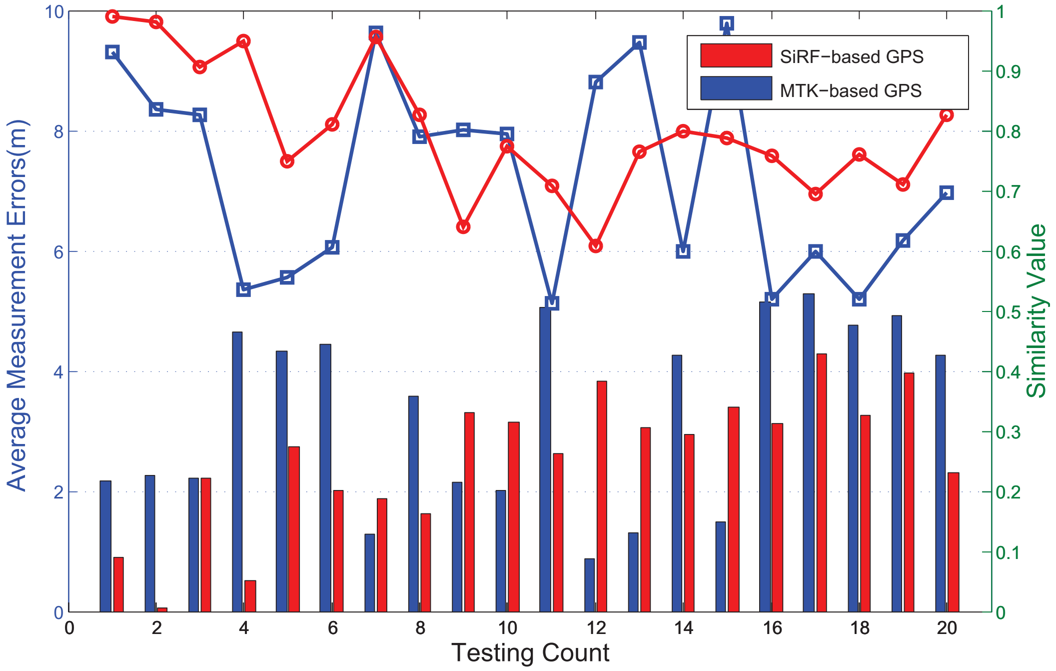


Figure 5.3: Average measurement error in moving scenario

We can easily find that in static scenario, when the similarity reaches to 0.9, the relative distance is almost equal to zero, with the possibility of 75%. From this observation, this method provide an accurate way to achieve local relative localization against the multi-path effect based on GPS raw data.

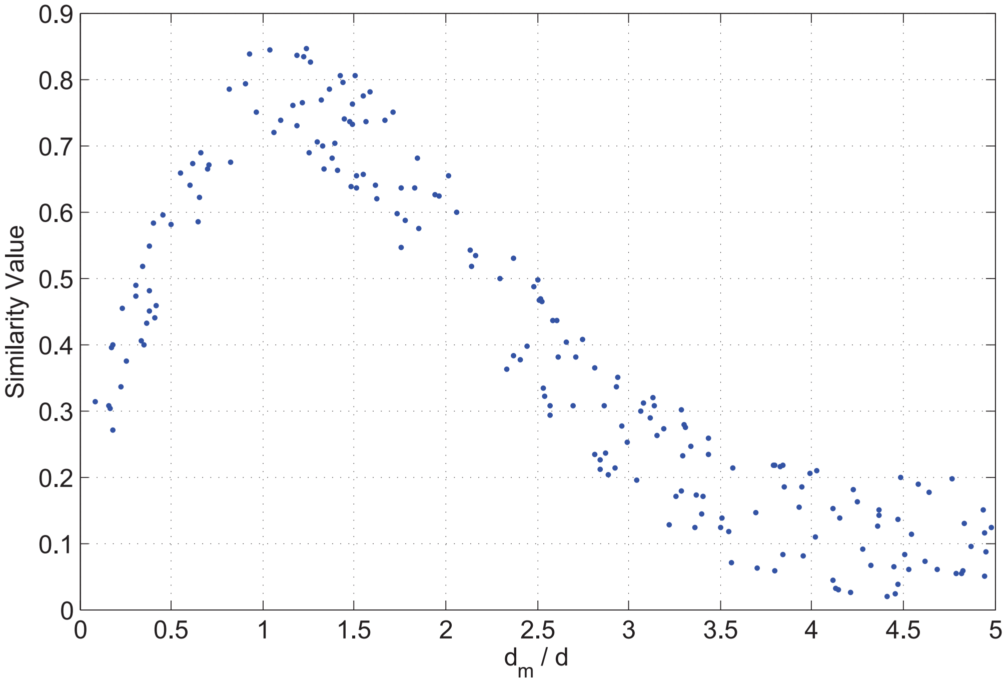


Figure 5.4: The correlation between distance variation with GPS signal similarity.

