

INFORMATION DISSEMINATION IN VEHICULAR AD HOC NETWORKS

Andreas Xeros

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Cyprus

Recommended for Acceptance

by the Department of Computer Science

November, 2012

© Copyright by

Andreas Xeros

All Rights Reserved

2012

APPROVAL PAGE

Doctor of Philosophy Dissertation

INFORMATION DISSEMINATION IN VEHICULAR AD HOC NETWORKS

Presented by

Andreas Xeros

Research Supervisor

Andreas Pitsillides

Research Co-supervisor

Marios Lestas

Committee Member

Marios Dikaiakos

Committee Member

Vasos Vassiliou

Committee Member

Carla Fabiana Chiasserini

Committee Member

Dimitri Konstantas

University of Cyprus

November, 2012

INFORMATION DISSEMINATION IN VEHICULAR AD HOC NETWORKS

Andreas Xeros

University of Cyprus, 2012

VANETs harness the potential of information and communication technologies to create a safer, smarter and more efficient transportation network. Many of the applications of VANETs, especially the safety related ones, set up requirements for information dissemination that are different from conventional networks and thus, difficult to fulfill with existing strategies. Safety applications pose stringent delay requirements on emergency message delivery and address geographical areas in which data needs to be cooperatively collected, distributed and maintained. Design challenges are then posed by the variable node density along the transportation network, the high mobility, the confined but often unpredictable movement of the vehicles, and the unreliable radio channel. In this thesis we address the information dissemination problem in VANETs, with particular interest in routing, information hovering and broadcast schemes, and propose new protocols and methods which overcome the design challenges outlined above.

We first address the information hovering problem in VANETs. We propose an adaptive information hovering scheme which is based on the application of epidemic routing within the hovering area and probabilistic flooding outside the hovering area. A unique feature of the proposed protocol is that it is adaptive in the sense that the rebroadcast probability outside the hovering area is adaptively regulated based on estimates of the vehicle density within the hovering area. The designed protocol is amenable to gradual deployment and has good robustness properties, as it does not rely on the existence of static infrastructure on the roadside, base stations or road assistance. It relies solely on V2V communication. The protocol is shown to achieve its design

objectives, as it achieves high reachability without overloading the network resources through a large number of exchanged messages. The methods used are attractive, as they lead to protocols with universal properties in the sense that they can be utilized as effective solutions in areas beyond the ad hoc vehicular technology area considered so far. In this thesis we integrate the derived information hovering protocol in a Data Broadcasting system with a Push Server, as a means of data dissemination off-loading that significantly improves performance. Such extensions can be further investigated in areas such as cloud computing, mobile cloud and information boomerang.

We then address the information routing problem in VANETs. We pursue efficient routing policies, which can ensure maximum probability of successful information delivery in target areas before a specific deadline expires. We model the considered road map as a directed weighted graph which is then used to calculate the probability of disseminating information along any given path, in a specific amount of time. So, for any two points on the road map we can find the path with the maximum probability of successful information delivery in the chosen time interval. This information is of great significance, since it can be utilized by routing protocols to optimally route packets in the vehicular network. The creation of the weighted graph requires the calculation of the lower bound on the probability of information propagation between two intersecting roads in a given time period. We show that the propagation probability is strongly related to the traffic conditions of the road where the information is to be transmitted. We use the derived formula to estimate, via simulations, the minimal conditions required to ensure that information propagation occurs with high probability on intersections. The research presented in this thesis has both theoretical analysis and simulation aspects. In every proposed solution supporting information routing

or information hovering there is a theoretical analysis that is also validated using realistic simulations. All simulations have been run on VISSIM, supplemented by the simulation of the mobile ad hoc network using a C++ application.

Andreas Xeros

INFORMATION DISSEMINATION IN VEHICULAR AD HOC NETWORKS

Andreas Xeros

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Cyprus

Recommended for Acceptance

by the Department of Computer Science

November, 2012

© Copyright by

Andreas Xeros

All Rights Reserved

2012

APPROVAL PAGE

Doctor of Philosophy Dissertation

INFORMATION DISSEMINATION IN VEHICULAR AD HOC NETWORKS

Presented by

Andreas Xeros

Andreas Xeros

University of Cyprus

November, 2012

ACKNOWLEDGEMENTS

First and foremost I would like to express my deep gratitude to my advisor, Dr. Andreas Pitsillides. I highly appreciate the way he contributed to me completing this dissertation, not despite, but because of his challenging comments. He helped me deepen my understanding of the subject matter, and provided me with encouragement from the beginning to the concluding stage. I feel privileged to have had him as my supervisor, and I am very thankful for the excellent example he provided me with.

Furthermore, I am greatly thankful to Dr. Marios Lestas and Dr. Maria Andreou, who have shown me, both consciously and unconsciously, that graduate school is worth the time and effort, contributing immensely to my personal and professional time as a Master's and a Ph.D. student. The joy and zeal they have for research was a strong motivation for me as well. In addition, I would also like to thank Dr. Vasos Vassiliou for the constructive advice he generously offered. Together we had extensive and very inspirational discussions, and his contribution has been truly valuable, as he helped made my experience working towards the Ph.D. both productive and stimulating.

This dissertation would never have been completed without the subtle and often times very evident support from my family. My mother, Maro, has always encouraged me to proceed in academia, as she, a teacher herself, could foresee the numerous benefits education can offer. And now, when the subject matter is not within her realm of knowledge, she continues to provide me with real and invaluable help, looking after my children when I, because of the various tasks I decided to undertake, could not. Likewise, I would like to express my biggest gratitude to my

sister, Effie, without whom my following and completing graduate studies would not have been possible.

Although during my Ph.D. studies I was blessed to have children of my own, joy came at the same time with sadness, as it often happens in life. My father may have passed away almost three years ago, but his memory has only grown stronger. He always believed strongly in the academic potential of his children, both of whom have followed Ph.D. studies. A part of me still hopes that he is not unaware of the big events that are occurring in our family, especially his newly born grandchildren.

Last but not least, I would like to thank my loving wife, Barbara. Without her I would be a very different, less improved, person today. I can't even begin to imagine how difficult giving birth to four children must have been for her, and yet she has always been immensely supportive of my academic and professional endeavors. Barbara and our four children, Giorgos, Theodora, Marios and Nicolas, have given me a purpose for everything I strive to do in my life, including this dissertation. And for that, my mere expression of thanks will never suffice.

LIST OF ACRONYMS, SYMBOLS & ABBREVIATIONS

Ad-Hoc	Network formed with little or no planning
A-STAR	Anchor based Street and Traffic Aware Routing
AU	Application Units
BF	Blind Flooding
BPSK	Binary Phase Shift Keying
CBF	Contention-Based Forwarding
CCH	Control CHannel
CCU	Communication & Control Unit
CodeON	Push-based popular content distribution scheme
CodeTorrent	Content Distribution using Network Coding
DAB	Digital Audio Broadcast
DGPS	Differential Global Positioning System
DSR	Dynamic Source Routing
DSRC	Dedicated Short Range Communication
DVB	Digital Video Broadcast
DVB/H	Digital Video Broadcast / Handheld
ECU	Electronic Control Units
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission
GPS	Global Positioning System
GPSR	Greedy Perimeter Stateless Routing
GSM	Global System for Mobile Communications
GSR	Geographic Source Routing
GyTAR	Greedy Traffic Aware Routing
HOV/Bus Lanes	High Occupancy Vehicle and Bus Lanes
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPv6	Internet Protocol Version 6
ISO	International Standards Organization
ISTEA	Intermodal Surface Transportation Efficiency Act
ITS	Intelligent Transportation Systems
ITSA	Intelligent Transportation Society of America
IVHS	Intelligent Vehicle Highway System
LAN	Local Area Network
LLC	Logical Link Control
MAC	Medium Access Control

MANETs	Mobile Ad-hoc Networks
MDDV	Mobility-Centric Data Dissemination Algorithm
MLME	MAC Layer Management Entity
MWT	Mean Waiting Time
NC	Network Coding
NIST	National Institute of Standards and Technology
OFDM	Orthogonal Frequency Division Multiplexing
P2P	Peer to Peer
PHY	PHYSical Layer
PLME	Physical Layer Management Entity
PRE-DRIVE	PREparation for DRIVING implementation and Evaluation
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying
RLS	Reactive Location Service
SCH	Service CHannels
SF	Selective Forwarding
TCP	Transmission Control Protocol
TO-GO	TOpology-assist Geo-Opportunistic Routing
TTL	Time To Live
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
V2I	Vehicle to Infrastructure
V2R	Vehicle to Roadside Unit
V2V	Vehicle to Vehicle
VANET	Vehicular Ad Hoc Network
VMS	Variable Message Signs
VSC	Vehicle Safety Communications
WAVE	Wireless Access in Vehicular Environment
WBSS	WAVE Basic Service Set
WME	WAVE Management Entity
WSM	WAVE Short Message
WSMP	WAVE Short Message Protocol

TABLE OF CONTENTS

Chapter 1: Introduction	1
1.1 Motivation	4
1.2 Thesis statement	8
1.3 Approach	9
1.4 Contributions	10
1.5 Acknowledgments of Contributions to this Thesis	17
1.6 Publications	18
1.6.1 List of publications related to this thesis	18
1.6.2 List of publications beyond the work presented in this thesis	20
1.7 Thesis Overview	21
Chapter 2: Background and Related Work	22
2.1 Intelligent Transportation Systems	22
2.2 Application Categories in VANETs	24
2.3 ITS Communication Architecture	27
2.3.1 ITS Vehicle Station	28
2.3.2 ITS Roadside Station	29
2.3.3 ITS Central Station	31
2.3.4 ITS Personal Station	31
2.4 Wireless Access in Vehicular Environment (WAVE)	32
2.4.1 WAVE Physical Layer	34
2.4.2 Wave Channel Coordination	35
2.4.3 WAVE Communication Protocol	35

2.4.4	WAVE Management Plane	37
2.5	Related Work	37
2.5.1	Information Hovering	39
2.5.2	Push Server - Data Broadcasting	41
2.5.3	Information Routing	42
Chapter 3:	Information Hovering in Vehicular AdHoc Networks	46
3.1	Introduction	46
3.1.1	Phase Transition, Percolation Theory and Random Graphs	48
3.2	Problem Formulation	50
3.3	Proposed Traffic Adaptive Information Hovering Protocol	51
3.4	Estimation of the Vehicle Density	64
3.5	Performance evaluation	71
3.6	Conclusions and Sum Up	86
Chapter 4:	Information Hovering: A New Approach for Performance Acceleration of Wireless Push Systems	88
4.1	Introduction	88
4.2	The Proposed Collaborative Scheme	90
4.3	Simulation	102
4.3.1	A. Simulation Setup and Results	102
4.4	Conclusions	110
Chapter 5:	Information Propagation Probability on Intersections in VANETs	111
5.1	Introduction	111
5.2	Problem Formulation	112

5.3	Theoretical Analysis of Message Propagation Probability on Intersections	114
5.3.1	Probability of Transmission of Information among vehicles on intersect- ing roads ($p_{tr}^{h_{ij}h_{jk}}$)	115
5.3.2	Driving Probability where veh_1 turns into road h_{jk} ($p_{dr}^{h_{ij}h_{jk}}$)	124
5.4	Simulation Validation	126
5.5	Conclusions / Sum up	131
Chapter 6:	Time Constrained Message Delivery Probability in VANETs	132
6.1	Introduction	132
6.2	Problem Formulation - Assumptions	132
6.3	Road-Graph Construction	134
6.3.1	Analysis of the Algorithm: "Construction of the Road-Graph"	136
6.4	MDP	142
6.5	Real World Example	145
6.5.1	Analytical Results	147
6.6	Simulation Validation	151
6.7	Conclusions and Sum Up	154
Chapter 7:	Conclusions and Future Work	156
Bibliography		171

LIST OF TABLES

1	Overview of applications in VANETs	25
2	WAVE Standard	33

Andreas Xeros

LIST OF FIGURES

1	Smart Vehicle with the basic modules	2
2	ITS Communication Architecture illustrates the communication system designed for ITS [110]. 1. Vehicle Station, 2. Personal Station, 3. Roadside Station, 4. Central Station.	28
3	ITS Vehicle Station	28
4	ITS Roadside Station (1.Management, 2.Facilities, 3.Networking & Transport, 4.Access Technologies, 5.Security, 6.Applications)	30
5	ITS Central Station	31
6	ITS Personal Station	32
7	Wave System Components	32
8	Wave Protocol Stack	33
9	Spectrum of Wave Channel	34
10	WSM Frame Format	36
11	Candidate probability functions. Their parameters have been tuned to exhibit similar decreasing behavior.	54
12	Road map of the reference model used in all simulation experiments. It represents a section of the road network in the cities of Bellevue and Redmond in Washington. Hovering areas are marked as black circles.	56
13	Number of messages exchanged in area A for each of the candidate probability functions.	58
14	Reachability achieved in area A for each of the candidate probability functions.	58

15 Number of messages exchanged and Reachability in hovering area A for different values of σ when the vehicle density is set to $6.7 \text{ Veh}/\text{Km}^2$ 61

16 Reachability vs number of message exchanged in hovering area B, for different values of σ when the traffic density is set to $15 \text{ Veh}/\text{Km}^2$ 62

17 Reachability vs number of message exchanged in hovering area C, for different values of σ when the traffic density is set to $4.9 \text{ Veh}/\text{Km}^2$ 62

18 Critical value of sigma vs traffic density graphs in different hovering areas. We observe similar behavior in all hovering areas. The relationship used in the proposed protocol is obtained by applying a least squares fit between the curves. 64

19 The road topology of the model used to relate analytically the vehicle density with the average number of neighbors. 65

20 Average number of neighbors obtained using simulations in different hovering areas, for different average number of vehicles in the areas and comparison with the predicted values using equation (13) 70

21 Road topology of the reference model used to evaluate the performance of the proposed information hovering protocol. The considered hovering areas are marked on the diagram as black circles. 72

22 Comparison of the proposed adaptive probabilistic scheme with other hovering schemes in terms of the reachability achieved and the number of messages exchanged in area H, for different average number of vehicles in the area. 74

23 Comparison of the proposed adaptive probabilistic scheme with other hovering schemes in terms of the reachability achieved and the number of messages exchanged in area G, for different average number of vehicles in the area. 75

24	Comparison of the proposed hovering scheme using the vehicle density estimation of equation (13), with a protocol that assumes the vehicle density within the hovering area to be known. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area H for different average number of vehicles in the area.	77
25	Comparison of the proposed hovering scheme using the vehicle density estimation of equation (13) with a protocol which assumes the vehicle density within the hovering area to be known. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area G for different average number of vehicles in the area.	78
26	Percentage increase of the critical radius relative to the radius of the hovering area vs traffic density graphs in different hovering areas. We observe similar behavior in all hovering areas. The relationship used in the adaptive version of the protocol proposed in [37] is obtained by applying a least squares fit between the curves. . .	80
27	Comparison of the proposed adaptive probabilistic flooding hovering scheme with the adaptive version of the hovering protocol proposed in [37]. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area H for different average number of vehicles in the area.	82
28	Comparison of the proposed adaptive probabilistic flooding hovering scheme with the adaptive version of the hovering protocol proposed in [37]. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area G for different average number of vehicles in the area.	83

29	Comparison of the four techniques achieving information hovering. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area A for different values of the radius of the circular hovering area	85
30	Illustration of the proposed load sharing between the broadcasting and the hovering systems. In a total of Q queries, $p_i \cdot Q$ refer to item i . A percentage $h_i \in [0, 1]$ of these are addressed by the hovering system, while the remaining are answered via the broadcasting.	92
31	The simulated VANET area of Bellevue and Redmond, Washington, as depicted in Google Maps (Coordinates: 47.626413, -122.149944). A number of events (dots) regarding traffic, road condition and parking availability are selected. Information on these events is disseminated via broadcasting (global coverage) and hovering around the corresponding locations.	103
32	Performance of the learning automaton-based adaptation technique, which is used for approximating the time-variant popularity of a dynamically formed data set. Random changes are set to occur in the cardinality of the studied data set and the popularity and size of the contained items. Even in the worst-case scenario (no content correlation between successive data set updates), the automaton adapts optimally in 50 - 250 sec.	104
33	Plots of the requested strain \bar{B}_u , versus the achieved mean waiting time \bar{W} and the actual imposed strain. The plots reflect the case of $\theta = 0.1$, i.e. a slightly skewed ZIPF distribution of the items' popularity. The waiting time is reduced by 10 times when compared to the case of stand-alone broadcasting ($\bar{B}_u = 0$). While the reduction follows the general theoretical form, some degree of divergence is present. Results are identical for other values of θ .	106

34	Time required for the initialization of the compared P2P schemes in an area of interest. Hovering requires a trivial amount of time before being established. CodeOn and CodeTorrent may be viable, requiring approximately 5sec for establishment in the best case.	107
35	The mean waiting time achieved by all compared approaches in a full simulation. Steady state is assumed, with the initialization and adaptation phases having been completed. The hovering-broadcasting combination yields the best results in the tested cases of locality of demand.	108
36	Average time required to obtain a data item, if the client is inside the corresponding area of interest. Hovering achieves minimal time, with CodeOn and CodeTorrent following. The results also imply that network coding may not be more beneficial than an efficient routing algorithm.	108
37	Alternative representation of the performance of the opposing schemes in terms of user satisfaction (try-best metric of [134]). The hovering-broadcasting combination achieves the best results, for all items $i = 6 - 100$ chosen for dissemination through P2P networking.	109
38	Representation of an intersection in a road network presenting the basic notations .	113
39	Influence of angle φ to $d(t)$ where the rest of the variables are constant. Their values are $r : 250m, V_1 : 11m/s, t: 7 \text{ sec.}$ and $X(t)$ vs time	117
40	Example of catching up case. Plot of $d(t)$ and $X(t)$ vs time	119
41	Case 2 scenario where veh_1 received the information at some point after passing from point R	120

42	The rare scenario where there is no informed vehicle in the area RI_j during the period of $[0, y]$. (a) veh_1 received the information message from veh_2 while being in the area RI_j . (b) veh_1 passed the intersection I_j before veh_2 had passed from point R resulting in no informed vehicles in the specific area.	122
43	Example showing how to calculate the probability of information propagation with the driving way	125
44	Intersection setup used in VISSIM	126
45	Theoretical results of Transmission Probability for different arrival rates	128
46	Comparing Theoretical with Simulation results of Transmission Probability for arrival rates 144 Veh/h and 540 Veh/h	129
47	Comparing Theoretical with Simulation results of Transmission Probability after 14 seconds	130
48	An example of message propagation	137
49	Catch-up Process	138
50	Theoretical and Simulation results of Information Propagation speed on a single lane road	140
51	Representation of a simple road map with 3 roads and 4 intersections	141
52	The Road-Graph that corresponds to the road map of figure 51	142
53	Major highways around the city of Los Angeles	146
54	The Road-Graph of major highways in the Los Angeles Area. S_i and D_i indicate the scenarios that we study (Source-Destination)	147
55	Information Delivery Probability versus Time from intersection I_{13} to intersection I_{53}	148

56	Information Delivery Probability versus Time from intersection I_1 to intersection I_{51}	148
57	Information Delivery Probability versus Time for three different portions of equipped vehicles	151
58	Information Delivery Probability versus Time from intersection I_{15} to intersection I_{41}	153
59	Information Delivery Probability versus Time from intersection I_3 to intersection I_{22}	154

Andreas Xeros

Chapter 1

Introduction

The rapid evolution of wireless data communication technologies, which emerged in the last few years, has led researchers to explore their application in Mobile Ad-hoc Networks (MANETs). MANETs are self-organized mobile wireless networks which are independent from infrastructure [1, 5, 80, 85, 120]. Mobile nodes are connected via wireless links forming networks of arbitrary topology [30, 33, 123]. Nodes are free to move randomly and organize themselves arbitrarily; thus, the network's wireless topology may change rapidly and unpredictably. Such a network may operate in a stand-alone fashion, or may be connected to a larger network.

The main task of a special class of these networks is to collect (e.g. by using on-board sensors) and propagate information among their nodes which finally has to be processed and transmitted by other stations/users and/or base stations. Vehicular Ad-hoc Networks (VANETs) constitute a subset of MANETs where vehicles serve as mobile nodes and their mobility is restricted by the roadway topology. VANETs consist of instrumented vehicles which are able to collect, process and communicate information among each other when their distance is within their transmission range.

As shown in figure 1, vehicles participating in VANETs -among others- are equipped with on-board sensors, a wireless communication system, a positioning system, a digital road map, a processor and a memory unit. Communicating vehicles exchange information messages which usually consist of a message header and a message body. Examples of header data include the ID of the vehicle which has initiated the transmission, the message ID, the time of creation, the time to live, the target area [124, 75], etc. The message body can consist of different types of data either raw or processed, depending on the application.

Vehicular ad-hoc networks, although being a subclass of mobile ad-hoc networks, have unique characteristics, which differentiate them from traditional MANETs. VANETs are not constrained by scarce energy resources, but are rather characterized by high mobility patterns and confined movement. High mobility is a result of the large speeds, which the vehicles can attain, leading to dynamic and rapidly changing network topologies and possible network fragmentation. The dynamic nature of the topology is exacerbated by the unpredictable nature of the drivers' response to various events. VANETs are also characterized by the constrained, largely one dimensional movement of the vehicles along the roadway network, which is fixed. In addition, unlike most MANET technologies, static infrastructure consisting mainly of roadside units is not uncommon in VANET systems and can, where available, actively support message transmission. The aforementioned

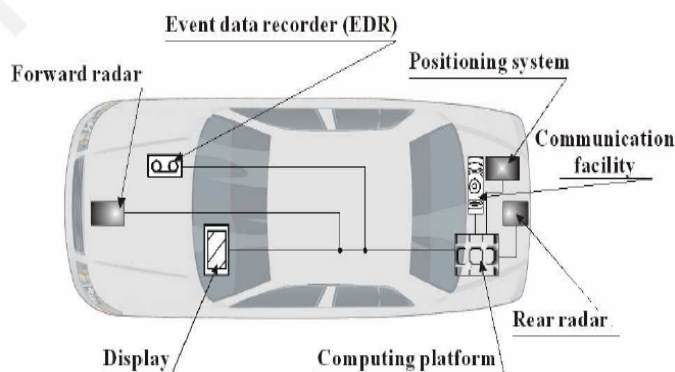


Figure 1: Smart Vehicle with the basic modules

VANETs characteristics pose design and modeling challenges different from traditional MANETs in the development of various network protocols as well as approaches to solve a variety of issues.

VANETs harness the potential of information and communication technologies to create a safer, smarter and more efficient transportation network. The recently published 802.11p [17] standard supports both vehicle-to-vehicle and vehicle-to-infrastructure communications, allowing the formation of vehicular ad hoc networks which are envisioned to accommodate the new generation of cooperative safety applications. The range of applications of VANETs goes beyond the safety related ones, so as to include traffic monitoring, platooning, text messaging, distributed passenger teleconferencing, file sharing (for example music downloading), roadside e-advertisements etc.

Many of the aforementioned applications in VANETs, especially the safety related ones, set up requirements for information dissemination [36, 56, 121, 125] which are different from conventional networks and are thus difficult to accomplish with existing strategies. Safety applications pose stringent delay requirements on emergency message delivery and address geographical areas in which data needs to be cooperatively collected, distributed and maintained. Design challenges are then posed by the variable node density along the transportation network, the high mobility, the confined but often unpredictable movement of the vehicles and the unreliable radio channel. Variations in traffic density are of particular importance, as low traffic densities cause the network to become intermittently connected, whereas high traffic densities lead to excessive contention. These phenomena significantly degrade the performance of data dissemination strategies whether these are routing protocols or broadcast-based schemes.

In this thesis we address the information dissemination problem in VANETs with particular interest in *information hovering* schemes and in *routing*, and we propose new protocols and methods which overcome the design challenges outlined above.