5.4 Design of *Networking-GPS*

In this section, we present the cooperative vehicle localization algorithm, named Networking-GPS, and analyze the impact of some key parameters on the localization accuracy of the proposed $Networking - GPS$ algorithm.

5.4.1 Overview

The Networking-GPS algorithm can run in a central server or can be executed distributedly in each vehicle. We will describe the centralized version for Networking-GPS algorithm in this work, and then we will investigate the distributed version in our future work.

Considering in a vehicular network, vehicles with Internet connection are equipped with commodity GPS receivers for localization or navigation service. GPS receivers calculate the positioning information and forward the raw GPS data to a central server. The raw GPS data are encapsulated in NMEA format. In the algorithm, we claim that the edge between two GPS receivers is a normal edge if a similarity index between them calculated by Equation 5.2 is greater than a threshold δ. Before describing the detailed algorithm design, we provide some preliminary knowledge about rigidity and redundantly rigidity.

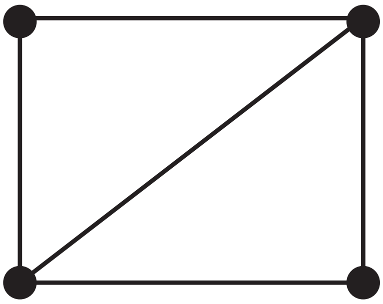
Definition 5. *A rigid graph is an embedding of a graph in a Euclidean space which is structurally rigid. A graph is redundantly rigid if it is rigid after one (any one) of its edges is removed.*

In our Networking-GPS algorithm, based on the similarity correlation, we first construct the atomic redundantly rigid graphs and then expand the local graphs to a global graph by including some anchor points. After that, the local positioning information can be transformed to a global localization information. The detailed algorithm is explicated as follows.

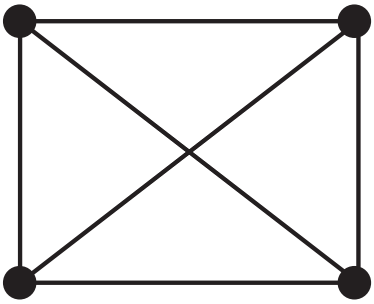
5.4.2 Construction of Atomic Redundantly Rigid Graphs

Since GPS signals have the significant variation against time and space diversity, it is very hard to form a global rigid graph for localization. To mitigate the impact of the time and space diversity, we have to first construct some atomic redundantly rigid graphs based on the nearby GPS signals. We take four-vertex redundantly rigid graphs as the atomic graphs, as shown in Figure 5.5b and 5.5c. The procedure is shown in Algorithm 4.

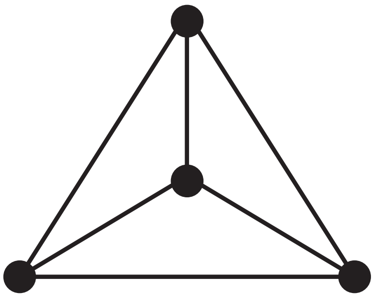
In the beginning, the similarity between each pair of GPS receivers in the considered grid is calculated by the Equation 5.2. We selected the top four receivers according to the sum of similarity values between them. If the similarity of the each pair among the selected



(a) Rigidity but not redundant rigidity



(b) Redundant rigidity



(c) redundant rigidity

Figure 5.5: Rigidity and redundant rigidity

Algorithm 4: Construction of atomic redundantly rigid graphs

```
Input  : A graph $G(V, E)$, the similarity threshold $\delta$ for defining normal links;
Output: Atomic redundantly rigid graph $G_r(V_r, E_r)$;
for $\forall v_i, v_j \in V$ do
    $s_{i,j} = Sim(v_i, v_j)$,
end
for $\forall v_i, v_j, v_k, v_l \in V$ do
    $t_{i,j,k,l} = \sum s_{p,q}$, where $p, q \in \{i, j, k, l\}$,
end
select $(a, b, c, d) = argmax\{t_{i,j,k,l}\}$, where $v_i, v_j, v_k, v_l \in V$ if $\forall s_{i,j} \geq \delta$, where
$i, j \in (a, b, c, d)$ then
    return $G_r$, where $V_r = a, b, c, d$ and $E_r = \{e_i, e_j | i, j \in \{a, b, c, d\}\}$
end
else
    return output $G_r = \emptyset$
end
```

four GPS receivers is greater than the threshold δ, the atomic redundantly rigid graph is found in the considered grid.

There are two important implications. First, due to the setting of similarity threshold δ, Algorithm 4 may find no satisfied results for the considered grid. In the performance evaluation part, we will discuss the impact of threshold δ to the localization accuracy. The other one is there may exist more than one atomic redundantly rigid graph in one considered grid and Algorithm 4 only pick one from the existing redundantly rigid graphs. It is our important assumption to conduct the following rigidity expansion operations.

5.4.3 Expansion Algorithm of Redundantly Rigidity

After forming atomic redundantly rigid graphs, we only obtain accurate relative positioning information between subset of the GPS receivers. To map the relative information to global information, we need to expand the atomic redundantly rigid graphs to a larger area to cover some anchor nodes, such as fixed nodes or moving nodes with high localization accuracy.

Before presenting the detailed expansion algorithm, we introduce the following theorem which shows the relationship between redundantly rigid graphs. From another perspective, it also provide an efficient method to do rigidity expansion.

Theorem

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two redundantly rigid graphs with $|V_1 \bigcap V_2| \geq 2$. Then,$|G1 \cup G2|$ is redundantly rigid.

Our expansion algorithm is a vertex-based solution which determines whether one vertex can be merged into a redundantly rigid graph. If the merged graph is still redundantly rigid, we select the redundantly rigid graph with the maximum cardinality and merge the vertex into the graph. If there exists no a merged graph that can merge the vertex to a larger redundantly rigid graph, we define a weight function $RW(\,\cdot\,)$, as shown in Equation 5.3 to characterize the similarity between the vertex and other redundantly rigid graphs. According to the RW values, we can merge those vertexes to weighted redundantly rigid graphs.

$$\begin{gathered} RW(v, G, \delta) = \sum_{i=1}^{3} \left(Sim(v, v_i^G) - \delta\right)^3, \\ \textbf{where } \forall Sim(v, v_i^G)|_{i=1,2,3} > Sim(v, v_j^G)|_{v_j \in V^G - \{v_1, v_2, v_3\}} \end{gathered} \tag{5.3}$$

The entire algorithm is shown in Algorithm 5. The steps 1 to 4 is responsible for merging vertexes into a larger existing redundantly rigid graphs. The merging process between redundantly rigid graphs are shown in steps from 5 to 11. The weighted redundantly rigid graphs is constructed from steps 12 to 15.

5.5 Performance Evaluation

We evaluate the performance of the Networking-GPS approach using real GPS samples. We first collected all the GPS raw data from receivers in the experiments, and then established a simulation environment for the offline evaluation.

Algorithm 5: Expansion algorithm of redundantly rigidity

```
Input  : A graph $G(V, E)$, a set of $n$ redundant rigid components
         $G^r = \{G_1, G_2, ..., G_n\}$, the similarity threshold $\delta$ for defining normal links;
Output: The maximum weighted redundantly rigid graph $G^R$;
for $\forall v \in G - \bigcup_{i=1}^{n} G_i$ do
    select $G_m = argmax\{|G_i + v|\}$ and $G_i + v$ is redundantly rigid $G_m = G_m + v$
end
$G^{rr} = \emptyset$ for $\forall G_i, G_j \in G^r$ do
    if $|V_i \bigcap V_j| \geqslant 2$ then
        $G_i \bigcup G_j \rightarrow G^{rr}$ $G^r = G^r - \{G_i, G_j\}$
    end
end
for $\forall v \in G - \bigcup_{i=1}^{n} G_i^{rr}$ do
    select $m = argmax\{RW(v, G_i^{rr}, \delta)\}$ $G_m^{rr} = G_m^{rr} + v$
end
return output $G^R = argmax|G_i^{rr}|$
```

5.5.1 Raw Data Collection

We still used the two kinds of GPS receivers as mentioned in the section 3.5.1 with two different GPS chips. To facilitate the process of data collection in vehicles, we built a small GPS collection box by packaging four functional elements, namely GPS receiver, portal power bank, data process unit and data storage unit. The data process unit is responsible for receiving the GPS data and converting the raw data to a human-friendly format.

In our experiment, we adopted a small portable router TP-Link TL-WR703N as the data process unit. TL-WR703N is a low cost 3G travel router and is very popular in developer community due to its open-source from hardware to software. We flashed a compatible OpenWrt firmware to the router, which is a opensource third-party router firmware system and compiled the necessary USB drivers and software package for the hardware. In our implementation, we utilized the TL-WR703N as the controller to receive and process the GPS raw data, and then forward to the data storage unit. The key components of a GPS collection box are shown in Figure 5.6.

We mounted our small GPS collection boxes on the top of vehicles and recorded the raw

Figure 5.6: Key components in a GPS collection box

GPS trace along the routes. In the end of the experiments, we exported all the data from the storage units. In addition, we obtained the ground true of the localization information, we collected the latitude and longitude values of ten waypoints along the route, as shown in Figure 5.7 and the waypoints are numbered from 1 to 10.