We first address the information hovering problem in VANETs. We propose an adaptive information hovering scheme which is based on the application of epidemic routing within the hovering area and probabilistic flooding outside the hovering area. A unique feature of the proposed protocol is that it is adaptive in the sense that the rebroadcast probability outside the hovering area is adaptively regulated based on estimates of the vehicle density within the hovering area. The designed protocol is amenable to gradual deployment and has good robustness properties, as it does not rely on the existence of static infrastructure on the roadside, base stations or road assistance. It relies solely on V2V communication. Furthermore the methods used are appealing as they lead to protocols with universal properties in the sense that they can be utilized as effective solutions in areas beyond the ad hoc vehicular technology area considered so far. As an example, in this thesis we integrate the derived information hovering protocol in a Data Broadcasting system with a Push Server, as a mean of data dissemination off-loading which significantly improves performance. Such extensions can be further investigated in areas such as cloud computing, mobile cloud and information boomerang.

We then address the information routing problem in VANETs. Our findings include the derivation of the probability of information propagation on intersections and its utilization in estimating the message delivery probability over long distances in a predetermined time interval. The latter is used to find the maximum probability of successful information delivery in a chosen time interval for a specific path. This information is of great significance, since it can be utilized by routing protocols to optimally route packets in a vehicular network.

1.1 Motivation

In the last twenty years there has been a huge increase of the number of passenger cars worldwide. In 2008 there were approximately 850 millions cars and light trucks on the roads (roughly

one car per 10 people) and the numbers are increasing rapidly, especially in China and India [84]. This number has been projected to be over 2 billions in 2030 [13]. The usage of automobiles provides many societal benefits, including economy benefits such as job and wealth creation of automobile production and maintenance, transportation provision, society wellbeing derived from leisure and travel opportunities, and revenue generation from the tax opportunities. The ability for humans to move flexibly from place to place has far reaching beneficial implications for the nature of societies [13]. However, there are many negative effects of the excessive use of automotive such as maintaining roads, land use, pollution and disposing of the vehicle at the end of its life. The most serious effect is, of course, car collisions, which kill 1.2 million people worldwide each year, and injure about forty times that number [92].

Government agencies and automotive industries are responding by investing billions of dollars in an effort to reduce these terrifying numbers, as well as the tremendous costs associated with vehicle damages and treating crash victims [113]. Wireless technologies are seen as playing a vital role in this effort. Dedicated Short Range Communication (DSRC) [106] is the variant IEEE 802.11a [41] standard which supports vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications forming Vehicular Ad Hoc Networks. VANETs are expected to accommodate a new generation of cooperative road safety applications.

Messages exchanged in VANETs increase the range of awareness of drivers beyond their line-of-sight, thus significantly improving safety and comfort of all passengers in a vehicle. Information on emergency actions like emergency braking or hazards on the road can provide measures for active safety. In addition, a plethora of other applications can be supported by the Vehicular Network such as collision avoidance and dynamic route scheduling [99, 108, 119], real-time traffic condition monitoring [4, 15, 82] and entertainment of passengers including interactive games, chatting, file sharing [14, 28, 90] etc.

Many applications in VANETs, especially safety related ones, require time sensitive message delivery, often over large distances. In many cases, information needs to propagate several kilometers away in a certain amount of time, otherwise message delivery loses its usefulness. For example, in case of a car accident, beyond its value to the nearest cars, information also needs to travel to the nearest hospital, as soon as possible. Another example is in the case of road hazard, in which information needs to reach the nearest police station, which can be far away from the incident. With recent advances in multimedia technologies, information may also include photos or even videos of the situation, helping the authorities to have a better understanding of the seriousness of the case. In addition, non-safety applications such as queries of weather conditions or traffic jams of a remote area, or even interactive games with remote users, involve the exchange of messages between vehicles that reside some distance away from each other. In all these cases, messages need to cover distances of several kilometers in a multi-hop fashion in variable density traffic conditions. These density variations can significantly affect the time taken to deliver the message. High traffic densities usually lead to smaller delivery times due to higher probabilities of the network being connected from source to destination. Of course, very high traffic densities can also cause delays due to higher probability of the network becoming congested. The large number of exchanged messages between cars in dense traffic can cause overloading of the available network resources and thus congestion related delays. Such delays, however, can be avoided by suitable design of congestion control protocols. Such protocols can be end-to-end or network centric. A recent protocol which addresses this problem and manages to reduce the number of exchanged messages and almost eliminate the congestion related delays without affecting the achieved reachability is proposed in [81].

On the other hand, low traffic densities often cause the network to be intermittently connected. In such cases the message delivery speed is bottlenecked by the speed of the vehicle which is

used to convey the information, which could even reach zero in case of a more prolonged network disconnection. So, when the application poses constraints on the delivery time, the probability to satisfy these constraints can be greatly affected by the chosen path. It is thus important to design efficient routing protocols that make routing decisions based on the need to increase the probability of successful delivery in a specific amount of time.

In addition to the applications related to message delivery in a remote area, there are scenarios in VANETs involving the exchange of information messages which are logically attached to a specific geographic area (vehicles outside this area are presumed not to be interested in this information, as this information only has local significance). For example, in case of an unexpected event, such as a traffic accident, it is important for all vehicles residing in the surrounding area to be notified of the hazard in order to take appropriate safety measures. Furthermore, the information must continue to lie in the area for a specific amount of time so that new vehicles entering the area are notified of the imminent danger. The same applies to the case of road works, where warning messages can be issued to all vehicles in the area around the work zone to alert drivers of the irregular situation.

Another example involves commercial enterprises that wish to advertise their products to possible customers within the transportation network. These enterprises can take advantage of the existing vehicular Ad-Hoc network to disseminate their advertisements in an area around their physical location. In all the aforementioned situations, useful information must be broadcasted to all vehicles in the specific geographical area, and this information must be maintained in the area for a finite time interval dictated by the application, in order to notify drivers entering the area of this useful information. The requirements as described above are closely related to the more general concept of Information Hovering, also appearing in the literature as *abiding* or *time-stable*

geocasting. In this thesis we adopt the term “Information Hovering”. Furthermore, similar concepts exist in areas such as cloud computing, mobile cloud and information boomerang, which are beyond the scope of this thesis.

1.2 Thesis statement

This thesis addresses the aforementioned information dissemination problems and proposes effective solutions that are derived with the use of probabilistic methods. The most prominent characteristic of these methods is that they are adaptive in nature, in the sense that some of the design parameters are regulated based on traffic density data, either measured or estimated online. In particular, three information dissemination schemes are investigated, new protocols are proposed and theoretical results are derived and validated. The information hovering problem is first addressed and a solution which employs epidemic routing within the hovering area and probabilistic flooding outside the hovering area is proposed and shown to be effective. The unique characteristic of the proposed protocol is that it regulates the rebroadcast probability outside the hovering area, based on estimates of the vehicle density within the hovering area, which are calculated online. Next, the information hovering protocol is integrated in a push-pull broadcast system aiming to reducing the mean waiting times of the clients by offloading a portion of the broadcasted data to the local hovering schemes. Finally, lower bounds on the probability of information dissemination on intersections are then calculated and used to find optimal source-destination routes that maximize the successful delivery probability in a constrained time period.

1.3 Approach

The research presented in this thesis has both theoretical analysis and simulation aspects. In every proposed solution supporting information routing or information hovering there is a theoretical analysis which is also validated using realistic simulations. The traffic related aspects of the considered scenarios have been simulated on the VISSIM simulator, whereas the networking aspects have been realized on C++ application developed by the authors. VISSIM [96], is a widely used microscopic multi-modal traffic flow simulation software. It is based on a time step and behavior simulation approach, developed to model urban traffic and public transport operations and flows of pedestrians. It simulates vehicles driving in realistic road systems (e.g. highways, city, etc) and conditions. The program offers multimodality, meaning that it models all modes of transport and their interactions in one simulation: motorised private transport, freight transport and public transport, as well as cyclists and pedestrians.

The accuracy of a traffic simulation model is mainly dependent on the quality of the vehicle modeling, e.g. the methodology of moving vehicles through the network. VISSIM uses the psycho-physical driver behavior model developed by WIEDEMANN (1974). The basic concept of this model is that the driver of a faster moving vehicle starts to decelerate as he/she reaches his/her individual perception threshold to a slower moving vehicle. Since the driver cannot exactly determine the speed of that vehicle, his/her speed will fall below that vehicles speed until he/she starts to slightly accelerate again after reaching another perception threshold. This results in an iterative process of acceleration and deceleration.

VISSIM simulates the traffic flow by moving “driver-vehicle-units” through a network. Every driver with his/her specific behavior characteristics is assigned to a specific vehicle. As a consequence, the driving behavior corresponds to the technical capabilities of his/her vehicle. More details about VISSIM and the simulation model used appear in Appendix A.

The simulation of the mobile ad hoc network is performed using a C++ application designed to process the output provided by the VISSIM simulator. The application is based on the assumption that a vehicle can transmit information message to vehicles residing in its transmission range with a probability p which is user defined. User can also define the assignment of data transmission range of each vehicle. Hovering areas can be defined given the center (X, Y, coordinates) and the radius of the area in meters. Furthermore, in order to make simulations more realistic one can define the percentage of vehicles that are appropriately equipped to participate in VANETs. Finally, users can give the coordinates of road site units and their transmission range, which improves information propagation. Depending on the desired outputs, the C++ application is tuned accordingly¹.

1.4 Contributions

In this thesis we address the information dissemination problem in VANETs with particular interest in information hovering, broadcast schemes and routing without the necessity for roadside units. We propose new protocols and methods that overcome the design challenges outlined above. Since the system is highly stochastic in nature, our approaches are based on probabilistic methods. The main contributions of this work are summarized below:

We first consider the information hovering problem in VANETs and we propose a traffic aware information hovering protocol that is successful in achieving a high percentage of vehicles receiving the relevant message while at the same time reducing the number of exchanged messages.

¹Simulation scripts can be found on [http : //www.netrl.cs.ucy.ac.cy/index.php?option = com_content&task = view&id = 76&Itemid = 54](http://www.netrl.cs.ucy.ac.cy/index.php?option=com_content&task=view&id=76&Itemid=54)

Information Hovering is a new concept of information dissemination over a mobile set of peers. It naturally applies in many applications in VANETs, where useful information needs to be made available to all vehicles within a confined geographical area for a specific time interval. A straightforward approach is to have all vehicles within the hovering area exchange messages with each other. However, this method does not guarantee that all vehicles within the hovering area will receive the message, because of potential partitioning of the network in areas with low traffic density and/or low market penetration rate. To alleviate this problem in this work we propose a scheme based on the application of epidemic routing within the hovering area and probabilistic flooding outside the hovering area. Informed vehicles outside the area can serve as information bridges towards partitioned uninformed areas within the zone of interest, thus leading to high reachability.

The design methodology has been simulative, and a major challenge in the overall procedure has been the design of the rebroadcast probability function for the probabilistic flooding scheme. Among a number of proposed candidate functions, we choose the one yielding superior performance and we tune its parameters taking advantage of phase transition phenomena, which are typical in probabilistic flooding schemes [61, 81]. It is known that probabilistic flooding in mobile ad hoc networks is characterized by phase transition phenomena [53], similar to the ones observed in the context of random graphs and percolation theory [19], which suggest the existence of a critical rebroadcast probability value beyond which high reachability is achieved with high probability. In this work we also demonstrate the existence of a vehicle traffic density dependant critical rebroadcast probability outside the hovering area, beyond which high reachability is achieved within the hovering area.

A unique feature of the proposed protocol is that it is adaptive, in the sense that the rebroadcast probability outside the hovering area is adaptively regulated based on estimates of the vehicle density within the hovering area. Estimates of the vehicle density within the hovering area are

obtained using measurements of the number of neighbors of each vehicle. The formula relating the two quantities is derived using a simple model of the transportation network within the hovering area. We demonstrate through simulations that the proposed scheme is successful in satisfying the design objectives and outperforms other candidate hovering protocols such as epidemic routing in the entire network, epidemic routing in the hovering area only, and the scheme proposed in [37].

The latter scheme applies epidemic routing in an extended area which includes the hovering area. By forming information bridges towards uninformed areas in the hovering area, the scheme increases the achieved reachability, at the expense of larger number of messages exchanged. However, the authors do not discuss the size of the extended area and how to regulate it based on some variables (e.g. traffic density, road topology) to achieve optimal performance.

We also demonstrate that, if we utilize the same design methodology to render the protocol proposed in [37] adaptive, its performance becomes comparable to the performance of the scheme proposed in this work. Thus, our contribution goes beyond the proposal of a specific information hovering protocol and extends to the introduction of a design procedure that can be used to design a class of density adaptive hovering protocols.

Next, we demonstrate the universality of the proposed information hovering protocol and its applicability in areas beyond the Vehicular Networking field, by integrating it in a Push-based wireless broadcast system as a load offloading tool. The combination of wireless broadcasting and information hovering is shown to significantly decrease the mean waiting times of clients in various wireless data broadcasting scenarios. In this setup, the system's architecture typically considers a large number of clients dispersed over an area covered by a single wireless system. The clients are interested in a common set of data, albeit with varying individual preferences. A central authority creates a common data broadcast schedule and transmits it over the covered area. The central authority is more or less initially unaware of the popularity of each data item. This

knowledge is acquired gradually, typically through a lightweight, client feedback mechanism, and the schedule is adapted accordingly. This technique is typically applied to Instant Messaging Services, Tweets and audio-video broadcasting (e.g. DVB/H, DAB+).

In [86] for example, the clients are required to emit a single pulse upon reception of a wanted item. The aggregate received power level is then mapped to the probability distribution by a learning automaton [91]. Once the item probabilities become known, a central authority creates a broadcast schedule, which optimizes a given criterion.

In order to be able to reduce the mean waiting time, we present a novel, MANET-oriented merging of wireless broadcasting and information hovering. The logic of the merge follows the fact that any wireless broadcast system is typically assigned bandwidth that is inversely proportional to its area of coverage. Thus, in order for it to achieve adequate mean waiting times, a portion of the data dissemination task can be offloaded to a local hovering scheme. The number of items to be offloaded is analytically evaluated. Simulation under realistic conditions yielded superior performance to individual solutions. It is worth noting that addressing the last problem using the proposed information hovering scheme highlights the generality of the proposed methodology.

Finally, we address the information routing problem in VANETs and utilize a simple mathematical model to find analytically lower bounds on the probability of information propagation on intersections. The probability, as expected, is found to depend heavily on the traffic state in the vicinity of the intersection. The derived result is interesting and significant in its own respect, but may also constitute a first step in developing solutions to a number of other information dissemination problems. It can, for example, be used to derive appropriate selection criteria for the installation of static infrastructure at roadway cross-sections. Intersections close to safety critical institutions such as hospitals, police stations and fire stations need to have high probability of

message delivery. If this probability is shown priori using the derived formula that it is low, then static infrastructure needs to be installed.

The result can also be used to develop a traffic-aware routing protocol. The probability of successful information delivery along a straight line road has been derived analytically in the literature [125]. Combining the latter with the result obtained in this thesis, one can find the probability of successful information delivery along any route in a city map. Such a probability can be used as a cost measure for the specific route in a graph theoretic formulation of the routing problem. Solutions of the routing problem are readily available in the literature, and one can then use such solutions as a baseline to develop traffic aware routing protocols for VANETs. In this thesis we adopt the aforementioned approach and examine whether the information propagation probability on intersections derived in this work combined with the information propagation probability along a straight line road can yield a formula that gives the probability to deliver messages to a specific area in a certain amount of time and in a setting where no static infrastructure is used. We use a graph theoretical approach where we model the road map and the traffic characteristics using a directed weighted graph, which we refer to as the Road-Graph. We use the Road-Graph to calculate the probability of disseminating information along any given path and in a specific amount of time. So, for any two points on the Road-Graph we can find the path with the maximum probability of successful information delivery in the chosen time interval. This information is of great significance, since it can be utilized by routing protocols to optimally route packets in the vehicular network. We validate our analytical findings by comparing them with data obtained using VISSIM.

As indicated above, in this thesis we address several aspects of the information dissemination problem in VANETs and we propose a number of solutions based on probabilistic methods. Our work is significant in that it offers practical solutions to a number of open problems in VANETs,

but it also demonstrates the potential of these probabilistic methods in developing effective solutions for a variety of problems. Below we list our main contributions:

1. We develop a new hovering scheme which is successful in accomplishing its design objectives as it achieves high reachability, while at the same time alleviating the broadcast storm problem by reducing the number of exchanged messages. In order to allow message delivery in case of partitioned hovering areas, we allow message dissemination outside the hovering area. For alleviating the broadcast storm problem, the protocol employs epidemic routing within the hovering area and probabilistic flooding outside the hovering area. A unique feature of the protocol is that it is traffic-adaptive, in the sense that the rebroadcast probability is calculated based on estimates of the vehicle density, which are in turn based on estimates of the average number of neighbors.
2. We utilize a simple model of the roadway topology to develop a formula that estimates the average vehicle density in the hovering area based on measurements of the average number of neighbors readily available to each node using a beaconing method. The theoretical result is validated using simulations.
3. The methodology used to design the rebroadcast probability function is based on the existence of phase transition phenomena, which are typical when using probabilistic flooding outside hovering area. In this thesis we verify the existence of phase transition phenomena using simulations. When utilizing probabilistic flooding, we observe that, as the rebroadcast probability increases, the reachability remains low until a critical value is reached beyond which the reachability attains a relatively high value. We thus identify a phase of low reachability and a phase of high reachability. The probability at which the transition takes place depends on the traffic state in the area of interest.

4. For comparison purposes, in this thesis we design an alternative traffic adaptive hovering scheme, which attempts to alleviate the low reachability problem in cases of low traffic density, by extending the area in which epidemic routing is allowed beyond the hovering area. The vehicles outside the hovering that can participate in the message exchange process serve as information bridges towards the partitioned uninformed areas. The size of the extended hovering area is adaptive, in the sense that it changes based on estimates of the traffic density. In addition, the design procedure has been the same with the procedure adopted to design the proposed hovering scheme that applies probabilistic flooding outside the hovering area. Comparing the extended area approach and the probabilistic approach we observe very similar behavior. So, our contribution goes beyond the proposition of a new hovering scheme and includes the proposition of a new design methodology, which can yield a class of information hovering schemes with similar characteristics.
5. We present a novel MANET-oriented merging of wireless broadcasting and information hovering. The logic of the merge follows the fact that any wireless broadcast system is typically assigned bandwidth that is inversely proportional to its area of coverage. Thus, in order for the system to achieve adequate mean waiting times, a portion of the data dissemination task can be offloaded to a local hovering scheme. The number of items to be offloaded is defined heuristically. Two cooperation schemes are proposed, targeting solely mean waiting time (MWT) minimization and MWT-hovering bandwidth conservation balancing respectively. Simulation under realistic conditions yielded performance superior to stand-alone approaches, in all cases. It is worth noting that addressing the last problem using the proposed information hovering scheme highlights the generality of the proposed methodology.

6. We use a simple model of the roadway network to develop a lower bound on the probability to propagate information along an intersection. This lower bound depends heavily on the traffic state and is validated using simulations. The result is important as it can be used in the development of traffic aware routing protocols.
7. We formulate the routing problem in VANETs using a graph theoretic approach and for a specific route we find analytically the probability of successful information delivery in a specific amount of time. We use this probability to find the optimal route which maximizes the probability of information delivery. Our analytical findings are again validated using simulations. The simulations are conducted on the VISSIM simulator and the reference model constitutes a section of the freeway system in the Los Angeles Area.

1.5 Acknowledgments of Contributions to this Thesis

Several people have offered significant assistance in achieving the aforementioned scientific contributions, to whom I am grateful for their cooperation and guidance.

Apart from my thesis advisor, Dr. Andreas Pitsillides, Professor at the University of Cyprus, whose help is invaluable, these are in alphabetical order: Dr. Maria Andreou, Tutor and researcher at Open University of Cyprus, Dr. Marios Lestas, Assistant Professor in Frederick University, Christos Liaskos, Ph. D Candidate in Aristotle University of Thessaloniki, Andreas Manoli Graduate student of University of Cyprus, Dr. Georgios Papadimitriou Associate Professor in Aristotle University of Thessaloniki, Dr. George Papageorgiou researcher in University of Cyprus and Dr. Vasos Vassiliou, Lecturer at the University of Cyprus.

In particular, Andreas Pitsillides, my thesis advisor, Dr. Marios Lestas and Dr. Maria Andreou have motivated me to explore the exciting area of VANETs during my Phd studies and it is their

experienced guidance which has enabled me to develop the main findings of this work. They have introduced me to exciting new research directions and have significantly improved my analytical capabilities and my technical writing and presentation skills. Dr. Vasos Vassiliou has provided important advice during the first stages of my studies. Dr. George Papageorgiou has helped me in setting up scenarios on the VISSIM simulator and has provided advice on how to use it. I would also like to thank Andrea Manoli for helping me write scripts and run multiple scenarios on VISSIM. Finally, I would like to thank Dr. Georgios Papadimitriou and his PhD student Christos Liaskos for their help in applying the information hovering protocol in the considered Broadcast application. More specifically, Liaskos wrote the theoretical analysis on how to offload data from the push server and has also been responsible for setting up simulations which include the push server.

1.6 Publications

This section provides a list of all publications achieved during the work of this thesis.

1.6.1 List of publications related to this thesis

REFEREED JOURNALS

- A. Xeros, M. Lestas, M. Andreou and A. Pitsillides, "Adaptive Probabilistic Flooding for Information Hovering in VANETs", *IEEE Transactions on Intelligent transportation systems*, accepted for publication 2012, subject to a revision.
- C. Liaskos, A. Xeros, G. Papadimitriou, M. Lestas, and A. Pitsillides, "Balancing Wireless Data Broadcasting and Information Hovering for Efficient Information Dissemination", *IEEE Transactions on Broadcasting*, vol. 58, pp. 66-76, March 2012.

REFEREED CONFERENCES

- C. Liaskos, A. Xeros, G. Papadimitriou, and A. Pitsillides, "Information Hovering: A New Approach for Performance Acceleration of Wireless Push Systems", *16th IEEE Symposium on Computers and Communications (IEEE ISCC 2011)*, pp. 809-814, Kerkyra (Corfu), Greece, 2011.
- A. Xeros, M. Lestas, M. Andreou, and A. Pitsillides, "Adaptive Probabilistic Flooding for Information Hovering in VANETs", *IEEE Vehicular Networking Conference (VNC 2010)*, pp. 239-246, Jersey City, NJ, USA, 2010.
- A. Xeros, M. Andreou, A. Pitsillides and M. Lestas, "Information Hovering in Vehicular AdHoc Networks", *IEEE Workshop on Intelligent Vehicles and Infrastructures (NiVi'09)*, pp. 1-6, Honolulu, Hawaii, USA, 2009.
- A. Xeros, M. Andreou, M. Lestas, G. Papageorgiou, V. Vassiliou, and A. Pitsillides, "Time Constrained Message Delivery Probability in VANETs", *4th Workshop on Vehicle to Vehicle Communications (V2VCOM 2008)*, 2008.
- A. Xeros, M. Andreou, A. Pitsillides, and M. Lestas, "Information Propagation Probability on Intersections in VANETs", *Third International Workshop on Vehicle-to-Vehicle Communications 2007 (V2VCOM 2007)*, pp. 57-64, 2007.

BOOK CHAPTERS

- Y. Mylonas, M. Lestas, A. Xeros, M. Andreou, and A. Pitsillides, "Probabilistic Information Dissemination in Vehicular Ad Hoc Networks", *Book Chapter in Advances in Vehicular Ad-Hoc Networks: Developments and Challenges (edited by Prof. Mohamed K. Watfa)*, IGI Global, chapter 12, pp. 233-256, 2010.

1.6.2 List of publications beyond the work presented in this thesis

REFEREED JOURNALS

- C. Liaskos, A. Xeros, G. Papadimitriou, M. Lestas and A. Pitsillides, "Broadcast Scheduling with Multiple Concurrent Costs", *IEEE Transactions on Broadcasting (ISSN 0018-9316)*, Vol 58, Number 2, June 2012, pp. 178-186. (DOI 10.1109/TBC.2012.2191028)

REFEREED CONFERENCES

- C. Liaskos, A. Xeros, G. I. Papadimitriou, M. Lestas, and A. Pitsillides, "Broadcast Scheduling with Costs: A Novel Approach to Wireless Data Broadcasting", *In Proceedings of the 10th International Conference on Wired/Wireless Internet Communications (WWIC 2012)*, Santorini, Greece, 2012.