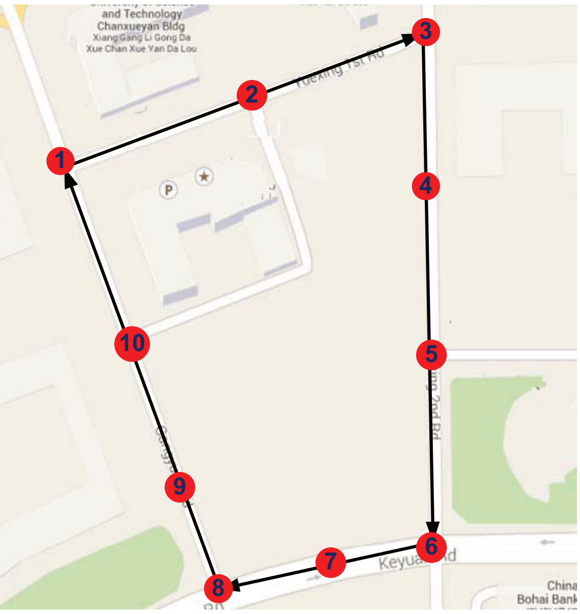


Figure 5.7: Route and waypoints

5.5.2 Simulation Results

We conducted offline simulations and processed GPS data traces from collection boxes. For ease of comparison, we only considered the GPS data collected around the ten waypoints in Figure 5.7. After applying our proposed Networking-GPS algorithm, the comparison of accuracy performance is shown in Figure 5.8. The blue bar represents the direct measurements from GPS receivers. And, the other two represent two different simulation results by choosing different threshold . We can easily find that the accuracy performance in direct measurements is degraded due to the multi-path effect and the variations are also great. For example, at waypoint 3, the average localization error is nearly 15m and the variation is about 20m. Networking-GPS provides the high accuracy and the localization information is much more stable than direct measurements.

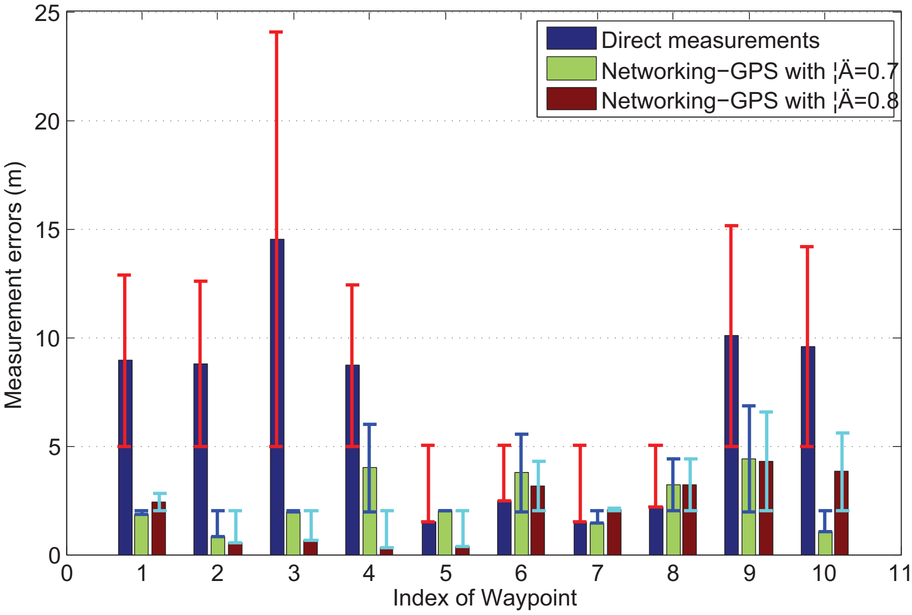


Figure 5.8: Accuracy comparison between direct measurements and Networking-GPS

In our Networking-GPS algorithm, we have a parameter, namely similarity threshold , which is involved in both Algorithm 1 and 2. To identify the impact of this parameter, we chose four waypoints into consideration, points 1, 3 ,6, 8. Around those waypoints, we executed Networking-GPS algorithms under different threshold δ, from 0.6 to 0.9 with step size of 0.3. The results are shown in Figure 5.9. We can find that the accuracy performance in different waypoints may present different trends. Therefore, the threshold δ should be dynamically adjusted in the entire route to achieve the best accuracy.

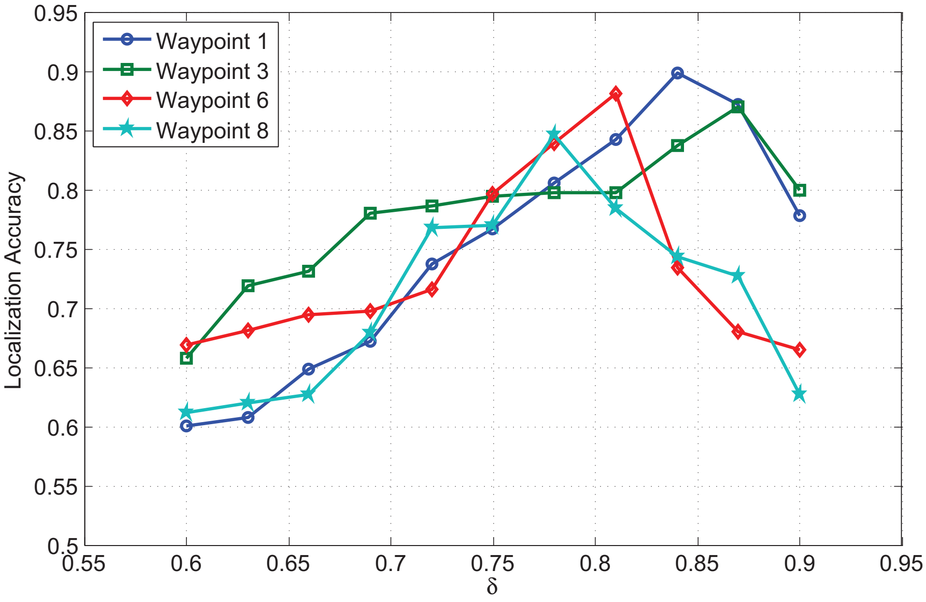


Figure 5.9: Impact of δ to accuracy performance

5.6 Summary

In this chapter, we develop a cooperative vehicle localization algorithm, named Networking-GPS, using commodity GPS in urban area. is developed based on some key observations, revealed by our trace data via data processing from our experiments. First, we show that multi-path effect imposes severe degradation in localization performance of GPS receivers. Then, we identified the correlation between the similarity of SNR values from different GPS satellites and relative distance among different GPS receivers. Based on the observation, we designed the Networking-GPS algorithm, which constructs atomic redundantly rigid graphs according to the signal similarity and expands atomic triangles to reach anchor points along the road side. Therefore, the local accuracy can be transited to the global accuracy through the redundantly rigid transformation. Our evaluation based on real GPS traces show that Networking-GPS can achieve higher accuracy against multi-path effect.

Chapter 6

Conclusion and Future Research

6.1 Conclusion

In this thesis, we identify the importance of Internet infrastructure on ever-increasing vehicular applications and clarify the fundamental Internet services matching the related services in the infrastructure. We conduct a systematical investigation and study on key issues in vehicular Internet access system. In each issue, we identified the problems which are important and not well addressed, and proposed corresponding solutions. We conclude these works as follows.

For the infrastructure planning issue, we identify the problem how to deploy different types of network components, such as base stations and relays, in a cost-efficient way to maximize the network performance. First, we propose a detailed analytical model for infrastructure-based vehicular Internet access in the presence of different network components. Then, through extensive simulations based on the proposed model, we obtain some guidelines on choosing and deploying different network components in vehicular Internet access networks. Finally, as a special case study of infrastructure planning, we investigate the incremental and cost-efficient network planning (ICNP) problem in vehicular Internet access networks with the objective of improving network capacity for an existing network under limited deployment cost. After defining a *cost efficiency* metric, namely incremental upgrade utility (IUU), we develop a heuristic algorithm, called Greedy Link Deployment algorithm (GLiD) with approximation ratio $(1-1/\sqrt{e})$ for deploying additional wireless links to the existing networks to achieve the maximum performance improvement. Through extensive simulations, the results show that our algorithm can effectively exploit the available resources and perform much better than other heuristic algorithms.

For the protocol adaption issue, we investigate the maximum network connectivity problem in vehicular Internet access networks, in which vehicles adopt both vehicle-to-infrastructure and vehicle-to-vehicle communications. We first analyze the characteristics of network connectivity in vehicular Internet access networks and derive a closed form to represent the connectivity probability under specific road-side units (RSU) deployment model and vehicle arrival rate. Then, we develop a connectivity-oriented data dissemination (CoDA) protocol to maximize the minimum network connectivity in vehicular Internet access systems based on the probabilistic analysis. Finally, we validate the CoDA protocol through extensive simulations and the results show that the CoDA protocol can achieve the maximum connectivity probability in vehicular communications in different network scenarios.

For the service enhancement issue, we take localization service, a fundamental service in vehicular applications, as an example. We study the vehicular localization problem in urban environments. First, we perform a series of experiments which shows that the vehicular mobility and multi-path effect impose severe degradation in localization performance of GPS receivers. We then identify the correlation between the similarity of signal-noise ratio (SNR) values of GPS signals from different GPS satellites and relative distance among those GPS receivers. Based on the observation, we design the Networking-GPS algorithm, which constructs atomic redundantly rigid graphs according to the signal similarity and expands atomic triangles to reach anchor points along the road side. Therefore, the local accuracy can be transited to the global accuracy through the redundantly rigid transformation. Our evaluation based on real GPS traces shows that Networking-GPS can achieve higher accuracy against multi-path effect in urban area.