1.7 Thesis Overview

The rest of the thesis is organized in the following manner: In chapter 2 we introduce the VANETs area by discussing the evolution of the Intelligent Transportation Systems (ITS), then we discuss the different categories of the VANETs applications, and subsequently we show the ITS communication architecture. Thereafter we discuss data dissemination methods for VANETs, and present the related work in the area of data dissemination in VANETs. In chapter 3 we study the applicability of Information Hovering in VANETs and we present a scheme, which is traffic adaptive, achieving high reachability and low number of messages exchanged. In chapter 4 we present a new approach for performance acceleration of Wireless Push Systems using Information Hovering in VANETs, thus highlighting applicability of results in other communication areas. In chapter 5 we present our contribution in “Information Propagation Probability on Intersections in VANETs”. In this work we derive a formula which gives a lower bound on the probability of information dissemination on intersections where no static infrastructure is used. In chapter 6 we investigate the problem of information routing based on the probability to deliver messages to a specific area in a certain amount of time, given a number of paths derived using a specific directed weighted graph, which we refer to as the Road-Graph, as discussed later. Finally, in chapter 7 we offer our conclusions and future research directions.

Chapter 2

Background and Related Work

In this chapter, we present background on VANETs research area, with a historical overview of Intelligent Transportation Systems (ITS). We also discuss the main application categories that VANETs can support. Finally, we briefly discuss the related work in the area of data routing and dissemination in VANETs.

2.1 Intelligent Transportation Systems

In 1989 the Mobility 2000 group was formed in the USA and led to the formation of IVHS America (Intelligent Vehicle Highway Systems) in 1990, whose function was to act as a Federal Advisory Committee for the US Department of Transportation. In 1991 the Intermodal Surface Transportation Efficiency Act (ISTEA), of which the national IVHS program was defined as an integral part, became law in order to develop a national intermodal transport system that is economically sound, to provide the foundation for the nation to compete in the global economy, and to move people and goods in an energy-efficient manner. With the help of the Intelligent Transportation Society of America (ITSA), a nonprofit organization whose members come from academia and industry, a procedural framework was developed wherein IVHS services could be

systematically planned, defined and integrated. In 1994 the IVHS program was renamed as ITS (Intelligent Transportation Systems), indicating that, besides car traffic, modes of transportation receive attention as well.

On May 19, 1997, ITSA petitioned the Federal Communications Commission (FCC) to allocate a band of radio frequency spectrum to Dedicated Short-Range Communications (DSRC), 5.85-5.925 GHz, a total of 75 MHz of bandwidth that would serve a diverse array of ITS applications. During the years of 2003-2004, FCC adopted the recommendation of the ITSA to use a single standard for the physical and medium access control layers of the architecture. The recommendation was to adopt an architecture that was based on 802.11 [47]. After that, the IEEE task group, TGp, started modifying this standardization to support the needs of the transportation systems. The outcome of this work was named IEEE802.11p [46]. Furthermore, another team, WG 1609, has undertaken the task of developing specifications to append more layers in the protocol suite. The IEEE 1609 standard consists of four documents IEEE 1609.1 [42], IEEE 1609.2 [43], IEEE 1609.3 [44], and IEEE 1609.4 [45]. Since both standards, the IEEE802.11p and IEEE1609.x, are used to facilitate the provision of wireless access in vehicular environment, the acronym WAVE is used to refer to both. In Europe, the Architecture Task Force group worked closely with Car2Car Communication Consortium [10] and other standardization bodies such as the Internet Engineering Task Force (IETF) [48] and the International Standards Organization (ISO) [50] to prepare a European standard to be submitted to the European Telecommunication Standards Institute (ETSI) [20]. The first public version of the document named “The European ITS Communication Architecture” was published in October of 2008. It entails an overview of the basics of the ITS architecture in Europe that was applied for a new research project named PREparation for DRIVING implementation and Evaluation (PRE-DRIVE C2X) [95].

2.2 Application Categories in VANETs

Motivated by the great need to reduce the constantly increasing number of traffic accidents, the majority of applications proposed in VANETs are designed to improve active safety in driving. However, messages exchanged between vehicles can serve other purposes as well, such as improving driving and entertainment. In this paragraph we present a representative collection of envisioned applications and we categorize them following the suggestions of [104]. The applications presented in Table 1 are compiled from several sources. A large collection of applications was gathered in a report in [118] by the Vehicle Safety Communications (VSC) project. A deeper description of virtual warning sign applications is given in [18]. Additionally, most publications in VANETs also contain examples of applications. The chosen classification scheme groups applications by their purpose, which leads to groups of logically similar applications.

A. Active Safety

Active safety applications are considered the typical and most desirable group of applications for VANETs with direct impact on road safety. The basic intention is to make driving safer by inter-vehicle communication. The benefits of this kind of communication include the drivers being warned about a dangerous situation and the vehicle avoiding an accident or reacting appropriately if the accident cannot be avoided. We categorize active safety applications according to the level of danger. Dangerous road features like curves are static and thus foreseeable, decreasing the danger level. Abnormal traffic and road conditions are still almost static, but have a dynamic notion, i.e. they differ from the expectation of drivers that regularly pass the event location. In these cases danger is elevated. Danger is high when applications are used to prevent collisions, e.g. if a vehicle brakes heavily in dense traffic. If this does not help any more, such as in the case of imminent danger, when a collision

	Situation/Purpose	Application Examples
A. Active Safety	1. Dangerous road features	1. Curve speed warning, 2. Low bridge warning, 3. Warning about violated traffic lights or stop signals
	2. Abnormal traffic and road conditions	1. Vehicle-based road condition warning, 2. Infrastructure-based road condition warning, 3. Visibility enhancer, 4. Work zone warning
	3. Danger of collision	2. Blind spot warning, 2. Lane change warning, 3. Intersection collision warning, 4. Forward/Rear collision warning, 5. Emergency electronic brake lights, 6. Rail collision warning, 7. Warning about pedestrians crossing
	4. Crash imminent	1. Pre-crash sensing
	5. Incident occurred	1. Post-crash warning, 2. Breakdown warning, 3. SOS service
B. Public Service	1. Emergency response	1. Approaching emergency vehicle warning, 2. Emergency vehicle signal preemption, 3. Emergency vehicle at scene warning
	2. Support for authorities	1. Electronic license plate, 2. Electronic drivers license, 3. Vehicle safety inspection, 4. Stolen vehicles tracking
C. Improved driving	1. Enhanced Driving	1. Highway merge assistant, 2. Left turn assistant, 3. Cooperative adaptive cruise control, 4. Cooperative glare reduction, 5. In-vehicle signage, 6. Adaptive drivetrain management
	2. Traffic Efficiency	1. Notification of crash or road surface conditions to a traffic iteration center, 2. Intelligent traffic flow control, 3. Enhance route guidance and navigation, 4. Map download/update, 5. Parking spot locator service
D. Business/ Entertainment	1. Vehicle Maintenance	1. Wireless diagnostics, 2. Software update/flashing, 3. Safety recall notice, 4. Just-in-time repair notification
	2. Mobile Services	1. Internet service provisioning, 2. Instant Messaging, 3. Point-of-interest notification
	3. Enterprise solutions	1. Fleet management, 2. Rental car processing, 3. Area access control, 4. Hazardous material cargo tracking
	4. E-Payment	1. Toll collections, 2. Parking Payment 3. Gas payment

Table 1: Overview of applications in VANETs

cannot be avoided any more, precrash-sensing will prepare the vehicle in order to minimize the impact of the impending crash, e.g. by closing windows or raising dampers. Finally, when danger is turned to an incident, it is important to warn approaching vehicles or call for help.

B. Public Service

Vehicular networks are also intended to support the work of public service such as, police or emergency recovery units. Prominent examples of this category are the support of emergency vehicles by virtual sirens or signal preemption capabilities. Using these applications, emergency vehicles may be able to reach their destination much faster than today. In addition, traffic surveillance could be simplified by applications such as electronic license plates. However, such an application must not be abused by anyone, which clearly underlines security requirements and the need for a discussion of the legal aspects of vehicular communication.

C. Improved Driving

This category contains applications which are intended to improve or to simplify driving by means of communication. The idea is comprised of microscopic scenarios in the immediate surrounding of a vehicle, as well as macroscopic optimization of traffic efficiency. In the first case, helper applications are intended to assist the driver in standard traffic situations, such as when entering a motorway and merging into the flowing traffic or the cooperative reduction of glare due to upper beam headlights. In the second case, traffic efficiency in a greater area is targeted. This can mean that an accident warning is disseminated in a larger area to inform vehicles about the potential obstacle, so that drivers can take a different

route. Another service is the dissemination of parking information or even the reservation of a parking space.

D. Mobile Business and Entertainment

A large block of applications can be embraced under the terms business and entertainment. Here, the focus is on delivering services to customers, automation of vehicle related tasks or payment applications, like download of music, fleet management, chatting, e-advertisement, simpler vehicle maintenance or payment of parking or road usage.

2.3 ITS Communication Architecture

In this paragraph we give an overview of the ITS communication architecture and its main components, which can communicate with each other using different communication technologies. The ITS communication architecture is a communication system designed for ITS and made of four separated subsystem components:

1. Vehicle Station. The vehicle subsystem component connected to vehicle domain architecture designed by vehicle manufacturer via the vehicle mobile gateway
2. Personal Station. The mobile subsystem component
3. Roadside Station. The roadside subsystem component
4. Central Station. The central subsystem component

All four components are described in figure 2 and each component separately in figures 3 to 6. Each component contains an ITS station, which basically consists of one or more Application Units (AU) and a Communication & Control Unit (CCU). The ITS station may be interconnected to additional components. These sub-components may be implemented in a single node, or in

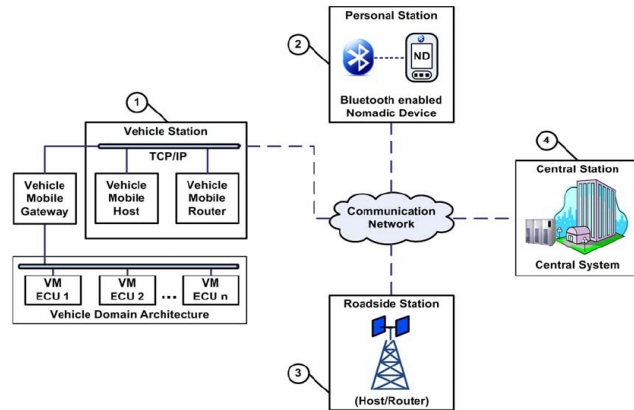


Figure 2: ITS Communication Architecture illustrates the communication system designed for ITS [110]. 1. Vehicle Station, 2. Personal Station, 3. Roadside Station, 4. Central Station.

separate nodes interconnected by an internal network. It may be also considered to have separate CCUs, e.g. for safety and non-safety related communication.

2.3.1 ITS Vehicle Station

The ITS Vehicle Station, shown in figure 3, is mounted within or fixed to a vehicle and enables the vehicle to participate in cooperative ITS applications. With respect to ITS communication it is irrelevant whether the vehicle station is fully integrated into a vehicle or whether it is a separate mobile component that can be connected to the vehicle.

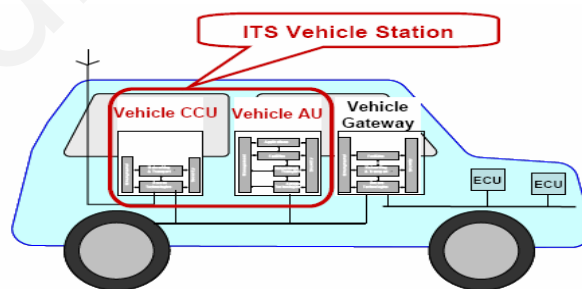


Figure 3: ITS Vehicle Station

An ITS Vehicle Station consists of two sub-components:

- A Vehicle CCU is in charge of communication with other vehicles (V2V) or with roadside infrastructure components (Vehicle to Roadside Unit - V2R). Depending on the profile, the Vehicle CCU may also provide access to the Internet Domain, e.g. by integrated UMTS (Universal Mobile Telecommunications System) communication hardware.
- One or several Vehicle AUs realising the ITS applications.

This way, the ITS Vehicle Station runs the cooperative ITS applications and is able to exchange information with other vehicles or with roadside infrastructure components. In order to provide vehicle specific ITS applications, the ITS Vehicle Station requires access to the vehicle “infrastructure” (e.g. data from ECUs (Electronic Control Units) or vehicle sensors, which are interconnected via a vehicle network (e.g., Controller Area Network bus). The link to the vehicle-specific network is provided by the vehicle gateway, which is able to access the vehicle information and provide it to the ITS Vehicle Station. In the reverse direction, the Vehicle Gateway may also provide data to vehicle controllers or to presentation devices integrated in the vehicle. As part of its role, the Vehicle Gateway may also include security and translations services.

2.3.2 ITS Roadside Station

An ITS Roadside Station, shown in figure 4, is a fixed installation along the road. From the communication component view, it consists of two sub-components:

- A Roadside CCU is responsible for the communication with ITS Vehicle Stations or other ITS Roadside Stations.
- One or several Roadside Application Units realise the ITS applications for the roadside equipment.

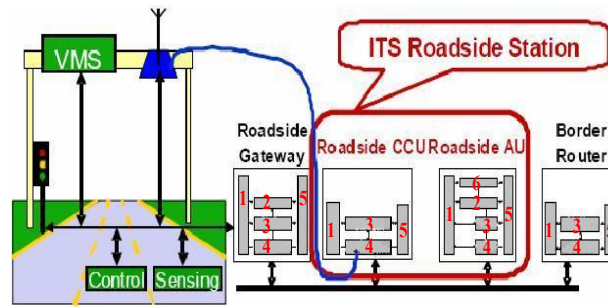


Figure 4: ITS Roadside Station (1.Management, 2.Facilities, 3.Networking & Transport, 4.Access Technologies, 5.Security, 6.Applications)

The ITS Roadside Station may be linked via a roadside gateway to road sensors or traffic control units, e.g. traffic lights or variable traffic signs and/or its controller. This enables the realization of ITS applications like in-vehicle signage, which provides the information of Variable Message Signs (VMS) to vehicles, as well as dynamically adapting the information displayed in the VMS, depending on the information transmitted by the vehicles.

Optionally, a border router connects the ITS Roadside Station to a backbone network, which can be an ITS roadside infrastructure network or the Internet Domain. Connection to the ITS roadside infrastructure network enables the direct exchange of information between different ITS Roadside Stations, whereas a connection to the Internet enables communication with hosts in the Internet. This connection can be used, e.g., to connect the ITS Roadside Station to a central traffic control server.

2.3.3 ITS Central Station

From a communication component view, the central system can realize for example business applications and processes. It therefore may provide information that can be used to realize cooperative ITS applications. The ITS applications themselves are implemented in the ITS Central Station (shown in figure 5), which basically comprises one or more application units.

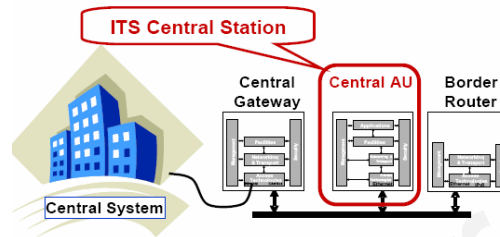


Figure 5: ITS Central Station

The available information of a central system is provided to the ITS Central Station via a Central Gateway, which ensures that the information of the central system can be accessed and used by the ITS Central Station. The ITS Central Station is typically connected to the Internet Domain using a border router. This way, the ITS Central Station is able to provide information useful for cooperative ITS applications to the vehicles, e.g. by sending them, via UMTS, to the vehicles, or, in case of traffic flow optimizations, by sending them via the ITS Roadside Stations to the vehicles.

2.3.4 ITS Personal Station

In the communication component point of view, the ITS Personal Station is a consumer electronics device, which is typically assigned to a person, e.g. pedestrian or cyclist. In case the consumer electronics device is linked to a vehicle, it is considered as ITS Vehicle Station. The ITS Personal Station, as shown in figure 6, typically consists of two components:

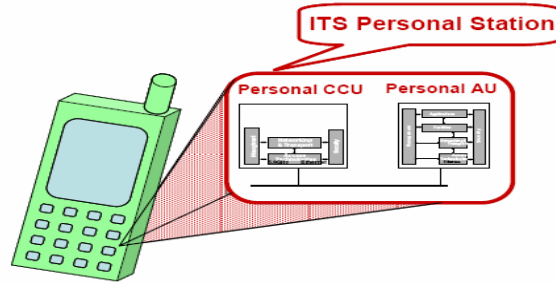


Figure 6: ITS Personal Station

- Personal CCU, which realizes the communication to participate within the ITS ad hoc network and to get access to the core network, e.g. via UMTS.
- One or more application units to run ITS applications

2.4 Wireless Access in Vehicular Environment (WAVE)

In this section we present a brief overview of the IEEE WAVE system architecture [109] as an indication of the current state of standardization process. WAVE system architecture is a set of standards that describes the communication stack of vehicular nodes and the physical airlink between them as shown in figure 7.

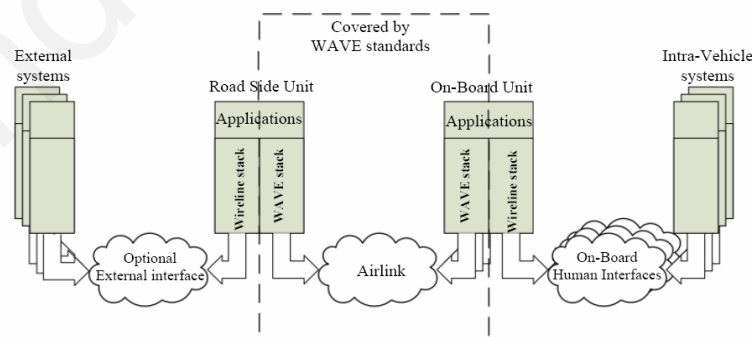


Figure 7: Wave System Components

	Name	Description
802.11p	Wireless Access in Vehicular Environments (WAVE)	An amendment to the well-known IEEE 802.11 Wireless LAN Standard and covers the physical layer of the system.
1609.1	Resource Manager	Covers optional recommendations for the application Layer.
1609.2	Security Services for Applications and Management Messages	Covers security, secure message formatting, processing, and exchange.
1609.3	Networking Services	Covers the WAVE communication stack.
1609.4	Multi-Channel Operation	Covers the arrangement of multiple channels and how they should be used.

Table 2: WAVE Standard

A road Side Unit may have two interfaces, one for the wireless WAVE stack and the other for external interfaces like wireline Ethernet that may be used to enable connectivity to the Internet. Similarly, each On Board Unit may have two interfaces, one for the wireless WAVE stack and the other for sensor-connections and human interaction.

WAVE standard consists of five complementary parts 802.11p [46], 1609.1 [42], 1609.2 [43], 1609.3 [44] and 1609.4 [45] as shown in Table 2.

The WAVE communication stack and the coordination between standards are shown in figure 8.

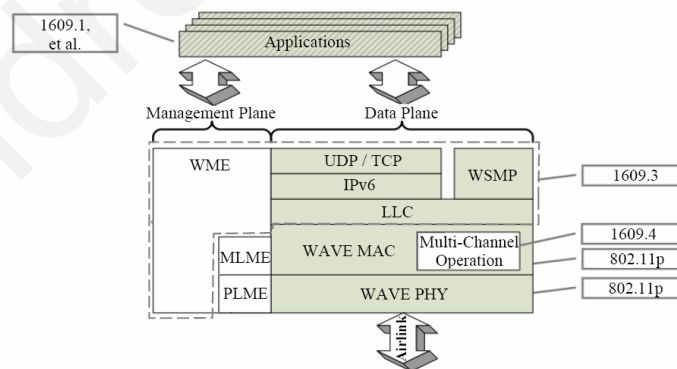


Figure 8: Wave Protocol Stack

2.4.1 WAVE Physical Layer

The WAVE spectrum is composed of seven channels of 10 MHz each in the 75MHz bandwidth of DSRC applications, as shown in figure 9, with an option of grouping two adjacent channels to have a spectrum of 20 MHz. Channel 178 is the only control channel (CCH), and other channels are service channels (SCH). Channels 175 and 181 are the 20 MHz channels. The channel number (CN) is derived by counting the number of 5-MHz spectrum in the frequency band from 5000 MHz to the center frequency $f(\text{CN})$ of the channel CN, i.e.

$$\text{Channel center frequency} = 5 \text{ GHz} + (5 \text{ channel number}) \text{ MHz}$$

For example, the Channel number of the frequency 5875 MHz to 5885 MHz (which has as channel center frequency 5880 MHz) is given by:

$$\text{Channel Number} = (5880 - 5000)/5 = 176$$

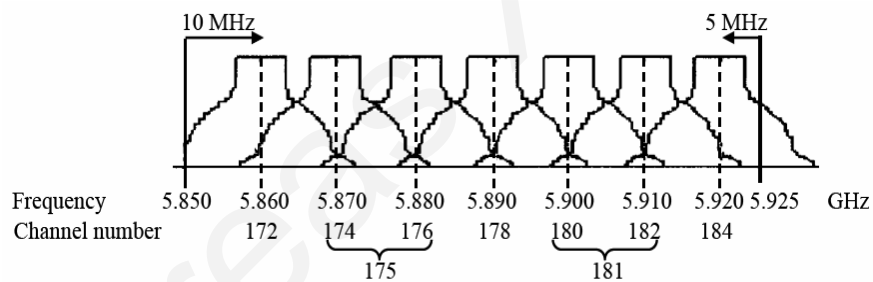


Figure 9: Spectrum of Wave Channel

The modulation scheme used by WAVE is the Orthogonal Frequency Division Multiplexing (OFDM) using 52 orthogonal subcarriers. The OFDM is a multi-carrier modulation scheme where data is split into multiple lower rate streams. Each stream is used to modulate one of the closely spaced orthogonal subcarriers. The primary advantage of OFDM is its ability to cope with frequency-selective fading due to multipath channels without complex equalization filters. This modulation scheme enables data rates of 3, 4.5, 6, 9, 12, 18, 24, and 27 Mbit/s in the 10 MHz channels and up to 54 Mbit/s in the 20 MHz channels. The orthogonal subcarriers should

be modulated using BPSK (Binary Phase Shift Keying), QPSK (Quadrature Phase-Shift Keying), 16-QAM (Quadrature Amplitude Modulation), or 64-QAM depending on the data rate required.

2.4.2 Wave Channel Coordination

The WAVE spectrum is composed of only one control channel (CCH) and six service channels (SCHs). The control channel is considered as the public room for all WAVE devices and it is a critical resource. Efficient organization and minimization of traffic on the CCH is a challenging problem. The CCH should only be used for service advertisement frames and broadcast messages (i.e. when the transmitter has not negotiated with a specific receiver yet). However, no active connections between two or more devices are allowed to exchange data over the CCH (i.e. after handshaking, the transmitter and the receiver must pursue talking in another channel).

The other six SCHs are considered as private rooms for any connection to exchange long streams of data. Before initiating a connection over a SCH, a node must first join an active logical private network (namely, the WAVE Basic Service Set, WBSS which is a set of cooperative WAVE stations). Advertisement of new services should be transmitted over the CCH, though actual data exchange of the service is done over any SCH.

2.4.3 WAVE Communication Protocol

WAVE supports two protocol stacks, the standard Internet Protocol Version 6 (IPv6) and a new specially designed WAVE Short Message Protocol (WSMP).

- **Internet Protocol Version 6 (IPv6).** WAVE networking services support data exchange using the Internet Protocol version 6 (IPv6) [49] with both TCP and UDP at the transport

layer. The existence of IPv6 protocol in the wireless device within vehicles opens the Internet access with a tremendous variety of possible applications. Connection using IPv6 is permitted only on SCHs after joining a WBSS.

- WAVE Short Message Protocol (WSMP).** The WAVE Short Message Protocol (WSMP) is a new protocol designed especially for an optimized operation in WAVE environments. If any node prefers not to join a WBSS (for example, a transmitter has a short data to broadcast) it will have to use only WSMP over the CCH. WSMP is used for direct transmission of short messages without joining WBSS. Messages of this protocol are designed to consume minimal channel capacity. Hence, it is the only protocol allowed over the CCH (and may be used on any SCH as well). The suggested frame format of a WAVE Short Message (WSM) is shown in figure 10 (lengths are in octets of bits).

1	1	1	1	1	4	2	variable
WSM Version	Security Type	Channel Number	Data Rate	Tx Power Level	Provider Service Identifier	WSM Length	WSM Data

Figure 10: WSM Frame Format

The “WSM Version” indicates the used version of WSMP (currently, its value is zero). The Security Type indicates the security processing of the WSM Data i.e. the transmitter application can sign or encrypt the message with an indication in security field. The Channel Number is used to identify the radio channel used for the WSM. The Data Rate indicates the data rate used for the WSM. The Tx Power Level indicates the transmit power used for the WSM. The Provider Service Identifier identifies the application that originated the WSM (each application will have a unique number). The WSM Length indicates the length in octets of the following WSM Data

field (limited to 1400 in its default value). The WSM Data contains the application data being transferred.

2.4.4 WAVE Management Plane

The WAVE management plane is considered a logical low-level database of the system and performs system configuration and maintenance functions. It consists of the WAVE Management Entity (WME) with a special part to serve the MAC layer, namely MAC Layer Management Entity (MLME), and another one to serve the physical layer, namely Physical Layer Management Entity (PLME). Examples of its use include:

- Prior to the first operation of the transceiver (i.e. network configuration phase) different system parameters are loaded into the device's WME. This field is known as "Local Information".
- Active applications register their parameters with the WME. Therefore, MAC layer can determine whether a received WBSS advertisement is of interest to any of its applications or not. This field is known as "User Service Information".
- The WME is responsible for generating the WAVE service advertisement frame on an application request. This field is known as "Provider Service Information".
- On the initiation or joining of a WBSS, network parameters are registered in the WME.

2.5 Related Work

Having discussed issues related to the IEEE WAVE system architecture which constitutes the latest set of standards for VANETs deployment, we next make a comprehensive overview of work that has appeared in the literature and is pertinent to the issues dealt with in this thesis. This thesis

addresses several information dissemination problems in VANETs and proposes novel solutions based on probabilistic methods, since the system is inherently stochastic in nature. The design methods used lead to solutions which are universal, in the sense that they can be applied to areas beyond the ones considered here. Such extensions have been investigated with emphasis on a Push based Broadcast System.

Data dissemination schemes in VANETs can have fundamentally different characteristics and design objectives. Consider, for example, a cooperative warning application which, upon detection of an unexpected event, such as a traffic accident, causes vehicles to generate warning messages with critical information regarding the unexpected event. A short range multi-hop broadcast scheme may be used to disseminate the emergency message to all neighboring vehicles, warning them of the imminent danger. A routing protocol can also be employed to facilitate the transfer of the warning message over long distances to emergency services such as hospitals and police stations. Finally, a hovering scheme can be used to maintain the message in a confined geographical area for a specific period of time, notifying all vehicles entering the area of the unexpected event. The applicability of the proposed data dissemination schemes goes beyond the safety related example mentioned above and encompasses other areas, as for example, push based broadcasting and mobile cloud. This universality property of the proposed scheme will be discussed later.

Below we review solutions which have appeared in the literature for each of the class of algorithms mentioned above. In addition, since in this thesis we explore the integration of the proposed information hovering scheme in a push based wireless broadcasting system, we also present a review of schemes pertinent to wireless broadcasting.

The methods proposed in this thesis are applicable to both safety and non-safety applications, although more emphasis is given to safety applications. So, our discussion is made with reference to both types of applications. Safety applications involve the exchange of messages which notify

vehicles of potential driving hazards and help prevent collisions. Some techniques of safety application can be found in [1, 6, 21, 22, 32, 56, 74]. Non-safety applications involve the exchange of messages which usually relate to accurate traffic monitoring, distributed passenger teleconferencing, music downloading and roadside e-advertisements. Some techniques of non-safety applications are presented in [16, 57, 112].

2.5.1 Information Hovering

The term Information Hovering was first introduced in [121], and formally defined later in [122]. A more detailed description is provided in [101]. The concept involves decoupling of the hovering information from its host and promotes coupling it directly with a specific geographical location, which is called the anchor location. In this regard, the hovering information stays “attached” to a specific geographical point, called the anchor location and to its vicinity area, called the anchor area. The information hovers from one mobile device to another, in a quest to remain within a specific vicinity and avail itself to users currently present or entering its anchoring geographical location. The Information Hovering concept can be used in a variety of applications in Mobile Adhoc Networks, with characteristic usage examples provided in [58]. A relevant concept in the networks literature is Geocasting, which addresses the problem of information dissemination in geographical target regions. Geocast protocols which have appeared in the literature [78] are intended to forward the relevant message in the target region and distribute the message once, to all devices residing in the region. However, this one time message delivery is not relevant in the information hovering paradigm, as the information needs to remain in the target area for a specific amount of time. This concept has appeared in the literature as abiding or time-stable geocasting [79]. Flooding based schemes, such as epidemic routing [116], can be applied to the entire network to effectively serve such a paradigm, achieving high percentage of nodes in the hovering area

receiving the relevant message (high reachability). However, this is done at the expense of a large number of redundant messages, which strain the communication channel and lead to extensive contention and high latency of message delivery. In cases of high traffic density, this is known as the broadcast storm problem, and a number of techniques have been proposed in literature to alleviate this problem [3, 9, 11, 25, 40, 59, 73, 103]. On the other hand, in areas with low vehicle density, dynamic network partitioning may be present. In this case, the message may be lost, as there is no path to bring the message back to the hovering area. One approach to address this challenge is to allow controlled exchange of messages outside the hovering area. The reasoning behind this approach is that informed vehicles outside the hovering area can serve as information bridges towards partitioned uninformed areas, thus increasing message reachability. In addition, since “epidemic routing” outside the hovering area is avoided, the number of exchanged messages is reduced. The authors in [37] follow this reasoning to allow epidemic dissemination of messages in an extended area beyond the hovering area. They demonstrate through simulations that such an approach can increase the recorded reachability in cases of low traffic density.

In this thesis we apply blind flooding within the hovering area and probabilistic flooding outside the hovering area. We use complex probability functions, which can be adaptively regulated based on estimates of the vehicle density within the hovering area. This increase in complexity is shown to be highly beneficial in terms of the high reachability and low number of messages achieved.

Note that inside the hovering zone where we apply blind flooding a number of solutions have been considered to eliminate broadcast storm problem [3, 9, 11, 25, 40, 59, 73, 103]. As this was extensively covered in the literature, in this thesis we concentrate on the design of the probabilistic flooding scheme outside the hovering zone. Inside the hovering zone we assume a simple flooding scheme.

2.5.2 Push Server - Data Broadcasting

Recent years have witnessed the wide-spread use of wireless push-based broadcasting where the clients continuously monitor a broadcast process from the server and retrieve their required data items. Simple in architecture and implementation, lightweight and energy efficient - especially from the client's point of view, both hardware and software - the push-based approach has been adopted for use in a variety of information dissemination applications and has been incorporated in almost every single mobile telecommunication device. Popular uses include airport and hospital information systems, instant messaging services, and multimedia on demand over the internet or cellular networks. Consequently, the growing interest of the telecommunications industry in these systems has spurred their research on the performance optimization.

Research on optimal data broadcast scheduling initially focused on the minimization of the clients' Mean Waiting Time (MWT). [34] presented the lower bound of the MWT, showing that the knowledge of the data items' request probabilities and sizes is sufficient for achieving it. A scheduling algorithm was also defined, which achieved the minimum MWT at the cost of inhibitory computational complexity. [102] introduced a "reservation" system that favors client waiting time in the case of big sized items. In light of the high complexity of the initial scheduling algorithms, several approaches have been proposed to simplify the data broadcasting problem, the most influential being the Broadcast Disks model [2]. A great deal of work has been done based on this model, studying data pre-fetching, caching and indexing, hybrid data broadcasting, as well as scheduling strategies and noise interference. An overview of all these concepts can be found in [131]. More recent works have been focused on the minimization of psychometrics [54] instead of the mean waiting time, in an effort to provide QoS. Authors in [87] have introduced adaptivity for wireless push systems with locality of demand by employing a learning automaton to process

the incoming client feedback. More efficient feedback schemes for dynamic environments were formulated in [67], employing the principle of maximum entropy to speed up convergence. Finally, broadcast scheduling algorithms with minimal complexity and optimal performance were presented in [68, 69]. It is clear from the above that previous state-of-the-art works have treated user networking simply as a means of extending the coverage of wireless access points [134]. In contrast, our work in [72] proposes parallel data dissemination through centralized, optimal broadcasting and through user networking (information hovering) at the same time. Data dissemination load can be freely shared between the centralized broadcasting and the hovering, following a pull-push balancing model. This data dissemination off-loading, offered by the information hovering scheme, is shown to greatly improve performance.

2.5.3 Information Routing

Routing has been studied thoroughly in VANETs. However, none has addressed the highly stochastic nature of the traffic system to design a routing scheme that maximizes the probability of successful message delivery in a specific amount of time. Therefore, below we present the main approaches adopted in literature to develop routing protocols for VANETs [7]. In [23] authors examine the applicability of existing ad hoc routing protocols to VANETs. Specifically, they compare Dynamic Source Routing (DSR) with the Greedy Perimeter Stateless Routing (GPSR) protocol. They conclude that, when communication sessions are comprised of more than 2 or 3 hops, position-based ad hoc routing is preferable over reactive non-position-based approaches.

Contention-Based Forwarding (CBF) [24] is a geographic routing protocol that does not require proactive transmission of beacon messages. Data packets are broadcast to all direct neighbors, who then, decide if they should forward the packet. The actual forwarder is selected by a distributed timer-based contention process, which allows the most-suitable node to forward the

packet and to suppress other potential forwarders. Receivers of the broadcast data would compare their distance to the destination to the last hop's distance to the destination. TOPOLOGY-ASSISTED GEO-OPPORTUNISTIC ROUTING (TO-GO) [62] is a geographic routing protocol that exploits topology knowledge acquired via 2-hop beaconing to select the best target forwarder and incorporates opportunistic forwarding with the best chance to reach it. It is different from CBF in three main aspects. Firstly, rather than picking the next forwarding node that makes the best progress to the destination, it picks the next forwarding node that makes the best progress to a target node. A target node is defined to be the node that the greedy algorithm or the recovery algorithm would normally pick. At the junction the choice of the target node either beyond the junction or at the junction is based upon whether the routing is in greedy mode or recovery mode. The reason for choosing the target node instead of the destination as the frame of reference is to take care of the city topology, where roads intersect and destination usually does not lie on the same street as the source as in the highway. Packets have to make multiple turns into different streets before arriving at the destination. The data is then broadcast to all direct neighbors. Whosever distance is closer to the target node gets picked to be the next forwarding node.

The second difference is that, unlike CBF, there is still the need of beacons, which are used for nodes to pick the target node. The fact that the data is broadcast and only the node that makes the furthest progress toward the target is chosen is to account for wireless channel errors and low packet delivery rate arising from multi-path fading, shadowing, and mobility. These problems makes the furthest node (the target node) which usually does not receive the data packet. Packets are therefore opportunistically making their best progress toward the target node and thus the destination. TO-GO uses a novel way to choose the forwarding set of nodes that are candidates for the next forwarding node. The set is chosen so that all nodes can hear one another (no hidden

terminals) and make a progress toward the target node. Lastly, TO-GO differs from CBF in providing routing decision for recovery. CBF on the highway works because the destination is always straight ahead. Thus, local maximum never occurs on the highway and the selection of the next forwarding node is always one. That is closest to the destination. However, in city environments, streets cross each other and destination does not lie on the same street as the source. Thus, local maximum frequently occurs. TO-GO adapts the concept of CBF that packets are opportunistically sent to the target node, calculated by the routing decision in both the greedy and recovery mode.

The Mobility-Centric Data Dissemination Algorithm (MDDV) [124] is one of the few that provide a complete architecture for vehicular routing. It combines the ideas of opportunistic forwarding, trajectory-based forwarding and geographical forwarding. The protocol disseminates data to intended receivers, while maintaining some design demands, e.g. high delivery ratio, low delay and low memory occupancy. The protocol can also be applied to hybrid architectures.

The Geographic Source Routing (GSR) protocol proposed in [75] examines the problems appearing with base-line position-based routing in two-dimensional urban scenarios. GSR combines position-based routing with topological information. The adoption of the Reactive Location Service (RLS) [55] system is assumed. The source uses flooding to request the position of a node identifier. After discovering the location of the recipient, the source uses a digital map of the roads to calculate the set of junctions that the packet will follow. In [64] authors present A-STAR, an Anchor based Street and Traffic Aware Routing scheme. They use information on city bus routes to identify an anchor path with high connectivity for packet delivery. The model is designed based on position-based routing in order to facilitate VANETs in urban areas. In such environments, vehicle density is larger in some well known -for their traffic- roads than in others. Connectivity in such roads can be higher and more stable due to regular bus passes. A-STAR constructs a graph, based on how many bus lines go through certain roads. Since each vehicle may be aware of the

bus route information through digital maps, an anchor route may be constructed using the Dijkstra's algorithm for the least weight. GyTAR [52] is an improved Greedy Traffic Aware Routing Protocol for VANETs in City Environments. It considers road traffic variations, vehicles' speeds, directions and multilanes. It improves Greedy Traffic aware routing protocol. Chisalita and Shahmehri in [12] propose a distributed protocol for decentralized network organization. The protocol requires the receivers to analyze the exchanged messages so as to figure out if they are the intended destinations. For this filtering the current traffic conditions are taken into account. The protocol includes mechanisms for message acceptance/denial, local maintenance of neighborhood information and transmission of basic safety - as well as nonsafety - messages. In [83] authors present a formal model of data dissemination in VANETs and study how the characteristics of Vehicular network, specifically the bidirectional mobility on well defined paths, affect the performance of data dissemination. Information can be disseminated using vehicles moving on the same direction, vehicles moving in the opposite direction, or vehicles moving in both directions.

Despite the various approaches which, as mentioned above, have been adopted in the literature, no scheme addresses the highly stochastic nature of the traffic system to design a routing scheme that maximizes the probability of successful message delivery in a specific amount of time. The stochastic nature of the system suggests that such an approach may lead to solutions with improved properties. However, calculating the probability of successful message delivery is not an easy task if one not only takes into account line roads but intersections as well. In this thesis we address the problem of calculating the probability of successful message delivery along an intersection and we demonstrate how such calculations can be used as a baseline for designing an effective routing protocol with improved properties.

Chapter 3

Information Hovering in Vehicular AdHoc Networks

3.1 Introduction

Information Hovering is a new concept of information dissemination over a mobile set of peers. The Information Hovering concept can be used in a variety of applications in Mobile Adhoc Networks, with characteristic usage examples provided in [58]. A relevant concept in the networks literature is Geocasting, which addresses the problem of information dissemination in geographical target regions. Geocast protocols which have appeared in the literature [78] are intended to forward the relevant message in the target region and distribute the message once, to all devices residing in the region. However, this one time message delivery is not relevant in the information hovering paradigm, as the information needs to remain in the target area for a specific amount of time. This concept has appeared in the literature as abiding or time-stable geocasting [79].

Information Hovering can naturally be applied to many applications in Vehicular Ad Hoc Networks, where useful information needs to be made available to all vehicles within a confined geographical area for a specific time interval. A straightforward approach is to have all vehicles within the hovering area exchange messages with each other. This approach has two potential

problems depending on the vehicle density within the hovering area. In cases of high traffic density the well known broadcast storm problem [114] can lead to a significant increase in the latency of message delivery. Several methods have been proposed in literature to alleviate this problem [3, 9, 11, 25, 40, 59, 73, 103] and is thus not further addressed in this work. The second problem of applying epidemic routing within the hovering area only is that, in cases of low traffic density and/or low market penetration rate, this method does not guarantee that all vehicles within the hovering area will receive the message due to potential partitioning of the network. To alleviate this problem in our work, we propose a scheme that is based on the application of epidemic routing within the hovering area and on the probabilistic flooding outside the hovering area [128]. Probabilistic flooding involves vehicles, upon receiving a message for the first time rebroadcasting the message with probability p . Informed vehicles outside the area can serve as information bridges towards partitioned uninformed areas, thus leading to high reachability, which is defined as the percentage of vehicles in the hovering area that have received the information. We demonstrate through simulations conducted using VISSIM [96] and a purpose build C++ application that our approach outperforms the scheme proposed in [37] by significantly decreasing the number of messages required in order to achieve high reachability.

Note that inside the hovering zone where we apply blind flooding a number of solutions have been considered to eliminate broadcast storm problem [3, 9, 11, 25, 40, 59, 73, 103]. As this was extensively covered in the literature, in this thesis we concentrate on the design of the probabilistic flooding scheme outside the hovering zone, and inside the hovering zone we assume a simple flooding scheme.

Another solution is proposed in [37] where authors follow this reasoning to allow epidemic dissemination of messages in an extended area beyond the hovering area. They demonstrate through simulations that such an approach can increase the recorded reachability in cases of low traffic

density. The main problem with the work in [37] is that authors do not discuss the issue of the traffic density and how mobile nodes can adapt to the changes of the traffic conditions.

The design methodology we adopt has been simulative, and a major challenge in the overall procedure has been the design of the rebroadcast probability function. The probability function maps local variables at each vehicle to the rebroadcast probability. Among a number of candidate functions, we choose the one that yields superior performance and we tune its parameters taking advantage of phase transition phenomena, which are typical in probabilistic flooding schemes.

A unique feature of the proposed protocol is that it is adaptive, in the sense that the rebroadcast probability outside the hovering area is adaptively regulated based on estimates of the vehicle density within the hovering area. Estimates of the vehicle density within the hovering area are obtained using measurements of the number of neighbors of each vehicle. The analytical formula relating the two quantities is derived using a simple model of the transportation network within the hovering area. We demonstrate through simulations that the proposed scheme is successful in satisfying the design objectives and outperforms other candidate hovering protocols, such as epidemic routing in the entire network, epidemic routing in the hovering area only, and the scheme proposed in [37]. We also demonstrate that if we utilize the same design methodology to render the protocol proposed in [37] adaptive, its performance becomes comparable to the performance of the scheme proposed in this work. Thus, our contribution goes beyond the proposal of a specific information hovering protocol and extends to the introduction of a design procedure that can be used to design a class of density adaptive hovering protocols.

3.1.1 Phase Transition, Percolation Theory and Random Graphs

As mentioned above, the design of the probability function of the information hovering protocol takes into account phase transition phenomena typical in schemes that employ probabilistic

flooding. Phase transition or phase change is a phenomenon where a system undergoes a sudden change of state or behavior in response to a small change in a chosen parameter set [19]. The phenomenon is of great importance, as it has been observed and analyzed in a number of systems pervading our physical and engineering world. Two areas of research where phase transition has been extensively studied are percolation theory and random graphs [8]. The study in these areas has led to abstract mathematical frameworks, which can serve as solid starting points to investigate other systems where phase transition is relevant, such as probabilistic flooding in wireless mobile ad hoc networks [100]. Percolation theory [8] deals with fluid flow¹ (or any other similar process) in random media (describes the behavior of connected clusters in a random graph). In a percolation model, given that there is a probability p that there is an open path and $1 - p$ that there is no path, it has been observed that a phase transition can occur at the change of state between having finite numbers of clusters and having one infinite cluster. The existence of a phase transition in percolation theory infers the existence of a critical threshold probability, p_c where the phase transition occurs [61]. Given this observation, we will investigate the possible application of the same theory to the probabilistic flooding problem in VANETs. Such an application can provide valuable insights to the observed behavior and can lead to analytically verifiable designs. Phase transition has also been observed and studied in the context of Random Graphs. In mathematics, a random graph is a graph that is generated by some random process [53]. Different choices of the random process yield different random graph models. Various models have been studied in literature such as the Fixed Edge Number Model, the Bernoulli model, the Fixed Radius model and the Dynamic model. In the case of a Bernoulli model with N vertices and probability of an edge existence between any two nodes equal to p , it has been shown that for values of p greater than $\log(N)/N$, the graph is connected with high probability [19]. This result not only verifies the existence of

¹A usual setting here is that some liquid is poured on top of a porous material. The question is whether the liquid is able to make its way from hole to hole and reach the bottom.

phase transition phenomena in random graphs but also quantifies the critical point. The problem of obtaining similar results for the Fixed Radius model, which is an ideal representation of Mobile Ad Hoc Networks, is a challenging open research problem. Since random graphs are relevant to wireless ad hoc networks, it is not surprising that phase transition phenomena have been recently reported in various contexts in wireless ad hoc networks [8]. In the context of probabilistic flooding, phase transition can be extremely cost-efficient to observe, as it implies that there exists a critical probability beyond which all nodes receive the transmitted message with high probability.

3.2 Problem Formulation

In this paragraph we provide a formal definition of the information hovering problem in VANETs and we present the methodology and approach adopted to develop and evaluate an effective solution.

We consider a bounded geographical area, which we refer to as the *hovering area* (also known as anchor area), and in which we assume that there initially exists at least one vehicle which possesses critical information.

We refer to vehicles possessing critical information as information sources whose objective is to disseminate this information to all vehicles residing in the hovering area. We do not discuss how information sources initially generate or receive the critical information. All equipped vehicles employ wireless communication to form a vehicular ad hoc network. The main objective is then for the information sources to utilize the vehicular ad-hoc network to persistently disseminate the useful messages to all vehicles in the hovering area. Persistent dissemination implies that the desired property of all vehicles receiving the message is not instantaneously achieved, but holds for an arbitrary time interval, during which vehicles dynamically enter and leave the hovering area.

A number of approaches can be developed to solve the problem. However, the solution space can be refined by choosing solutions that make efficient use of the available resources, such as storage, power, bandwidth etc. Towards this goal we pose the additional objective of minimizing the number of exchanged messages. Small numbers of exchanged messages imply less storage requirements, less power consumption, less bandwidth use, and, above all, smaller latency of information delivery, since phenomena like severe congestion and contention which lead to collisions are avoided. The problem can thus be formally defined as follows:

Problem Definition:

Given a road map area, which we refer to as the Hovering Area, a single message information data and a time to live period, it is required to find an Information Dissemination Protocol, which when applied for the specific time to live period, almost all vehicles residing in the area receive the message with a probability approaching 1, with as low number of messages exchanged as possible.

3.3 Proposed Traffic Adaptive Information Hovering Protocol

A straightforward solution to the above problem is the application of epidemic routing merely within the hovering area. However, this approach leads to excessive congestion and contention in the case of high traffic density, and, in the case of low traffic density, it leads to low reachability. Furthermore, low traffic densities cause the network to become intermittently connected and it is not possible for partitioned uninformed areas to receive the critical information data. In such a scenario, vehicles outside the hovering area can serve as information bridges towards the partitioned uninformed areas. One can thus increase the achieved reachability by allowing epidemic routing in the entire network. This, however, leads to a huge number of redundant messages. For

this reason, in this work we present a solution which employs epidemic routing within the hovering area and probabilistic flooding outside the hovering area. Since epidemic routing outside the hovering area is avoided, the number of exchanged messages is greatly reduced.

The proposed information hovering protocol works as follows: We assume that all vehicles are equipped with a positioning system such as GPS. Vehicles exchange beacon messages that enable them to discover their neighbors. Information sources within the hovering area initiate the dissemination process by broadcasting a packet that includes the critical message and the following fields in its header: a TTL field, which determines the expiration time of the message, and a hovering area field, which contains information that can be used to determine the hovering area.

Then they mark all vehicles currently in their neighbor list as recipients of the critical message. Upon receiving the packet, a vehicle stores the critical message and the fields contained in the packet header and checks whether it resides within the advertised hovering area. If it does lie in the hovering area, it periodically checks whether its neighbors have been marked as recipients of the critical message. The frequency with which the check is conducted is referred to as the scanning frequency. If there exists at least one neighbor that has not received the critical message, then the vehicle rebroadcasts the message and marks all neighbors as recipients of the critical message. If the vehicle does not lie within the hovering area, it decides to initiate the rebroadcast procedure with probability p , otherwise it rejects the received message (i.e. with probability $1 - p$). The probability p is calculated based on a probability function f . Although this function can have several input variables, in this work we restrict our choices to decreasing functions of the distance from the hovering area only. With the term “distance from the hovering area” we mean the shortest distance from the boundaries of the area.

The reason behind this design choice is that, the further an informed vehicle is from the hovering area, the less probable it is that it will contribute to the finding of a yet unexplored path. The TTL field in the packet header is used to determine the expiration time of the critical message. The initial value of TTL field, which is the time to live period of the hovering protocol in seconds, depends on the application generating the message and it is determined at the beginning of the Hovering procedure. When a packet is received, the value in the TTL field is stored locally and appropriately decreased based on a local timer. If the vehicle rebroadcasts the critical message, it stores the updated local value in the corresponding TTL field. Once the TTL field expires, vehicles stop rebroadcasting the message and delete it from their memory. In this way we prevent a message from circulating indefinitely.