In summary, vehicular Internet access is a promising networking application in future vehicular communication networks. We identified several important problems in different aspects of vehicular Internet access system, and proposed corresponding solutions. The evaluation results show that our approaches can increase the network performance in vehicular Internet access networks.

6.2 Future Research

We close this thesis by providing some suggestions for future research. Specifically, we believe that the following aspects are worth further investigations.

6.2.1 Opportunities Communication in Heterogenous Vehicular Access Networks

With the advance of wireless communication, more and more wireless access technologies are adopted in wide area networks. Cellular networks have evolved from 3G to 4G, even 5G networks and Wi-Fi networks achieve the high high penetration rate in people's daily life. It is expected that in the near future, vehicular networks have to be established in such heterogenous wireless networks. For each radio access technology, its distinguished characteristics can be exploited to improve the user experience during the in-vehicle time. To make full use of heterogenous wireless networks around vehicles, opportunities access is one helpful communication paradigm in such network scenario. There are two main issues to be cared to conduct opportunities communications in heterogenous vehicular networks. First, a uniform wireless radio module should be developed to establish wireless links to multiple radio networks. It is not realistic for vehicles to install every radio interface cards for each wireless access networks. The other one is that the dynamic of network topology and wireless spectrum should be jointly considered in the protocol design. Cross-layer information is helpful to achieve the optimization.

6.2.2 Big Data among Vehicular Access Networks

In conventional vehicles without V2V and V2I communication capability, the information generated is quite limited and hard to share to others. Vehicular access networks pave the way to Internet of vehicles and huge amount of data can be exchanged among vehicles and anywhere in the world. There are several categories of information collected in vehicular networks, including inner-vehicle information, environment information and inter-vehicle information. Inner-vehicle information can be used to telematics system to detect and avoid failures or accidents. Vehicles can be regarded as mobile sensor nodes scattered in the cities and collect information along the route. Inter-vehicle information mainly focus on the interactions between vehicles and vehicle to infrastructure, including communication behaviour, social relation and etc. Some efforts have been made to analyze the data from single source. It is more important that the correlation mining on data from different categories.

Bibliography

[3GP09] 3GPP. Evolved universal terrestrial radio access (e-utra) and evolved universal terrestrial radio access network (e-utran);overall description; stage 2 (release 10), 3 2009.

[Ara] New wireless networking system brings eye care to thousands in india. `http://www.berkeley.edu/news/media/releases/2006/06/06_telemedicine.shtml,june2006.u.c.berkeleynews.`

[Ash] Project ashwini. `http://www.byrrajufoundation.org/ashwini_overview.htm`.

[AWW05] Ian F. Akyildiz, Xudong Wang, and Weilin Wang. Wireless mesh networks: a survey. *Comput. Netw. ISDN Syst.*, 47(4):445–487, March 2005.

[AZ11] Atef Abdrabou and Weihua Zhuang. Probabilistic delay control and road side unit placement for vehicular ad hoc networks with disrupted connectivity. *IEEE Journal on Selected Areas in Communications*, 29(1):129–139, 2011.

[Ben05] Abderrahim Benslimane. Localization in vehicular ad hoc networks. In *Proceedings of the 2005 Systems Communications*, ICW '05, pages 19–25, Washington, DC, USA, 2005. IEEE Computer Society.

[BGJL06] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine. Maxprop: Routing for vehicle-based disruption-tolerant networks. In *INFOCOM 2006. 25th IEEE*

International Conference on Computer Communications. Proceedings, pages 1 –11, april 2006.

[BHM$^+$06] Vladimir Bychkovsky, Bret Hull, Allen Miu, Hari Balakrishnan, and Samuel Madden. A measurement study of vehicular internet access using in situ wi-fi networks. In *MOBICOM 2006*, pages 50–61, 2006.

[BL07] Andria Bilich and Kristine M. Larson. Mapping the gps multipath environment using the signal-to-noise ratio (snr). *Radio Science*, 42(6):n/a–n/a, 2007. RS6003.

[BLA06] Andria Bilich, Kristine M. Larson, and Penina Axelrad. Observations of signal-to-noise ratios (snr) at geode- tic gps site casa: Implications for phase multipath, 2006.

[CC09] Bin Bin Chen and Mun Choon Chan. Mobtorrent: A framework for mobile internet access from vehicles. In *Proc. IEEE INFOCOM 2009*, pages 1404–1412, 2009.

[Che90] Kan Chen. Driver information systems a north american perspective. *Transportation*, 17(3):251–262, 1990.

[CZ14] Jiannong Cao and Chisheng Zhang. *Seamless and Secure Communications over Heterogeneous Wireless Networks.* Springer Briefs in Computer Science. Springer, 2014.