The most significant part of the overall design procedure is the determination of the probability function f . In this work, the method used to determine this function has been simulative. Our initial objective has been to determine the type of function to be utilized, and our subsequent efforts focus on tuning its parameters. The probability function, as discussed above, is required to be non-increasing. We thus consider four commonly used non-increasing candidate functions, which we evaluate using simulations in order to select the one that exhibits the best performance. The candidate functions chosen are the following:

- Gaussian-like function:

$$p = e^{-\frac{d^2}{2\sigma^2}} \quad (1)$$

- Step function:

$$p = \begin{cases} 1 & \text{if } d \leq \left(\frac{r}{9}\right) \\ 0.90 & \text{if } \left(\frac{r}{9}\right) < d \leq \left(\frac{r}{4}\right) \\ 0.80 & \text{if } \left(\frac{r}{4}\right) < d \leq \left(\frac{r}{2}\right) \\ 0.70 & \text{if } \left(\frac{r}{2}\right) < d \leq \left(\frac{7r}{9}\right) \\ 0.60 & \text{if } \left(\frac{7r}{9}\right) < d \leq \left(\frac{17r}{18}\right) \\ 0.50 & \text{if } \left(\frac{17r}{18}\right) < d \leq \left(\frac{10r}{9}\right) \\ 0.40 & \text{if } \left(\frac{10r}{9}\right) < d \leq \left(\frac{13r}{9}\right) \\ 0.30 & \text{if } \left(\frac{13r}{9}\right) < d \leq \left(\frac{5r}{3}\right) \\ 0.20 & \text{if } \left(\frac{5r}{3}\right) < d \leq (2r) \\ 0.10 & \text{if } (2r) < d \leq \left(\frac{43r}{18}\right) \\ 0 & \text{if } d > \left(\frac{43r}{18}\right) \end{cases} \quad (2)$$

- Linear function:

$$p = 1 - \frac{1.3d}{3r} \quad (3)$$

- Exponential-Like function:

$$p = e^{-\frac{0.7d}{r}} \quad (4)$$

where σ , d and r are design parameters. These parameters are tuned so that the candidate probability functions exhibit similar decreasing behavior as shown in figure 11.

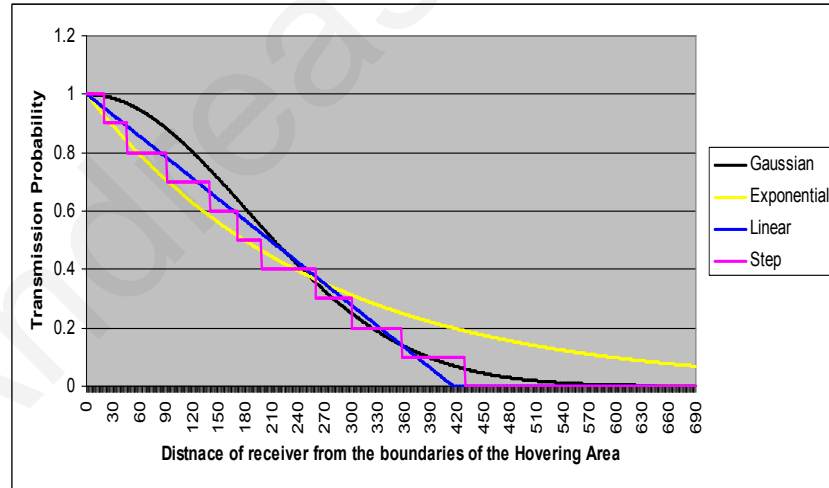


Figure 11: Candidate probability functions. Their parameters have been tuned to exhibit similar decreasing behavior.

The reference model used in all simulation experiments represents a section of the road network in the cities of Bellevue and Redmond in Washington. It includes congested arterial streets

and a saturated freeway. It models 37 semi-actuated coordinated and fully-actuated signalized intersections, converted to fixed time operation, 1 all-way stop controlled intersection, and numerous stop-controlled driveways throughout the model. The freeway modeled includes lanes for general traffic, high-occupancy vehicle (HOV)/Bus lanes, 3 closely spaced interchanges, 6 ramp meters, and 2 collector-distributor roads. Volumes were dynamically assigned and a 52-zone origin-destination trip table was generated using the VISUM interface with the EMME2 travel demand forecast model for the local 3-city area. VISUM is a comprehensive, flexible software system for transportation planning, travel demand modeling and network data management. Once the file reached convergence, Static routes were created and Dynamic Assignment was disabled. Baseline volumes from several cordon boundaries and screenlines were calibrated within 10 percent of those from the area travel demand forecast model. The model is setup to provide travel times for several key routes, system, intersection and approach delay, intersection and approach level of service (LOS), queue length, vehicle miles traveled, and volume output.

In this road network we identify six hovering areas which are shown as circles in figure 12. We select areas having different road topologies and traffic characteristics (traffic densities, vehicles' speeds, traffic lights etc.) in different portions of the given road network. In this way, our analysis aims to represent a large number of different realistic conditions of road networks where information hovering can be applied. Below we give a small description of each area:

- Area A: includes an intersection with medium traffic density.
- Area B: includes two intersections with medium traffic which also includes traffic lights and low vehicles' speed.
- Area C: includes a section of the freeway with high traffic density and high speeds.

- Area D: shows similar characteristics to area A but is located in a different section of the road map
- Area E: includes a section of the freeway and at the same time has an intersection of two other roads
- Area F: includes an intersection and several other smaller roads with smaller traffic density

For this set of experiments we consider area A. We conduct all the simulation experiments using VISSIM.

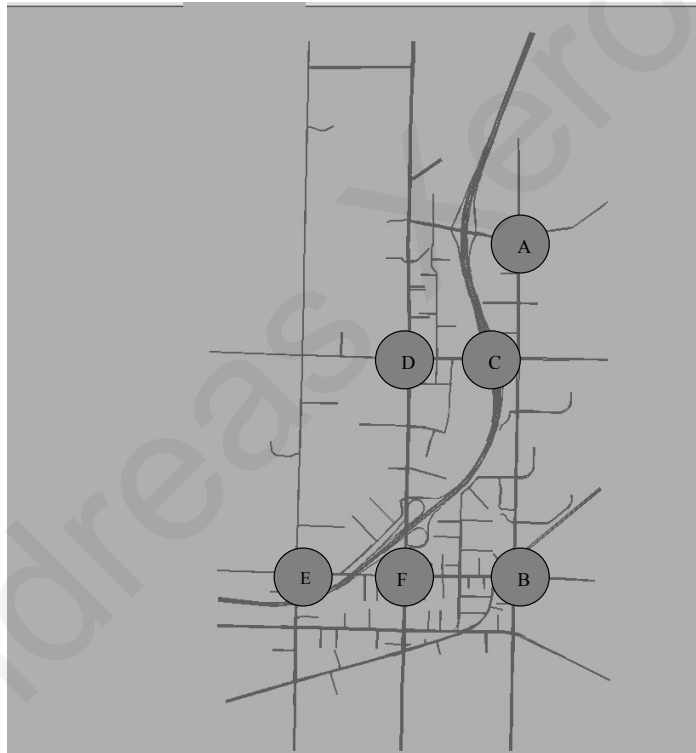


Figure 12: Road map of the reference model used in all simulation experiments. It represents a section of the road network in the cities of Bellevue and Redmond in Washington. Hovering areas are marked as black circles.

The four candidate probability functions are integrated in the information hovering protocol described above, and the resulting schemes are evaluated in terms of the average reachability achieved and the total number of messages received by the vehicles. The average reachability

is calculated over the simulation time after which the system has settled down to its equilibrium state, meaning that the achieved reachability is almost constant. The transmission range of each vehicle is selected randomly in the range from $140m$ to $220m$ and it remains constant through the time of simulation.

The performance of the four schemes is also compared to two other protocols. These protocols can be considered as extreme cases of the class of the probabilistic hovering protocols that can be designed: epidemic routing in the entire network, which implies a probability function equal to 1, and epidemic routing in the hovering area only, which implies a probability function equal to 0. Epidemic routing in the entire network yields high reachability, at the expense of a large number of redundant messages, whereas epidemic routing in the hovering area only achieves a lower number of exchanged messages, though at the expense of reduced reachability in cases of low traffic densities.

The performance of the considered information hovering schemes is investigated in area A for different values of the traffic density. We consider traffic density values in the range $1.2 - 18.3$ Vehicles/ Km^2 . We perform 9 different runs of the simulation with similar traffic density. Figure 13 shows the average number of received messages and the corresponding standard deviation values, whereas figure 14 shows the corresponding reachability values. The standard deviation when considering the number of message exchanged is relatively small. On the other hand, reachability measurements report a much larger spread. The reason for this is that the low traffic densities that must be taken into consideration yield small number of vehicles. For such small numbers any change in the number of vehicles is reflected in large percentage changes.

The results, as expected, show that the highest reachability is reported by the epidemic routing scheme, though at the expense of the highest number of exchanged messages. On the other end, epidemic routing within the hovering area and the linearly decreasing probability function report

the lowest reachability values. The remaining protocols exhibit similar reachability values close to the one achieved by epidemic routing. From these remaining protocols, the Gaussian-like probability function reports the smallest number of exchanged messages, thus making it the probability function of choice for the information hovering protocol to be designed [129].

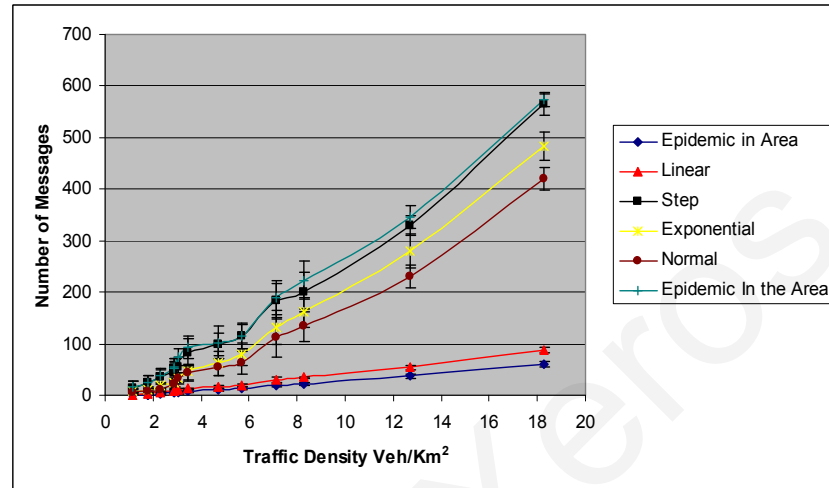


Figure 13: Number of messages exchanged in area A for each of the candidate probability functions.

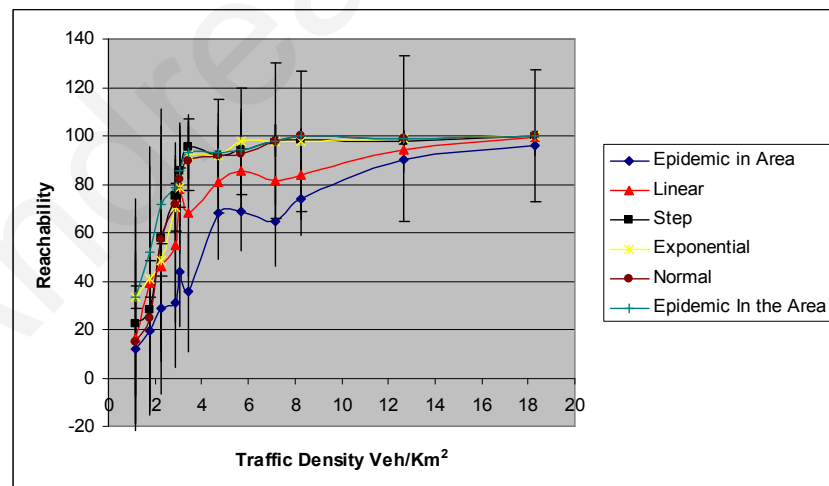


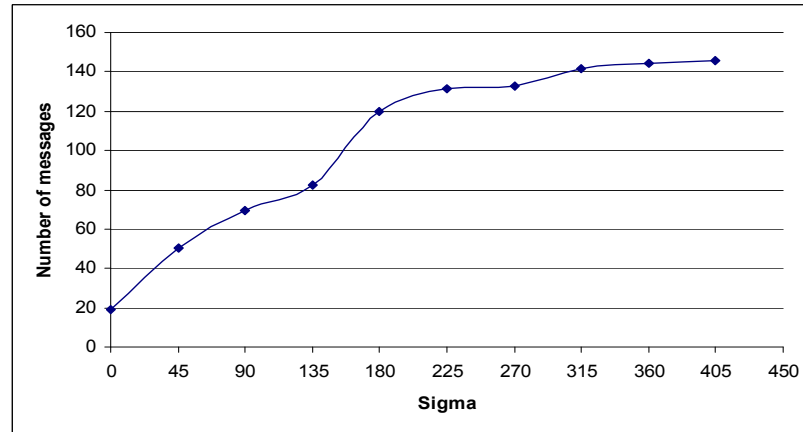
Figure 14: Reachability achieved in area A for each of the candidate probability functions.

It is worth noting that for low traffic densities the reachability does not increase smoothly. The reason is the very small number of vehicles in the area (2-3 vehicles) that render the results sensitive to the positions of these vehicles, which are random in nature. Depending on the positions of the vehicles, the connectivity of the vehicular network can become extremely unreliable, thus leading to low connectivity values. The phenomenon is more intense in the case of the epidemic routing in the area and in the case of the linear probability function.

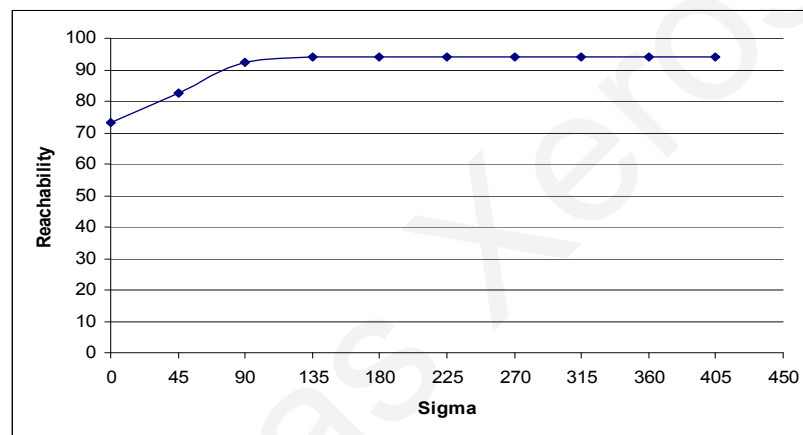
Having decided on the type of function to be used, the next step is to tune the parameters of this function in order to achieve the best possible performance. The equation of the chosen Gaussian-like probability function is given by equation (1). This equation is obtained by considering a normal curve $p = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{d^2}{2\sigma^2}}$, where the mean is equal to zero and the standard deviation is equal to σ , which is then multiplied by a factor $\sigma\sqrt{2\pi}$. The latter multiplication is necessary to ensure that the transmission probability on the boundary of the hovering area is equal to 1. The standard deviation σ is the only parameter of the function the value of which value needs to be regulated. The value of σ is a measure of the area in which critical message forwarding is allowed with high probability. The size of this area and, thus, the value of σ must be adaptively regulated based on the connectivity of the network within the hovering area. The smaller the connectivity is, the larger σ must be, in order to provide more message forwarding opportunities. More message forwarding opportunities enable additional paths to be explored, which in turn lead to increased connectivity and thus higher reachability. However, the connectivity within the hovering area is a difficult quantity to measure directly. Thus, the value of σ is adaptively regulated based on network parameters which affect the achieved connectivity, and their values are either known or estimated online. One such parameter is the transmission range of each vehicle. In this work we assume that the transmission range is constant to a value of 180 m, which is typical in vehicular ad hoc networks.

Another parameter which is known to affect network connectivity is the vehicle density within the hovering area. The higher the vehicle density is, the higher is the probability of high connectivity. Thus, a lower value of σ suffices to guarantee high reachability and low number of exchanged messages. The exact function, mapping the vehicle density to the desired value of σ , is determined using simulations. Various vehicle density values are considered in area A, and for each vehicle density the value of σ with the best possible performance is extracted. Despite the fact that the original simulation study is conducted in area A, the same design procedure is later applied in other hovering areas, demonstrating the robustness of the derived function with respect to changing topologies. In order to characterize the value of σ which yields, for a particular traffic density, the best possible performance, the traffic density in area A is first set equal to $6.7 \text{ Veh}/\text{Km}^2$ and the number of received messages and the reachability achieved by the proposed information hovering protocol are recorded for a predetermined set of values of σ in the range 0 to 405. The recorded values are shown in figure 15.

We observe a strictly increasing number of messages as the value of σ increases. In addition, the achieved reachability also increases with increasing σ until a critical value is reached, beyond which the reachability is almost constant attaining values close to 95%. For a particular vehicle density, this critical value of σ is the desired one, as it achieves high reachability with the minimum possible number of exchanged messages. The existence of this critical value is a strong indication of the presence of phase transition phenomena, which are typical in the theory of percolation theory and random graphs and are also observed in VANETs in [81]. The same type of behavior is observed in other hovering areas and for different traffic density values. Figure 16 shows the achieved reachability vs the number of exchanged messages when the vehicle density in area B is set to $15 \text{ Veh}/\text{Km}^2$, a setup with very high vehicle density, whereas figure 17 shows the achieved reachability vs the number of exchanged messages when the vehicle density in area C is set to



(a) Number of Exchanged Messages



(b) Reachability

Figure 15: Number of messages exchanged and Reachability in hovering area A for different values of σ when the vehicle density is set to $6.7 \text{ Veh}/\text{Km}^2$.

$4.9 \text{ Veh}/\text{Km}^2$. Each point on these graphs corresponds to a particular value of σ . However, due to the strictly increasing relationship between σ and the number of exchanged messages, higher values of exchanged messages also imply higher values of σ . We consider values of σ in the range 0-405. A value of σ equal to zero corresponds to the application of epidemic routing in the hovering area only. We also consider the application of epidemic routing in the entire network, which corresponds to a value of sigma equal to ∞ . The graphs indicate the existence of a critical value of σ , beyond which the achieved reachability is almost constant. Such critical density values

are obtained for values of the traffic density in the range $2 - 16 \text{ Veh}/\text{Km}^2$ in all hovering areas depicted in figure 12.

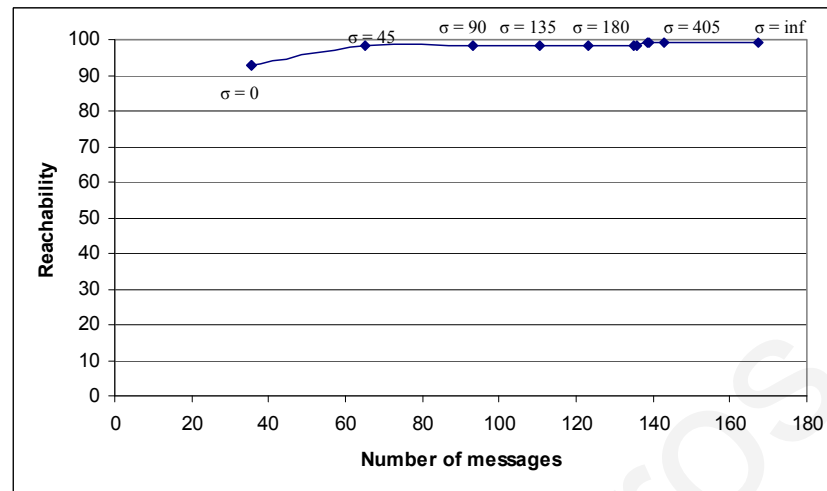


Figure 16: Reachability vs number of message exchanged in hovering area B, for different values of σ when the traffic density is set to $15 \text{ Veh}/\text{Km}^2$.

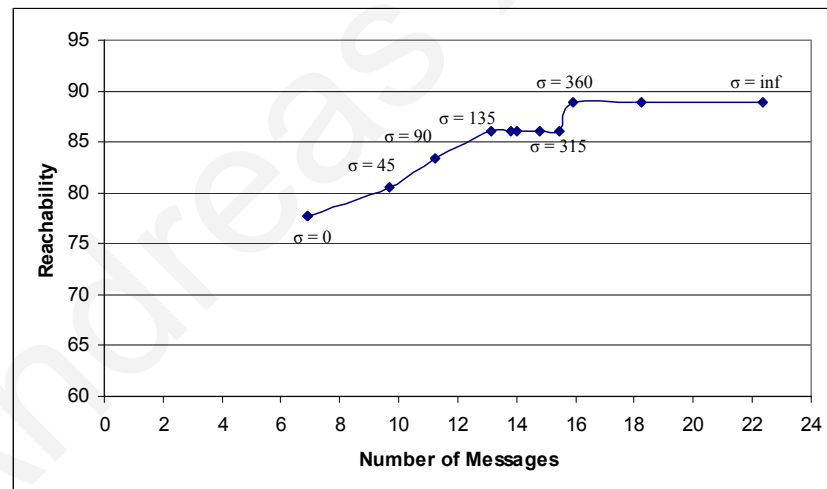


Figure 17: Reachability vs number of message exchanged in hovering area C, for different values of σ when the traffic density is set to $4.9 \text{ Veh}/\text{Km}^2$.

The extracted critical σ values vs the traffic density are shown graphically in figure 18. We observe that the relationship between the desired σ and the vehicle density exhibits an exponentially

decreasing behavior which is similar in all hovering areas. This demonstrates that the relationship is almost independent of the considered topology, which implies that a universal σ vs vehicle density function can be derived. We derive such a function by considering a least squares fit between the curves. The resulting function is shown graphically in figure 18. It must be noted that the proposed method is mostly effective for relatively low traffic densities within the hovering area which lead to low reachability values. In cases of high traffic density (above approximately 16veh/km²), the proposed protocol, as shown by the rebroadcast probability function of figure 18, does not rebroadcast messages outside the hovering area. This makes it equivalent to the blind flooding in the area only case and can be observed in the figure 22, where the two protocols exhibit identical behavior at the highest traffic densities considered. The information hovering protocol can thus be further enhanced by applying probabilistic flooding within the hovering area in cases of high traffic density. The switching between probabilistic flooding outside the hovering area and probabilistic flooding inside the hovering area can be done at the density where the rebroadcast probability in figure 18 first becomes equal to zero. This enhancement will be the topic of future research.

The vehicle density within the hovering area is an unknown quantity that needs to be estimated online. In the subsequent section we discuss the estimation method.

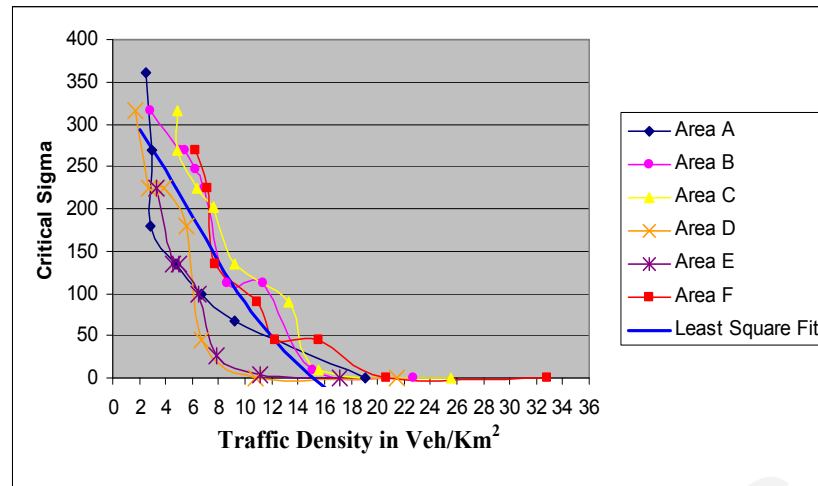


Figure 18: Critical value of sigma vs traffic density graphs in different hovering areas. We observe similar behavior in all hovering areas. The relationship used in the proposed protocol is obtained by applying a least squares fit between the curves.