[DB10] N.M. Drawil and O. Basir. Intervehicle-communication-assisted localization. *Intelligent Transportation Systems, IEEE Transactions on*, 11(3):678–691, Sept 2010.

[Dju] The story of a project to support the weak it infrastructure in an low populated area of denmark. `http://djurslands.net/biblioteket/international/djurslands_net_english_presentation.ppt`.

[DNP04] Shirshanka Das, Alok Nandan, and Giovanni Pau. Spawn: a swarming protocol for vehicular ad-hoc wireless networks. In *Proceedings of the First International Workshop on Vehicular Ad Hoc Networks, 2004, Philadelphia, PA, USA, October 1, 2004*, pages 93–94, 2004.

[DPZ04] Richard Draves, Jitendra Padhye, and Brian Zill. Comparison of routing metrics for static multi-hop wireless networks. In *SIGCOMM '04*, pages 133–144, New York, NY, USA, 2004. ACM.

[DTH02] O. Dousse, P. Thiran, and M. Hasler. Connectivity in ad-hoc and hybrid networks. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 1079 – 1088 vol.2, 2002.

[Eri11] Ericsson. Mobile data traffic growth doubled over one year, 10 2011.

[FL04] Chuan Heng Foh and Bu-Sung Lee. A closed form network connectivity formula one-dimensional manets. In *Proceedings of IEEE International Conference on Communications, ICC 2004, Paris, France, 20-24 June 2004*, pages 3739–3742, 2004.

[GP15] Ankit Gupta and Nitin Pundir. Pedestrian flow characteristics studies: A review. *Transport Reviews*, 35(4):445–465, 2015.

[HBZ$^+$06] Bret Hull, Vladimir Bychkovsky, Yang Zhang, Kevin Chen, Michel Goraczko, Allen Miu, Eugene Shih, Hari Balakrishnan, and Samuel Madden. Cartel: A distributed mobile sensor computing system. In *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems*, SenSys '06, pages 125–138, New York, NY, USA, 2006. ACM.

[HGNH09] Ying Huang, Yan Gao, K. Nahrstedt, and Wenbo He. Optimizing file retrieval in delay-tolerant content distribution community. In *Distributed Computing Systems, 2009. ICDCS '09. 29th IEEE International Conference on*, pages 308–316, June 2009.

[JGG$^+$11] Jaehoon Jeong, Shuo Guo, Yu Gu, Tian He, and D.H.C. Du. Trajectory-based data forwarding for light-traffic vehicular ad hoc networks. *Parallel and Distributed Systems, IEEE Transactions on*, 22(5):743–757, May 2011.

[JGG$^+$12] Jaehoon Jeong, Shuo Guo, Yu Gu, Tian He, and D.H.C. Du. Trajectory-based statistical forwarding for multihop infrastructure-to-vehicle data delivery. *Mobile Computing, IEEE Transactions on*, 11(10):1523–1537, Oct 2012.

[Kaw90] Hironao Kawashima. Japanese perspective of driver information systems. *Transportation*, 17(3):263–284, 1990.

[KMN99] Samir Khuller, Anna Moss, and Joseph (Seffi) Naor. The budgeted maximum coverage problem. *Information Processing Letters*, 70(1):39 – 45, 1999.

[KPD$^+$09] M. Kafsi, Panagiotis Papadimitratos, Olivier Dousse, Tansu Alpcan, and Jean-Pierre Hubaux. Vanet connectivity analysis. *CoRR*, abs/0912.5527, 2009.

[LCG09] Uichin Lee, Ryan Cheung, and Mario Gerla. Emerging vehicular applications. In S. Olariu and M. C. Weigle, editors, Vehicular Networks: From Theory to Practice. Chapman and Hall/CRC, 2009.

[LLPG10] Uichin Lee, Jiyeon Lee, Joon-Sang Park, and M. Gerla. Fleanet: A virtual market place on vehicular networks. *Vehicular Technology, IEEE Transactions on*, 59(1):344–355, Jan 2010.

[LLS12] T.H. Luan, Xinhua Ling, and Xuemin Shen. Mac in motion: Impact of mobility on the mac of drive-thru internet. *Mobile Computing, IEEE Transactions on*, 11(2):305–319, Feb 2012.

[LMG+09] Uichin Lee, E. Magistretti, M. Gerla, P. Bellavista, and A. Corradi. Dissemination and harvesting of urban data using vehicular sensing platforms. *Vehicular Technology, IEEE Transactions on*, 58(2):882–901, Feb 2009.

[LSK+09] SuKyoung Lee, K. Sriram, Kyungsoo Kim, Yoon Hyuk Kim, and N. Golmie. Vertical handoff decision algorithms for providing optimized performance in heterogeneous wireless networks. *Vehicular Technology, IEEE Transactions on*, 58(2):865 –881, feb. 2009.