3.4 Estimation of the Vehicle Density

The vehicle density within a confined area is a global quantity that is impossible to calculate accurately in a distributed fashion, especially in cases where the vehicular network is intermittently connected. Since in the considered scenarios the vehicular network is often partitioned, distributed algorithms must be developed to estimate the quantity online. As described in the previous section, these estimates are required by the developed information hovering protocol to calculate the rebroadcast probability outside the hovering area. The metric used in this work to estimate the vehicular density is the average number of neighbors of the vehicles residing in the hovering area. The basic question we pose is whether the vehicular density relates to the average number of neighbors in a roadway setting and how. The plethora of possible roadway topologies makes any general analysis intractable, so in this work we derive the relationship theoretically considering the simple case of two two-way straight roads crossing perpendicularly at the center of the circular hovering area, as shown in figure 19.

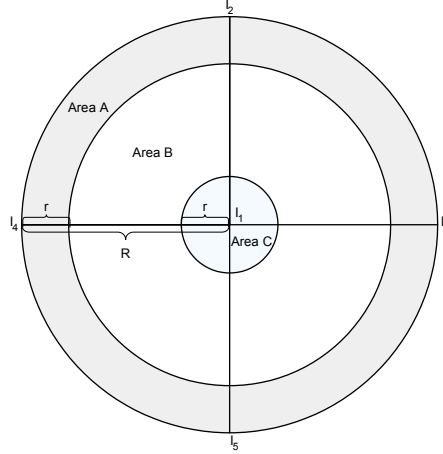


Figure 19: The road topology of the model used to relate analytically the vehicle density with the average number of neighbors.

Our objective is to calculate the expected number of neighbors of each vehicle when the number of vehicles residing in the hovering area is equal to $n+1$. In such case the vehicle density is given by $\frac{n+1}{\pi R^2}$, where R is the radius of the hovering area. We assume that at any time instant the location of each vehicle is a uniformly distributed random variable over all possible locations in the considered roadway system. Two vehicles are considered to be neighbors when they lie within their transmission range. It follows from our previous assumption that the events of any two vehicles being neighbors are independent. The event of any two vehicles being neighbors is modeled as a bernoulli trial with probability of success equal to p . It follows from the independence of the bernoulli trials that the number of neighbors of any vehicle is a binomially distributed random variable whose expected value is equal to np . So, in order to calculate the expected number of neighbors when the number of vehicles in the hovering area is equal to $n+1$, it suffices to calculate the probability p of two vehicles being neighbors. We calculate this probability below.

We first define basic notations utilized in the subsequent analysis. As mentioned above, we assume that in the considered hovering area the roadways systems consist of two two-way straight-line roads intersecting perpendicularly. I_1 represents the point of intersection where I_2 up to I_5

represent the edge-points on the considered roadway system as shown in figure 19. Among the $n + 1$ vehicles residing in the hovering area we randomly select two, referred to as veh_1 and veh_2 . Our objective is to calculate the probability that the distance between these vehicles is less than their transmission range r , which is assumed to be constant. h_{i-j} represents the directional roadway link between points I_i and I_j . The direction of the traffic flow is from I_i to I_j . We partition the hovering area in three subareas A, B, C , as shown in figure 19. Area A represents the circular ring enclosed by the perimeters of the concentric circles with radius R and $R - r$ respectively. Area B represents the circular ring enclosed by the perimeters of the concentric circles with radius $R - r$ and r respectively. Area C represents the circle with radius r . The probability of vehicle veh_1 residing in area A is denoted by P_1^A . In a similar way we define P_1^B and P_1^C . Given that veh_1 lies in area A, the probability that it has veh_2 as a neighbor is constant and is denoted by Pn_A . Similarly, we define Pn_B and Pn_C .

Due to the areas A, B and C constituting a partition of the hovering area, the desired probability p is given by

$$p = P_1^A * Pn_A + P_1^B * Pn_B + P_1^C * Pn_C \quad (5)$$

Probability P_1^A is equal to the ratio of the total road segment lying in area A over the total road segment within the hovering area. Since the road is bidirectional, the length of the road segment in Area A is given by $8r$, whereas the total length of the road segment within the hovering area is given by $8R$. It follows that:

$$P_1^A = \frac{r}{R} \quad (6)$$

Similarly, one can deduce that:

$$P_1^B = \frac{R - 2r}{R} \quad (7)$$

$$P_1^C = \frac{r}{R} \quad (8)$$

We now calculate the probability of Pn_A . Given that veh_1 is in area A, we assume that it is located on road h_{1-2} distance k from the perimeter of the area and in this case denote by $Pn_{1-2}^A(k)$ the probability that veh_2 is its neighbor. k can attain values between 0 and r and so the conditional probability that veh_1 has veh_2 as its neighbor when it lies on road h_{1-2} within area A, Pn_{1-2}^A , is given by $\frac{1}{r} \int_0^r Pn_{1-2}^A(k) dk$. veh_2 will be neighbor of veh_1 in two cases: if it is located on any side of the road between veh_1 and I_2 and if it is located on any side of the road within a distance r from veh_1 in between veh_1 and I_1 . The probability that the first case is valid is equal to the length of the roadway section between veh_1 and I_2 over the total length of the roadway in the hovering area, whereas the probability of the second case being valid due to its geometry is equal to $\frac{1}{4}P_1^A$. So, $Pn_{1-2}^A(k)$ is equal to $\frac{2r}{8R} + \frac{2k}{8R}$ and Pn_{1-2}^A is given by $\frac{1}{r} \int_0^r Pn_{1-2}^A(k) dk = \frac{3r}{8R}$. Similarly, we define Pn_{1-j}^A and Pn_{j-1}^A for $j = \{2, 3, 4, 5\}$. Moreover, we define P_{1-j}^A and P_{j-1}^A $j = \{2, 3, 4, 5\}$ to be the probabilities that veh_1 lies in road section h_{1-j} , h_{j-1} respectively in area A. It follows that $Pn_A = \sum_{j=2}^5 Pn_{1-j}^A P_{1-j}^A + Pn_{j-1}^A P_{j-1}^A$. Due to the symmetry of the problem $Pn_{1-j}^A = Pn_{j-1}^A = Pn_{1-2}^A \forall j = \{2, 3, 4, 5\}$ and $P_{1-j}^A = P_{j-1}^A = \frac{1}{8} \forall j = \{2, 3, 4, 5\}$. It follows that:

$$Pn_A = \frac{1}{8} * Pn_{1-2}^A * 8 = \frac{3r}{8R} \quad (9)$$

In area B, given that veh_1 is placed on road h_{1-2} , then veh_2 will be a neighbor of veh_1 if it is less than r apart from veh_1 in both directions on h_{1-2} or on road h_{2-1} . In this case, $Pn_{1-2}^B(k)$

is equal to $\frac{4r}{8R}$ and Pn_{1-2}^B is given by $\frac{1}{R-2r} \int_r^{R-r} Pn_{1-2}^B(k)dk = \frac{4r}{8R}$. Using the same arguments leading to equation (9), it follows that

$$Pn_B = \frac{1}{8} * Pn_{1-2}^B * 8 = \frac{r}{2R} \quad (10)$$

Finally, when veh_1 is placed in area C on road h_{1-2} , then veh_2 will be its neighbor if it lies a distance less than r from veh_1 on roads h_{1-2} and h_{2-1} . In addition, assuming that veh_1 is located a distance k from intersection I_1 on road h_{1-2} , then veh_2 can be a neighbor of veh_1 if it is located within a distance $\sqrt{r^2 - k^2}$ from I_1 on roads $h_{1-3}, h_{3-1}, h_{1-4}, h_{4-1}$. So, given that veh_1 is in area C, the probability that the two vehicles are neighbors Pn_{1-2}^C is given by $\frac{4r}{8R} + 4 \int_0^r \frac{\sqrt{r^2 - k^2}}{8Rr} dk = \frac{r}{2R} + \frac{1}{2Rr} \int_0^r r \sqrt{1 - (\frac{k}{r})^2}$. By setting $\frac{k}{r} = \sin \theta$, we get: $Pn_{1-2}^C = \frac{r}{2R} + \frac{1}{2Rr} r \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 \theta} r \cos \theta d\theta = \frac{r}{2R} + \frac{r}{4R} \int_0^{\frac{\pi}{2}} 1 + \cos 2\theta d\theta = \frac{r}{2R} + \frac{r\pi}{8R} = \frac{4r+r\pi}{8R}$. Using the same arguments leading to equation (9), it follows that

$$Pn_C = \frac{1}{8} * Pn_{1-2}^C * 8 = \frac{4r + r\pi}{8R} \quad (11)$$

Substituting equations (6)-(11) in (5) we obtain

$$p = \frac{r}{R} \frac{3r}{8R} + \frac{R-2r}{R} \frac{r}{2R} + \frac{r}{R} \frac{4r+r\pi}{8R} = \frac{4Rr + \pi r^2 - r^2}{8R^2} \quad (12)$$

which implies that the expected number of neighbors $E[X]$, where X denotes the random variable of the number of neighbors of a vehicle in the hovering area, when the vehicle density is equal to $\frac{n+1}{\pi R^2}$ is given by:

$$E[X] = n \frac{4Rr + \pi r^2 - r^2}{8R^2} \quad (13)$$

The above probability is derived using a simple model of the roadway system which consists of two intersecting roads only. In the remainder of this section we compare our theoretical findings with simulation results extracted from scenarios that relax the simplifying assumptions of our theoretical model. We observe a surprisingly good agreement between the theoretical findings and the simulation results, despite the simplicity of the theoretical model.

The reference model used in the simulation experiments is the one considered in the previous section, which is shown in figure 12. We utilize all the hovering areas (A-F) shown in figure 12. All hovering areas have identical topology, being circular with radius $R = 500m$. The roadway topologies within the hovering areas, however, differ significantly and are more complex than the topology of the theoretical model, involving a considerable number of intersecting roads. In all simulation experiments, the vehicle transmission range r is set equal to 180m. In each hovering area, we consider various values for the number of vehicles residing in the area and for each number we use the simulation results to find the average number of neighbors. We compare these results with the function of equation (13), which becomes $E[X] = 0.21n$ when R and r are assigned the values considered in the simulation experiments.

The results are shown in figure 20. We observe very good agreement between the simulation results and our theoretical findings. This demonstrates the validity of equation (13) and that it can be successfully used in the proposed hovering scheme to accurately estimate the vehicle density within the hovering area when the average number of neighbors is known priori. But how does each vehicle obtain estimates of the average number of neighbors? Each vehicle in the hovering area maintains the following state variables: the estimate of the average number of neighbors in the hovering area (S_t), the number of its neighbors n , the number of neighbors of each of its neighbors (H_t^j for the neighbor j), and the estimates of the average number of neighbors of each of its neighbors (S_t^j for the neighbor j). The latter three quantities are obtained by each vehicle

by means of a beacon exchange mechanism. We also use a as the learning factor of the algorithm.

The following exponentially weighted moving average like algorithm is used to update S_t :

$$S_t = a * \frac{(n + \sum_{j=1}^n H_t^j)}{n + 1} + (1 - a) * \frac{(S_{t-1} + \sum_{j=1}^n S_{t-1}^j)}{n + 1} \quad (14)$$

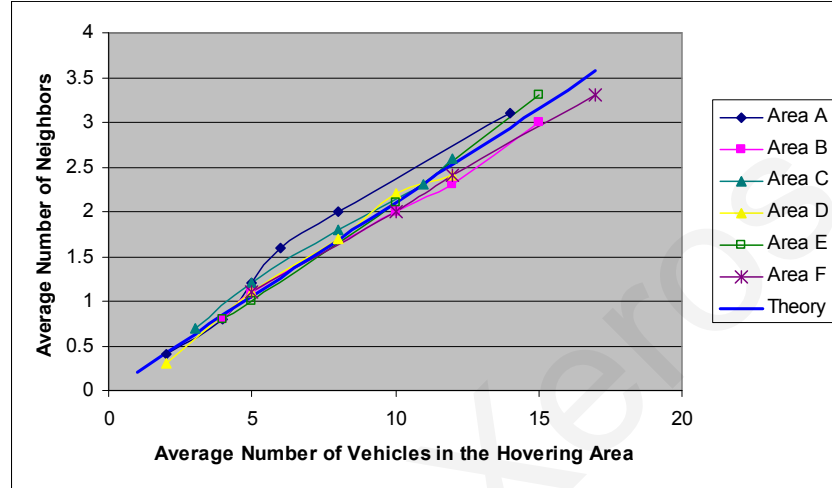


Figure 20: Average number of neighbors obtained using simulations in different hovering areas, for different average number of vehicles in the areas and comparison with the predicted values using equation (13)

It must be noted that the above calculation is only conducted by vehicles within the hovering area. Outside the hovering area, when vehicles exchange neighbor update beacon messages, they also exchange their estimates of the average number of neighbors within the hovering area. Upon receiving an average neighbor value (S_{t-1}^j) from neighbor j , a vehicle updates its estimate of the average number of neighbors within the hovering area S_t according to the equation:

$$S_t = \frac{(S_{t-1} + \sum_{j=1}^n S_{t-1}^j)}{n + 1} \quad (15)$$

Updates outside the hovering area are essential. They ensure that most recent updates, which are provided by vehicles exiting the hovering area, quickly spread in the entire network, while also improving the robustness of the system with respect to erroneous measurements.

Obviously, the weighted moving like algorithm described is applied by each vehicle locally. This will result in different values of the estimation of the average number of neighbors between vehicles. Subsequently each vehicle will have a different estimation of the traffic density which will leads to a different value of σ meaning that the Gaussian-like probabilistic function will be different. This actually works for the benefit of the proposed solution which handles efficiently the cases of having a variety of traffic densities inside a hovering area. For example, we consider a defined hovering area where there are two partitioned subareas of vehicles, one with high traffic density and one with low. Messages sent by vehicles originated from the high traffic density subarea will have lower rebroadcast probability reducing, in this way, the redundant messages exchanged. On the other hand, messages transmitted from vehicles in the lower traffic density will have higher rebroadcast probability increasing in this way its reachability.

3.5 Performance evaluation

In this section we evaluate the performance of the proposed information hovering protocol using simulations. The reference model used in all simulation experiments is drawn from the same area considered in the design procedure i.e. a section of the road network including congested arterial streets and saturated freeway in the cities of Bellevue and Redmond in Washington. However, we consider different hovering areas than the ones used in the design procedure, which are shown in figure 21 and are denoted by the letters H and G. We conduct all the simulation experiments using VISSIM. The performance metrics used are the reachability within the hovering area and the number of received messages. As discussed in previous sections, the design objective has been to achieve the highest possible reachability with the minimum number of exchanged messages.

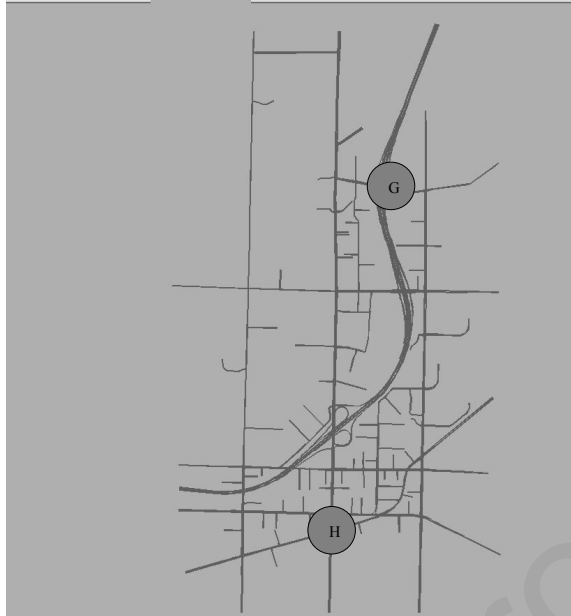
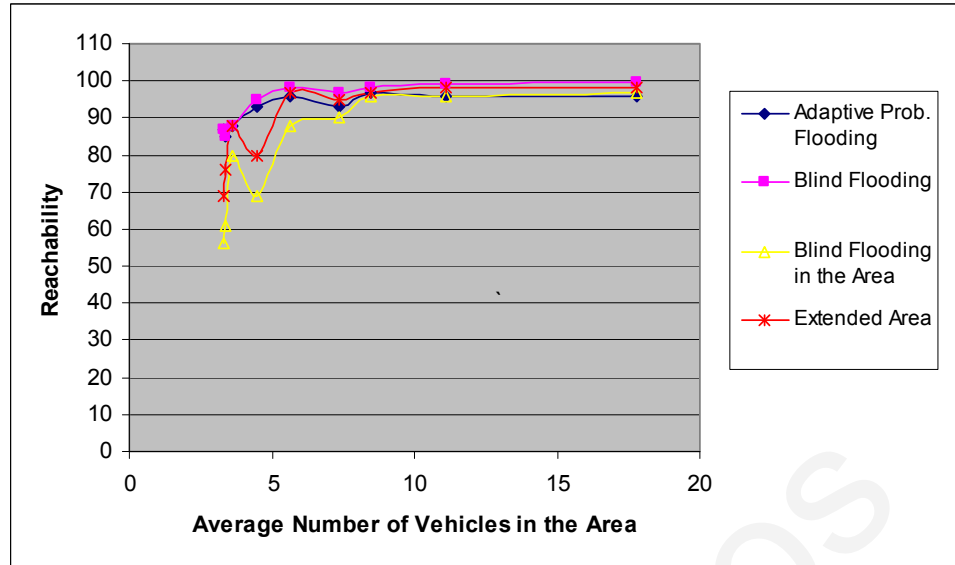


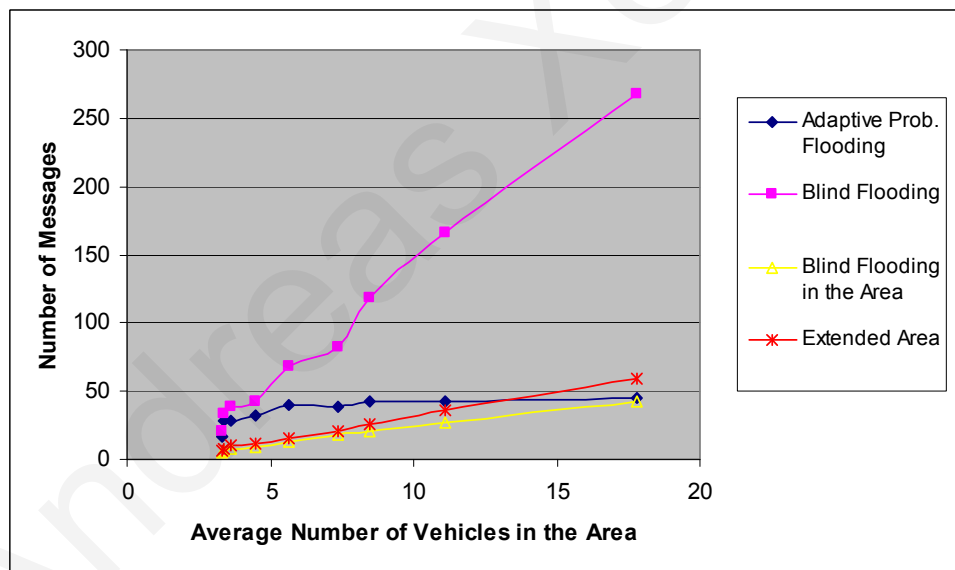
Figure 21: Road topology of the reference model used to evaluate the performance of the proposed information hovering protocol. The considered hovering areas are marked on the diagram as black circles.

We first conduct a comparative study in order to investigate to what extent the proposed scheme achieves its design objectives relative to other schemes that have appeared in the literature: epidemic routing in the hovering area only, epidemic routing in the entire vehicular network, and the scheme proposed in [37], which allows exchange of messages in a closed area outside the hovering area. Since the authors in [37] do not give guidelines on how to choose the size of this area, we have set the extended area to be equal in size to the hovering area. This is achieved by setting the radius of the area in which message exchange is allowed equal to $\sqrt{2}$ times the radius of the hovering area. The number of exchanged messages and the reachability reported in area H are shown in figure 22 as we increase the average number of vehicles residing in the area. We observe that for low traffic densities, the reachability achieved by the four schemes is significantly less than 100% indicating the existence of partitioned areas which prevent some vehicles to be informed of the critical message. As the vehicle density increases, both the achieved reachability and the number of exchanged messages increases. However, we observe that the proposed

adaptive probabilistic flooding scheme exhibits superior performance as it achieves high reachability values close to the ones reported by the scheme that applies epidemic routing in the entire network and low number of exchanged messages similar to the ones reported by the scheme that applies epidemic routing in the hovering area only. The proposed scheme thus gets the best of the two aforementioned schemes. The extended hovering area approach is successful in achieving low number of exchanged messages, at the expense, however, of low reachability values. Similar results are obtained in area G and are shown in figure 23. This demonstrates the ability of the proposed scheme to work effectively in various hovering areas with different road topologies and traffic characteristics.

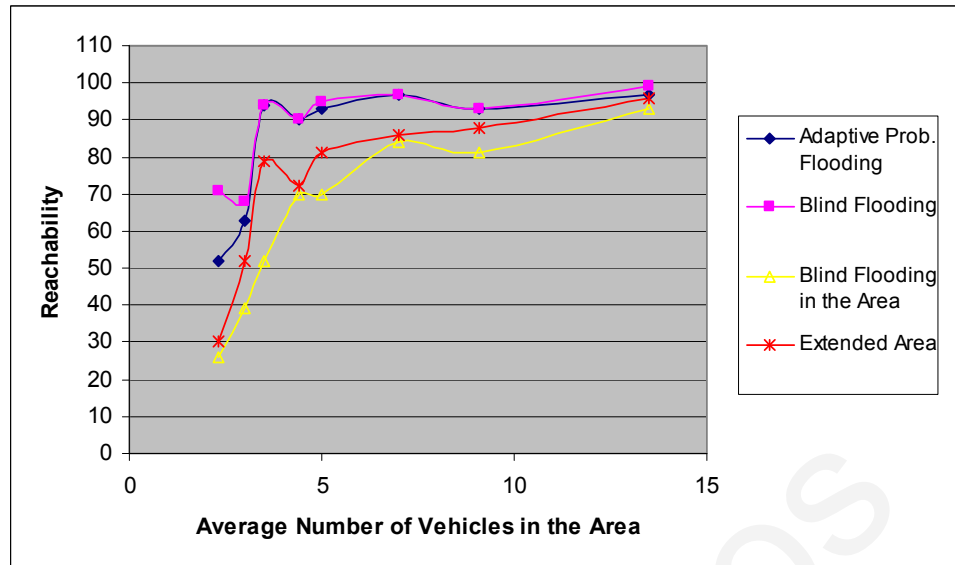


(a) Reachability

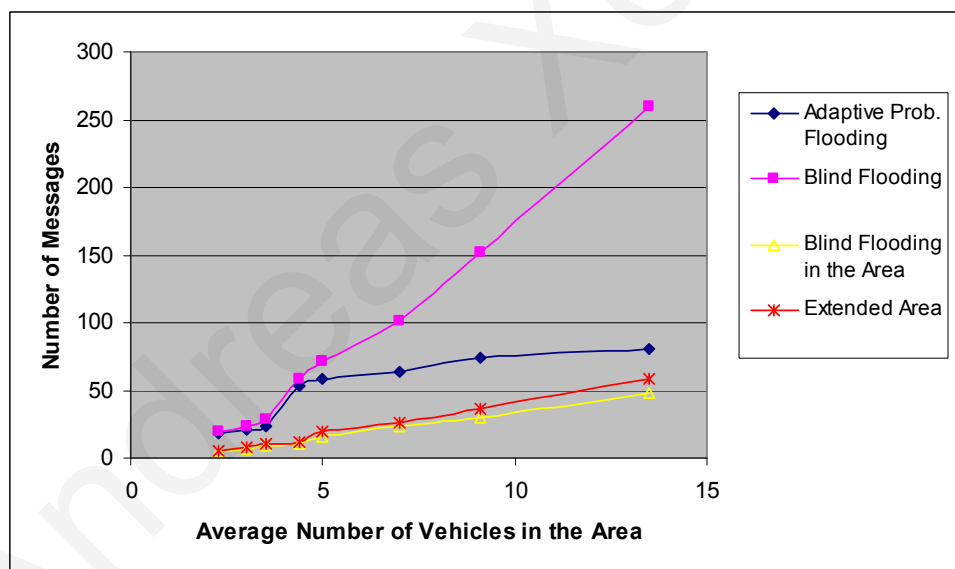


(b) Number of Exchanged Messages

Figure 22: Comparison of the proposed adaptive probabilistic scheme with other hovering schemes in terms of the reachability achieved and the number of messages exchanged in area H, for different average number of vehicles in the area.