[LsPY+06] Uichin Lee, Joon sang Park, Joseph Yeh, Giovanni Pau, and Mario Gerla. Codetorrent: Content distribution using network coding in vanets. In *In MobiShare06*, 2006.

[Ltd] Cisco Co Ltd. Cisco wireless mesh networking solution overview. Technical report.

[LZG+06] Uichin Lee, Biao Zhou, M. Gerla, E. Magistretti, P. Bellavista, and A. Corradi. Mobeyes: smart mobs for urban monitoring with a vehicular sensor network. *Wireless Communications, IEEE*, 13(5):52–57, October 2006.

[MA06] Daniele Miorandi and Eitan Altman. Connectivity in one-dimensional ad hoc networks: A queueing theoretical approach. *Wireless Networks*, 12(5):573–587, 2006.

[Mal11] Om Malik. Mobile data traffic doubled in past 12 months, 10 2011.

[Nep] Nepal wireless. `http://www.nepalwireless.net/`.

[NZZ+11] Seh Chun Ng, Wuxiong Zhang, Yu Zhang, Yang Yang, and Guoqiang Mao. Analysis of access and connectivity probabilities in vehicular relay networks. *Selected Areas in Communications, IEEE Journal on*, 29(1):140 –150, january 2011.

[OK04] J. Ott and D. Kutscher. Drive-thru internet: Ieee 802.11b for "automobile" users. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 1, pages 4 vol. (xxxv+2866), 2004.

[OK05] J. Ott and D. Kutscher. A disconnection-tolerant transport for drive-thru internet environments. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 3, pages 1849–1862 vol. 3, March 2005.

[PBN03] J. Snyde P. Bouvard, L. Rosin and J. Noel. The arbition national in-car study. 2003.

[PLMD12] Daan Pareit, Bart Lannoo, Ingrid Moerman, and Piet Demeester. The history of wimax: A complete survey of the evolution in certification and standardization for ieee 802.16 and wimax. *Communications Surveys Tutorials, IEEE*, 14(4):1183 –1211, quarter 2012.

[Ram] Bhaskaran Raman. Digital gangetic plains: 802.11-based low-cost networking for rural areas, 2001-2004: A report. `http://www.cse.iitk.ac.in/users/braman/dgp.html,2004.`

[RS07] Anton Riedl and Dominic A. Schupke. Routing optimization in ip networks utilizing additive and concave link metrics. *IEEE/ACM Trans. Netw.*, 15(5):1136–1148, 2007.

[Sik10] B. Sikdar. Characterization and abatement of the reassociation overhead in vehicle to roadside networks. *Communications, IEEE Transactions on*, 58(11):3296–3304, November 2010.

[SR07] Sayandeep Sen and Bhaskaran Raman. Long distance wireless mesh network planning: problem formulation and solution. In *WWW 2007*, pages 893–902, 2007.

[Sta] Wayne Stargardt. Condition-based maintenance using wireless monitoring: Developments and examples. EAM2008.

[TLYH11] Wee Lum Tan, Wing Cheong Lau, OnChing Yue, and Tan Hing Hui. Analytical models and performance evaluation of drive-thru internet systems. *Selected Areas in Communications, IEEE Journal on*, 29(1):207–222, January 2011.

[WFGH04] Hao Wu, Richard Fujimoto, Randall Guensler, and Michael Hunter. Mddv: a mobility-centric data dissemination algorithm for vehicular networks. In *Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, VANET '04, pages 47–56, New York, NY, USA, 2004. ACM.

[WIL11] JO WILKES. Ericsson: Mobile data traffic will grow 10-fold by 2016, 11 2011.

[YW08] Yaling Yang and Jun Wang. Design guidelines for routing metrics in multihop wireless networks. In *INFOCOM 2008.*, pages 1615–1623, 2008.

[ZC08] Jing Zhao and Guohong Cao. Vadd: Vehicle-assisted data delivery in vehicular ad hoc networks. *Vehicular Technology, IEEE Transactions on*, 57(3):1910 – 1922, may 2008.

[ZKL+07] Xiaolan Zhang, Jim Kurose, Brian Neil Levine, Don Towsley, and Honggang Zhang. Study of a bus-based disruption-tolerant network: Mobility modeling and impact on routing. In *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*, MobiCom '07, pages 195–206, New York, NY, USA, 2007. ACM.

[ZLSK10] Zizhan Zheng, Zhixue Lu, P. Sinha, and S. Kumar. Maximizing the contact opportunity for vehicular internet access. In *Proc. IEEE INFOCOM 2010*, pages 1–9, 2010.

[ZZC09] Yang Zhang, Jing Zhao, and Guohong Cao. Roadcast: A popularity aware content sharing scheme in vanets. In *Proc. 29th IEEE Int. Conf. Distributed Computing Systems ICDCS '09*, pages 223–230, 2009.