(a) Reachability

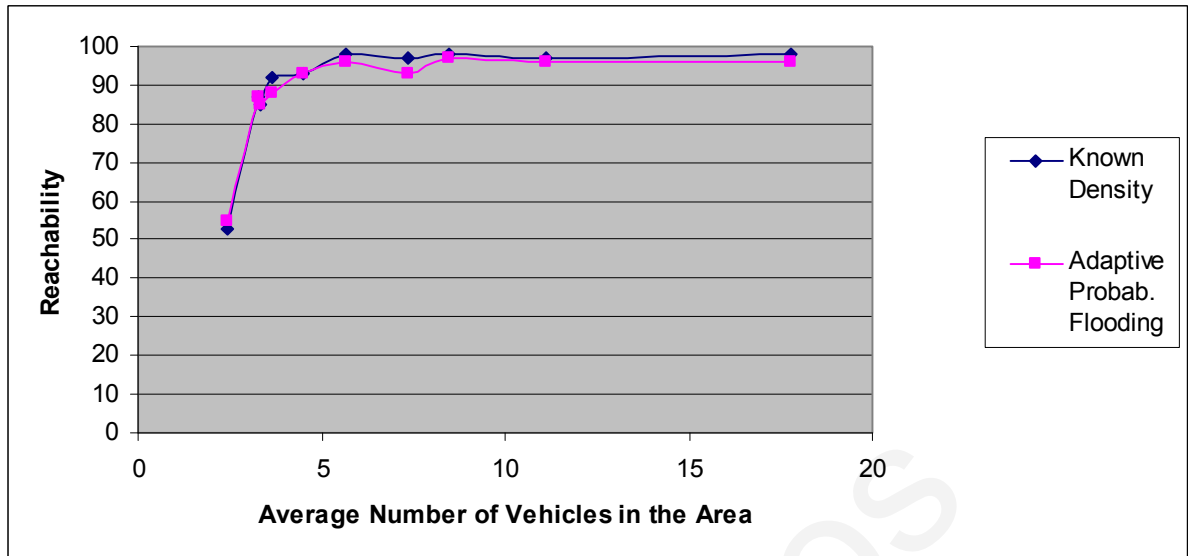


(b) Number of Exchanged Messages

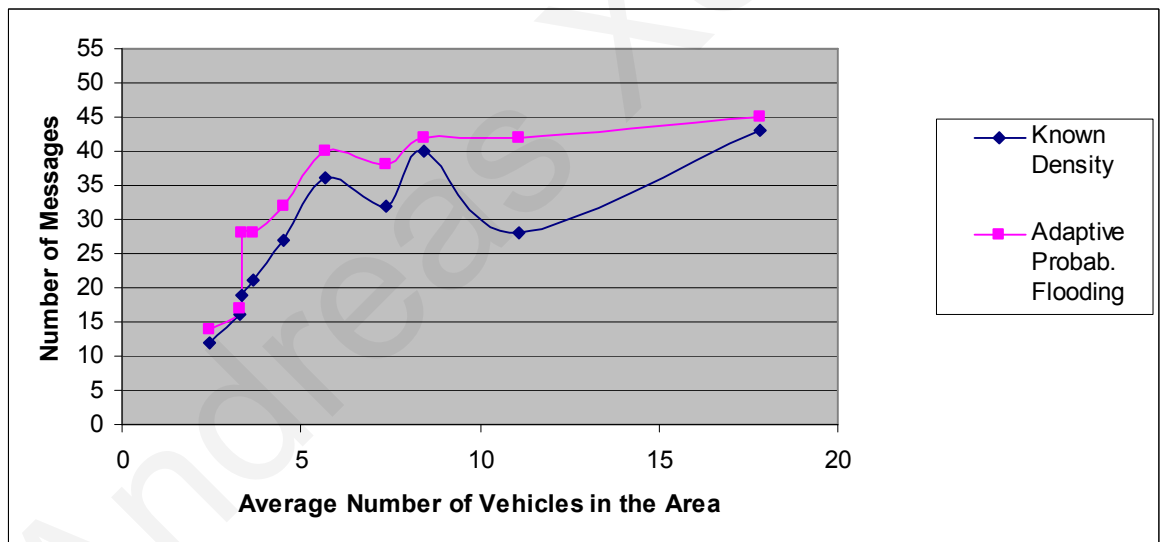
Figure 23: Comparison of the proposed adaptive probabilistic scheme with other hovering schemes in terms of the reachability achieved and the number of messages exchanged in area G, for different average number of vehicles in the area.

The next set of experiments aims at investigating the effect of the estimation algorithm on the performance of the proposed information hovering protocol. How would the hovering protocol perform, had we known the actual vehicle density within the hovering area and to what extent, would this performance be compromised when the actual values are replaced by estimates generated by the estimation algorithm proposed in the previous section?

Towards this end we consider an information hovering protocol that assumes the vehicle density within the hovering to be known and utilizes the known density value to tune the parameter σ of the rebroadcast probability function according to the function depicted on figure 18. We compare the performance of this protocol with the performance of the proposed information hovering protocol, which replaces the unknown density values with estimates generated online. Simulation experiments are conducted in both areas G and H and the results are summarized in figure 24 and in figure 25. The figures show the number of exchanged messages and the reachability achieved by the two protocols as we increase the vehicle density. The simulation results indicate that the two protocols exhibit similar behavior, demonstrating that the estimation algorithm does not compromise the performance of the information hovering which could be achieved, had the vehicle density values be known. The estimation algorithm is thus shown to be successful in generating accurate enough estimates of the unknown vehicle density.

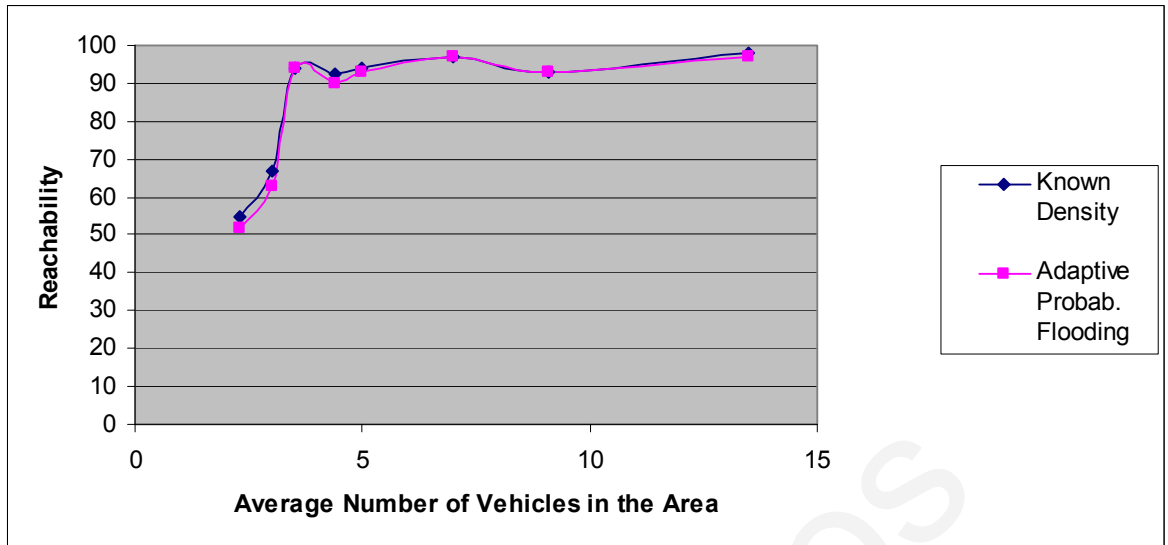


(a) Reachability

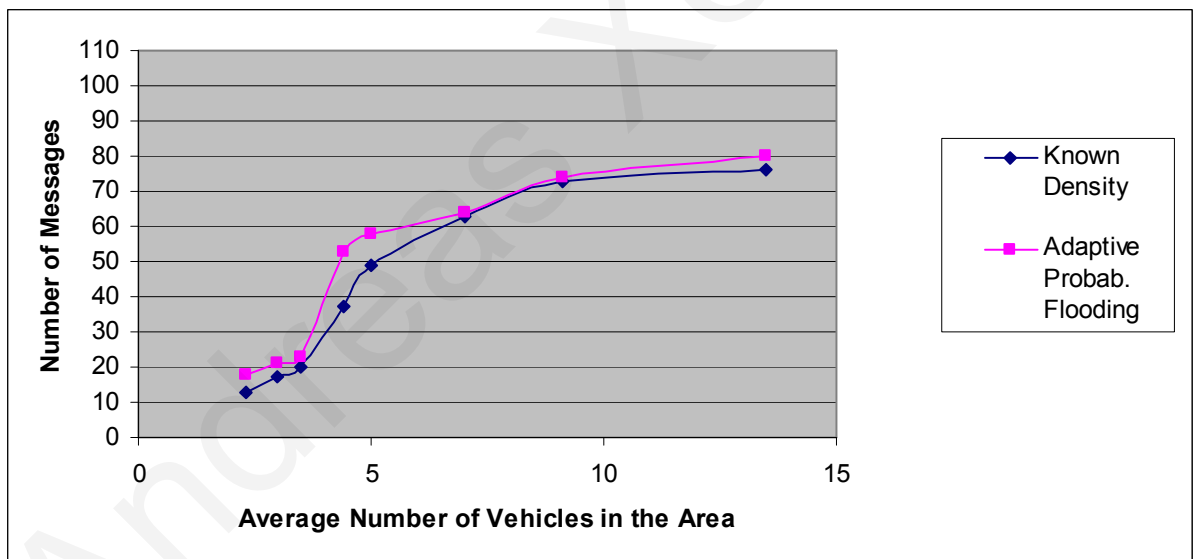


(b) Number of Exchanged Messages

Figure 24: Comparison of the proposed hovering scheme using the vehicle density estimation of equation (13), with a protocol that assumes the vehicle density within the hovering area to be known. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area H for different average number of vehicles in the area.



(a) Reachability



(b) Number of Exchanged Messages

Figure 25: Comparison of the proposed hovering scheme using the vehicle density estimation of equation (13) with a protocol which assumes the vehicle density within the hovering area to be known. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area G for different average number of vehicles in the area.

The information hovering protocol developed in this work attempts to solve the problem of low reachability in cases of low traffic density, by allowing controlled exchange of messages outside the hovering area. The latter is achieved by employing the idea of probabilistic flooding outside the hovering area. A unique feature of the protocol is that it is adaptive, in the sense that the rebroadcast probability is regulated based on estimates of vehicle density calculated online. The information hovering protocol proposed in [37] is based on a different approach, which allows message exchange in a restricted area outside the hovering area. As shown earlier in this section, the information hovering protocol proposed in this work outperforms the protocol proposed in [37] when the size of the restricted area is fixed. The issue now, however, is how the two approaches would compare had the size of the restricted area been adaptively regulated based on the vehicle density. The idea of adaptively regulating the area in which message exchange is allowed, based on estimates of the vehicle density, has been briefly discussed in [37]. However, no guidelines have been given on how to design the estimation algorithm and how to use the estimates that it generates to regulate the size of the area. In this section, we modify the algorithm proposed in [37] to make it adaptive, and we compare it with the proposed scheme. In order to render the algorithm proposed in [37] adaptive, we use similar ideas and almost the same design procedure used to design the proposed information hovering protocol.

We consider 12 different configurations of the protocol that applies epidemic routing in an extended circular area encompassing the hovering area itself. The hovering area is also circular with radius R . The protocol differs in the size of the extended area. One implementation applies epidemic routing in the hovering area only, while the radius of the extended area of the rest of the implementations differs by $0.2R$. For a particular traffic density in a specific hovering area we compare the performance of the 12 implementations and we choose the one which exhibits superior performance in the sense that it reports high reachability with the smallest possible number of

exchanged messages. The radius of the extended area of this implementation is referred to as the critical radius. We repeat the same procedure for different traffic densities and different hovering areas, and the resulting critical radius obtained are shown graphically in figure 26. Similar to the critical sigma curve of figure 18, the critical radius curve exhibits an exponentially decreasing pattern as the vehicle density increases. This pattern does not vary significantly in each hovering area, indicating that a universal curve may be obtained that is independent of the chosen road topology. Such a curve is extracted by applying a least squares fit between the curves of figure 26. The least squares fit curve is also shown on the same diagram. Each vehicle estimates the vehicle density within the hovering area utilizing the methodology described in the previous section and uses the least squares curve to calculate the radius of the area in which message exchange is allowed. Upon receiving a relevant message, each vehicle decides to rebroadcast the message if it lies within the calculated extended area.

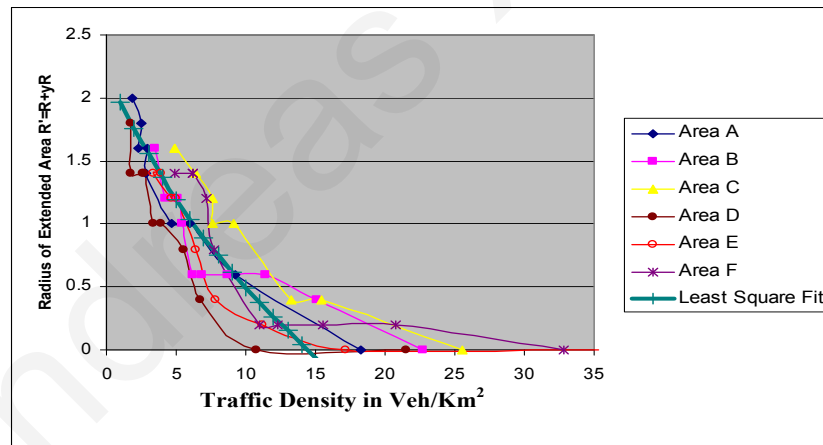


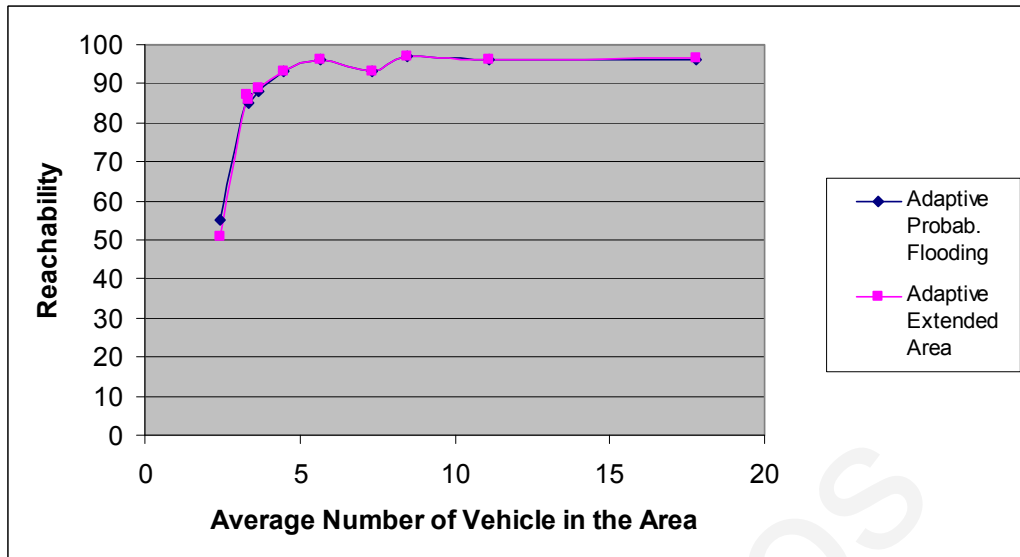
Figure 26: Percentage increase of the critical radius relative to the radius of the hovering area vs traffic density graphs in different hovering areas. We observe similar behavior in all hovering areas. The relationship used in the adaptive version of the protocol proposed in [37] is obtained by applying a least squares fit between the curves.

The proposed information hovering protocol and the adaptive version of the protocol proposed in [37] are compared using simulations in areas G and H.

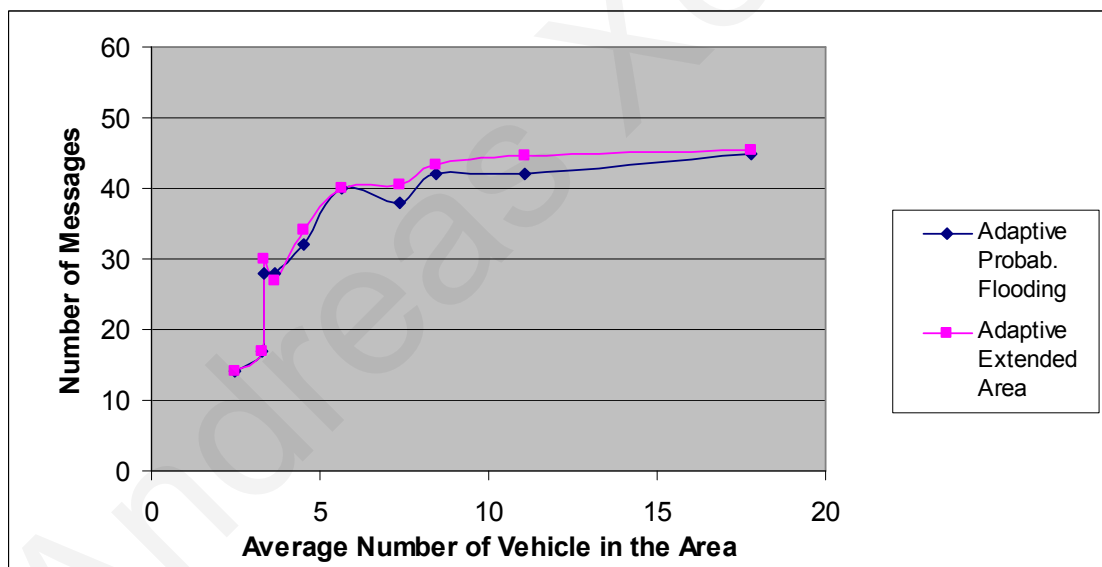
Similar to the adaptive probabilistic flooding approach, the traffic density is estimated locally by each vehicle using the weighted moving average-like algorithm, as described earlier. This means that each vehicle might come up with a different value for the extended hovering area which improves performance in cases of traffic density variations within the hovering area.

The simulation results for each area are shown in figure 27 and in figure 28 respectively. We observe almost identical behavior, indicating that any of the two approaches can be used to satisfy the posed design objectives. Thus, our contribution goes beyond the proposal of a specific information hovering protocol, but extends to the introduction of a design procedure that can be used to design a class of density adaptive hovering protocols.

It worth noting that the extended area solution is simpler to implement than probabilistic flooding. The two solutions have similar complexity when calculating the critical value of σ and the critical radius respectively. However, probabilistic flooding employs a parameterized Gaussian function to calculate the rebroadcast probability based on the distance from the hovering area. In addition, it uses a random generator to decide whether to retransmit or not. These introduce additional complexity relative to the extended area solution, which merely checks whether a vehicle's location lies in the extended area dictated by the critical radius to decide whether to rebroadcast or not.

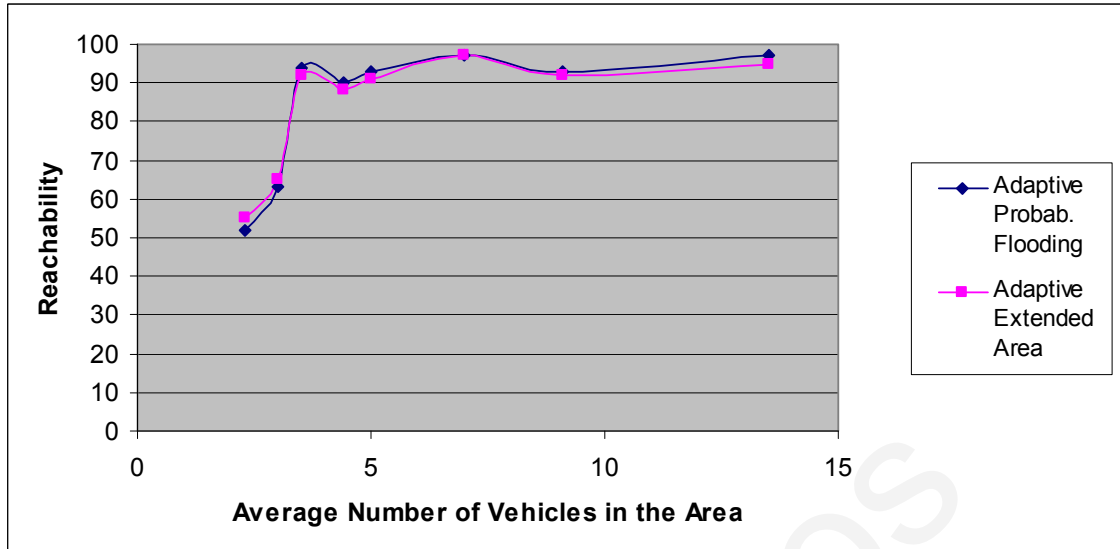


(a) Reachability

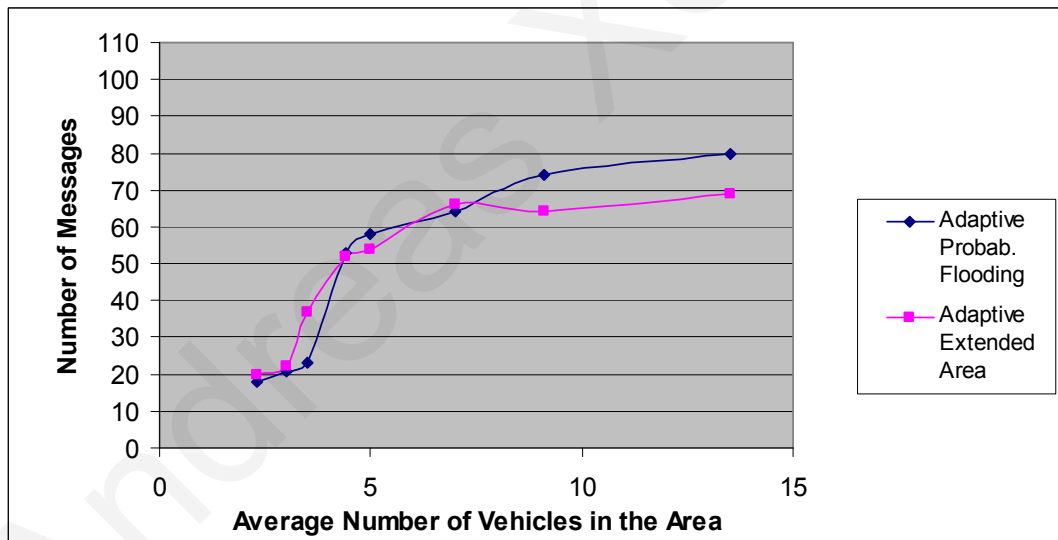


(b) Number of Exchanged Messages

Figure 27: Comparison of the proposed adaptive probabilistic flooding hovering scheme with the adaptive version of the hovering protocol proposed in [37]. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area H for different average number of vehicles in the area.



(a) Reachability

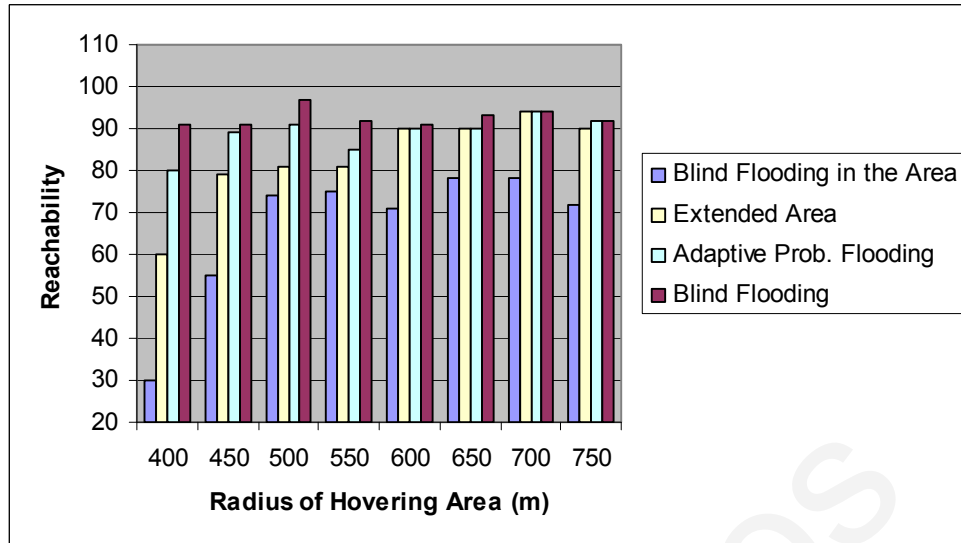


(b) Number of Exchanged Messages

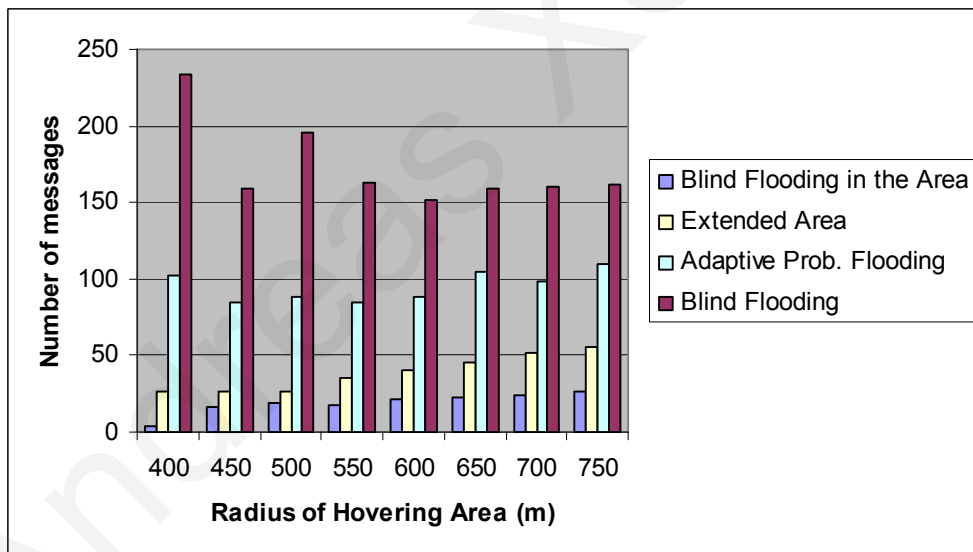
Figure 28: Comparison of the proposed adaptive probabilistic flooding hovering scheme with the adaptive version of the hovering protocol proposed in [37]. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area G for different average number of vehicles in the area.

The analysis presented in this chapter is based on a fixed circular hovering area with radius equal to $500m$. The specific size can be considered adequate for a number of applications including warning messages generated by car accidents, road constructions and collision avoidance, and also for non-safety applications such as advertisements, available pharmacies, gas prices of local petrol stations etc. However, other hovering area sizes may be considered in practice and so we examine the robustness of the proposed algorithm with respect to the size of the hovering area.

In particular, we examine the performance of the proposed scheme for circular hovering area sizes whose radius attains values in the the range $400m$ to $750m$ in steps of $50m$. The performance is evaluated in terms of the reachability achieved and the number of messages exchanged. For each hovering area the proposed scheme is compared with the performance of three other candidate protocols: blind flooding scheme in the entire network, blind flooding in the hovering area only and blind flooding in an extended area beyond the hovering area. In all simulation experiments the traffic density is in fairly constant in the range 5 - 8 vehicles per Km^2 . The simulation results are summarized in figure 29.



(a) Reachability



(b) Number of Exchanged Messages

Figure 29: Comparison of the four techniques achieving information hovering. The comparison is made in terms of the reachability achieved and the number of messages exchanged in area A for different values of the radius of the circular hovering area

The results indicate that for all hovering area sizes, the proposed scheme achieves high reachability values comparable to the ones of blind flooding in the entire network. In addition, it manages to significantly reduce the number of messages exchanged. The other schemes can further reduce the number of exchanged messages, at the expense, however, of lower reachability values at small hovering areas. Since high reachability is critical, the proposed adaptive probabilistic scheme can be considered to exhibit superior performance for these hovering areas. It must be noted that for large hovering areas the proposed scheme continues to perform well. However, it can be outperformed by the extended area approach, which can further reduce the number of exchanged messages. The reason for this is that, as the hovering size increases the simulation model (which is restricted in total size), it does not allow the proposed scheme to explore all the potential paths that could lead to increased reachability. So, had the simulation model been larger it would have demonstrated the superiority of the proposed scheme.

3.6 Conclusions and Sum Up

In this chapter we address the Information Hovering problem in VANETs. The problem naturally applies in many applications where useful information needs to be made available to all vehicles within a confined geographical area for a specific time interval. Similar to many problems in VANETs, the performance of the information hovering protocol is affected by the traffic density of the considered transportation network. In cases of low traffic density, partitioned uninformed areas may lead to low reachability. In this work, we propose a novel scheme that overcomes this problem by applying probabilistic flooding outside the hovering area based on a probability function using simulation among a number of candidate implementations. Informed vehicles outside the area can make use of possible paths that may inform vehicles in partitioned areas of the hovering area leading to high reachability. Rebroadcast probability outside the hovering area is adaptively

regulated based on estimates of the vehicle density within the hovering area. The estimation of the vehicle density is based on an analytically derived function. Simulation results indicate that the proposed protocol is successful in satisfying its design objectives and that it outperforms other candidate hovering protocols appearing in the literature. We also demonstrate that our contribution goes beyond the proposed information hovering protocol and extends to the introduction of a design procedure that can be used to design a class of density adaptive hovering protocols. Finally, we check the robustness of our proposed solution on different sizes of hovering area. The results indicate that our traffic adaptive probabilistic flooding achieves high reachability with a reduced number of messages exchanged.

Chapter 4

Information Hovering: A New Approach for Performance Acceleration of Wireless Push Systems

4.1 Introduction

Ubiquitous inter-networking has enabled access to an ever-growing amount of data by a steadily increasing number of clients. As the client-base expands, similarities in user preferences become more noticeable. Data broadcasting offers a comprehensive way of exploiting this fact, enabling the realization of efficient, bandwidth preserving dissemination schemes. Broadcast scheduling, the process of serializing the data delivery in an optimal way, constitutes a well-founded field with every-day applications, such as television, radio and teletext scheduling. While broadcasting favors wireless transmission, it has been incorporated successfully in wired scenarios as well. Well known applications include broadcasting for data availability advertisement and cache scheduling in named content networking [51].

Wireless data broadcast scheduling typically assumes a single frequency or cellular network, which covers a densely populated area. The clients therein are assumed to be interested in a common set of discrete data items. Each item is associated with a request probability and a size

measured in bytes. While each item's size is static and given, its request probability varies with time. In order to approximate the probability distribution of the items, the clients are usually required to provide some lightweight feedback regarding their preferences, either explicitly or indirectly. In [86], for example, the clients are required to emit a single pulse upon reception of a wanted item. The aggregate received power level is then mapped to the probability distribution by a learning automaton [91]. Once the item probabilities become known, a central authority creates a broadcast schedule that optimizes a given criterion. Examples are the minimization of the mean waiting time in [117] and of the mean incurred impatience in [97]. The schedule is transmitted over the covered area, and the clients retrieve needed data items in a streaming fashion.

While wireless broadcasting offers perfect scalability in terms of served clients, its performance does not scale as well with the increase of the amount of data. Furthermore, broadcast systems with large spatial coverage are typically assigned narrow bands in low frequencies. For example, according to [115], channel 2 for national TV broadcast in the U.S. is assigned a 6MHz band starting at 54MHz, while the amateur zone at 2.4GHz is assigned a band of 17MHz. In addition, the achievable mean waiting time in a broadcasting system is known to have an analytical lower bound [117]. The emerging challenge can be expressed as follows:

Is it possible to improve the performance of a wireless broadcast system via architectural modifications or client networking?

Architectural modifications include the deployment of multiple, spatially distributed broadcast systems, each one handling a subset of the original data [86]. In a similar fashion, data subsets can be assigned to different frequency channels, as in [133]. Client networking has been proposed as another alternative. In this case, data items are retrieved via collaborative caching and P2P networking [66, 77, 134]. Architectural modifications and client networking typically imply backbone and client device upgrades, extra bandwidth or strict locality of demand. With the

utilization of the Information Hovering scheme proposed in [128], users entering or present in this location have rapid access to the hovered data. Due to its simplicity, information hovering constitutes a viable, cost effective alternative over classic P2P networking.

In this chapter we present a novel collaborative scheme between centralized wireless broadcasting and user networking (information hovering) [72]. Previous state-of-the-art works have treated user networking simply as a means of extending the coverage of wireless access points [134]. In [66] this concept is repeated and network coding at symbol level is employed for boosting the throughput of the users' network. In contrast, our work [72] proposes parallel data dissemination through centralized, optimal broadcasting and through user networking (information hovering) at the same time. Thus, a two-tier architecture is formed: at tier 1 (broadcasting) data is disseminated through optimal periodic scheduling. The scheduling considers time-variant data content and user preferences. At tier 2 data with locality of demand are hovered around time-variant anchoring points. A user inside the hovering zone can then retrieve the information needed before its next scheduled broadcast (tier 1). Data dissemination load can be freely shared between the centralized broadcasting and the hovering, following a pull-push balancing model. As an additional contribution, the classic balancing method of "optimal cut-off point" [31] is shown to produce suboptimal results. The analytically optimal alternative is also presented.

4.2 The Proposed Collaborative Scheme

The goal of the present study is the collaborative data dissemination through broadcasting and information hovering. The methodology consists of the following steps:

- Represent the collaboration as a push/pull balancing problem. Then, define analytically the load of data to be handled by the hovering system, minimizing the mean waiting time of the

broadcasting system. It is assumed that the time required to answer a query via pull is trivial compared to the push alternative.

- The hovering system must be notified of the data it should handle. Therefore, a signaling protocol between the broadcasting and the hovering systems is presented. These steps are studied in paragraphs 4.2-A (*Load Sharing*) and 4.2-B (*Collaboration between subsystems*) respectively.

Notation and standard assumptions

We assume a set of N data items arbitrarily indexed by $i = 1 \dots N$. Each item i is associated with its size l_i (in *bytes*) and its request probability p_i . Therefore, it holds that $\sum_{i=1}^N p_i = 1$. No assumptions are made concerning the nature of a data item during the analysis. In accordance with the related work on scheduling [2, 26, 68, 69, 70, 71, 86, 87, 97, 117, 132, 133], an item is simply a piece of information a client may acquire through a single query. During the simulations of paragraph 4.3, items are specialized to carry information about traffic, parking availability and road conditions in VANETs covering an urban area. It is clarified that, in push-based broadcast scheduling, the term “client query” does not imply posting a request to a server, but rather waiting for the broadcast of a specific item.

According to [117], the minimum mean waiting time of a stand-alone, push-based broadcast system is given by:

$$\overline{W} = \frac{(\sum_{i=1}^N \sqrt{p_i \cdot l_i})^2}{2} \quad (16)$$

Notice that \overline{W} does not depend on the number of clients [117]. This bound is achievable in practice, as shown in [69, 117]. Finally, it is noted that broadcast scheduling and information hovering are application-layer techniques, independent of the underlying physical implementation

parameters. This is a standard assumption, in accordance with [2, 26, 58, 68, 69, 70, 71, 86, 87, 97, 117, 128, 132, 133]. The reader is referred to [65, 126] for exemplary studies regarding physical aspects of digital content dissemination via broadcasting and P2P networking respectively.

A. Load sharing

Load sharing between hovering and broadcasting implies a probabilistic approach. Push-based broadcasting assumes per item request probabilities (see equation (16)). In addition, a hovering scheme can disseminate a data item with a certain success probability, which depends on the number of clients participating in the hovering process [129]. A query posed by a mobile client will be answered by listening to the broadcast schedule, unless the hovering supplies the needed item sooner. This configuration calls for a probabilistic pull/push balancing model, which is illustrated in figure 30. The goal is to define the optimal percentage of queries that should be answered successfully by the hovering scheme.

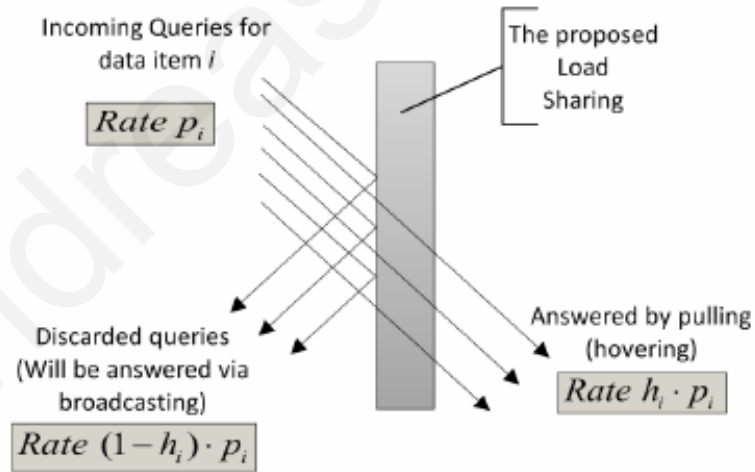


Figure 30: Illustration of the proposed load sharing between the broadcasting and the hovering systems. In a total of Q queries, $p_i \cdot Q$ refer to item i . A percentage $h_i \in [0, 1]$ of these are addressed by the hovering system, while the remaining are answered via the broadcasting.

The optimality refers to *minimizing the lower bound expressed by equation (16) for a given strain on the hovering system.*

The minimization of equation (16) expresses the fact that, since certain items may be available through external sources (i.e. hovering), their broadcast frequency can be lowered. Thus, more “air time” becomes available to other items. Notice that equation (16) regards item request probabilities and sizes only. This means that the broadcast scheduling authority must alter the original probability distribution p_i by offloading items to the hovering system, in order for the mean waiting time \bar{W} to be lowered.

The proposed load sharing cannot occur regardless of the imposed strain on the hovering system. Indeed, one could require that all clients hover all available data items with perfect efficiency. This would effectively nullify equation (16) in theory, but would imply that each mobile client had vast memory, processing power and available bandwidth. It is therefore logical to make the load sharing tunable in terms of imposed strain. Since a query for item i receives a reply of size l_i , the strain can be expressed as the mean amount of pulled data \bar{B} :

$$\bar{B} = \sum_{\forall i: \{0 < h_i < 1\}} p_i \cdot h_i \cdot l_i \quad (17)$$

where $h_i \in [0, 1]$ denotes the percentage of queries for item i that are addressed by the hovering system. An expression involving only the rate of pull queries is also introduced:

$$\bar{B}_u = \sum_{\forall i: \{0 < h_i < 1\}} p_i \cdot h_i, \bar{B}_u \in [0, 1] \quad (18)$$

The goal of the study can be defined as follows: *given a desired total strain \bar{B} , where should hovering occur and with what efficiency (i.e. define optimal values for h_i , $i = 1 \dots N$, $h_i \in [0, 1]$) in order to minimize the mean waiting time of the clients (equation (16))?*

We assume that hovering occurs instantaneously and intermediate transient phenomena are negligible. With these remarks in mind, the analysis is initiated by assuming two discrete states:

STATE 1: The starting state, in which the item request probabilities are unaltered: $p_i^{(1)} = p_i, \forall i$. This state corresponds to total absence of hovering not initiated yet. In this phase the broadcast system assumes the total data dissemination load.

STATE 2: The ending state, in which:

$$\left\{ \begin{array}{l} p_I^{(2)} \rightarrow 1 \\ p_i^{(2)} \rightarrow 0, i \neq I \end{array} \right. \quad (19)$$

In other words, this state represents the extreme case where hovering has assumed most of the dissemination load, handling all data items but item I . Notice that one could had also considered the total absence of broadcasting as the ending state. However, this would entail that the central broadcasting infrastructure has stopped functioning, for no apparent reason. In other words, the final load attributed to broadcasting can be minimal, but a null value would yield inefficient use of resources and therefore bad overall design. However, the results of the analysis remain unaltered in any case.

We proceed to introduce a set of N continuous variational flows $p_i(t), t \in [0, T], i = 1 \dots N$, which fulfill the conditions:

$$\left\{ \begin{array}{l} p_i(0) = p_i^{(1)} \\ p_i(T) = p_i^{(2)} \end{array} \right. \quad (20)$$

In this sense, expression 16 can also be rewritten as a function of the flow control variable t , by simply substituting p_i with $p_i(t)$. The same applies to 17 and 18 when substituting

h_i with $h_i(t)$ and p_i with $p_i^{(1)}$. The notation $\overline{W}(t)$, $\overline{B}(t)$ and $\overline{B}_u(t)$ corresponds to this expression. The goal is to define the flows $p_i(t)$, $i = 1 \dots N$ which minimize $\overline{W}(t)$ for any give $t \in [0, T]$.

Lemma 1. *The continuous flows $p_i(t)$, $i = 1 \dots N$ which minimize $\overline{W}(t)$, with regards to the chosen item I of STATE 2, are linear functions of $\overline{B}_u(t)$:*

$$p_i(t) = p_i^{(1)} + \overline{B}_u(t), i = I, 0 \leq t \leq T \quad (21)$$

$$p_i(t) \begin{cases} p_i^{(1)}, 0 \leq t \leq T \\ p_I^{(1)} - \overline{B}_u(t), t_o \leq t \leq t_1, i \neq I \\ p_i^{(2)}, t_1 \leq t \leq T \end{cases} \quad (22)$$

where $t_0 : \sum_{j \in O_i^*} p_j^{(1)} = \overline{B}_u(t_0)$, $t_1 : \sum_{j \in O_i} p_j^{(1)} = \overline{B}_u(t_1)$ and $\overline{B}_u(T) = 1 - p_I^{(1)}$. O_i^* denotes the ordering of all items but I by ascending $\frac{l_i}{p_i^{(1)}}$ ratio, up to (but without) item i . O_i denotes the corresponding inclusive ordering.

Proof: Assume an appropriately small interval $\Delta t \rightarrow 0^+$ for which it holds that

$$p_i(t) = p_i(0) + \left. \frac{dp_i}{dt} \right|_{t=0} \cdot \Delta t, t \in [0, \Delta t], i = 1 \dots N \quad (23)$$

For brevity, the notation $d_i = \left. \frac{dp_i}{dt} \right|_{t=0}$ will be employed. Since $\sum_{i=1}^N p_i(t) = 1, \forall t$, it is derived from (23) that

$$\sum_{i=1}^N d_i = 0 \quad (24)$$

Furthermore, assume that in this first infinitesimal step $t \in [0, \Delta t]$, the chosen flow I (referring to the sole data item in STATE 2) increases while the rest decrease:

$$d_I \geq 0 \text{ and } d_i \leq 0, \forall i \neq I \quad (25)$$

Up to this point the flow control parameter t was an auxiliary variable bounded in an arbitrary range $[0, T]$ and deprived of any physical meaning. We will now proceed to assign a specific physical context to variable t . From this point on, t will express the normalized strain imposed on the hovering subsystem and will be a synonym to the $\bar{B}u \in [0, 1]$ metric.

Therefore, Δt expresses the increase of the total rate of pulled items. Consequently:

$$\Delta t = - \sum_{i \neq I} d_i \cdot \Delta t \Leftrightarrow - \sum_{i \neq I} d_i = 1 \quad (26)$$

In other words, the increase in the rate of pulled items is equal (by absolute value) to the decrease in the rate of pushed items. From (24) and (26) we deduce that:

$$d_I = 1 \quad (27)$$

Next, we calculate $\frac{d\bar{W}}{dt}$, $t \in [0, \Delta t]$ with regard to equations (23):

$$\frac{d\bar{W}}{dt} = \sum_{i=1}^N \frac{d_i \cdot \sqrt{l_i}}{2 \cdot \sqrt{p_i(0) + d_i \cdot \Delta t}} \quad (28)$$

The goal is to make \bar{W} reduced as fast as possible, i.e. to find the $d_i, i \neq I$ values which minimize (28). Consider the ordering O_i of all items by ascending $\frac{p_i}{l_i}$ ratio, an arbitrary item k in the ordering of which it holds:

$$\frac{l_k}{p_k(0)} \geq \frac{l_i}{p_i(0)}, i \neq I \quad (29)$$

Let $d_k = -1$, $d_i = 0 \forall i \neq k, I$. The derivative (28) then yields the value:

$$V = \frac{\sqrt{l_I}}{2 \cdot \sqrt{p_I(0) + \Delta t}} - \frac{\sqrt{l_k}}{2 \cdot \sqrt{p_k(0) - \Delta t}} \quad (30)$$

We will prove through reduction that $\frac{d\bar{W}}{dt} \geq V$, for any $d_i, i \neq I$ values abiding by the statement (25):

$$\begin{aligned} \frac{d\bar{W}}{dt} \geq V &\Leftrightarrow \sum_{i=1}^N \frac{d_i \cdot \sqrt{l_i}}{2 \cdot \sqrt{p_i(0) + d_i \cdot \Delta t}} \geq \dots \\ &\dots \frac{\sqrt{l_i}}{2 \cdot \sqrt{p_i(0) + \Delta t}} - \frac{\sqrt{l_k}}{2 \cdot \sqrt{p_k(0) - \Delta t}} \Leftrightarrow \end{aligned} \quad (31)$$

$$\begin{aligned} \sum_{i \neq I} \frac{d_i \cdot \sqrt{l_i}}{2 \cdot \sqrt{p_i(0) + d_i \cdot \Delta t}} &\geq - \frac{\sqrt{l_k}}{2 \cdot \sqrt{p_k(0) - \Delta t}} \Leftrightarrow \\ \text{via (25): } \sum_{i \neq I} \frac{d_i \cdot \sqrt{l_i}}{2 \cdot \sqrt{p_i(0) + d_i \cdot \Delta t}} &\geq \left(\sum_{i \neq I} d_i \right) \frac{\sqrt{l_k}}{2 \cdot \sqrt{p_k(0) - \Delta t}} \Leftrightarrow \quad (32) \\ \text{via } \Delta t \rightarrow 0: \sum_{i \neq I} \frac{d_i}{2} \left(\sqrt{\frac{l_i}{p_i(0)}} - \sqrt{\frac{l_k}{p_k(0)}} \right) &\geq 0 \end{aligned}$$

which holds always, because of (24) and (29). Therefore, $\frac{d\bar{W}}{dt} \geq V$ for any $d_i, i \neq I$, *QED*.

Notice that the best possible values ($d_I = 1$ and $d_k = -1$, $d_i = 0$, $i \neq k, I$) do not present any dependence on t , reminding that t expresses the normalized hovering strain. Therefore, we proceed to the next infinitesimal Δt step and repeat the process (continuity assumption).

Observe that the place of k in the $\frac{l}{p}$ ordering remains unchanged. Repeat the process until $p_k = 0$, i.e until $t = p_k(0)$. At that point, consider that the original set of N items has been reduced to $N - 1$, i.e. the now improbable item k has been removed. Name this state

as STATE 1 and repeat the proof recursively until only item I remains. This concludes the proof of Lemma 1.

Notice that the flows defined in Lemma 1 can be directly expressed as functions of \bar{B}_u only, discarding the redundant variable t . This was expected, since t was introduced only for facilitating the analysis. For the remainder of this work, the expression of flows will follow this convention.

Lemma 1 refers to flows with respect to the chosen item I of *STATE 2*. For a given strain \bar{B}_u we examine all possible values of $I = 1 \dots N$ and keep the one which produces the lowest mean waiting time (16). This optimal value, I^{opt} , corresponds to optimal flows, which will be denoted as $p_i^{opt}(\bar{B}_u)$. Concerning the optimal rate $h_i^{opt}(\bar{B}_u) \cdot p_i^{(1)}$ of queries that should be answered via pulling (see figure 30), it holds that:

$$h_i^{opt}(\bar{B}_u) \cdot p_i^{(1)} = p_i^{(1)} - p_i^{opt}(\bar{B}_u) \quad (33)$$

Equation (33) essentially states that for each item i , the queries answered via the hovering are equal to the total number of queries for item i , minus the ones answered via the centralized broadcasting. Finally, solving (33) for h_i^{opt} yields:

$$h_i^{opt}(\bar{B}_u) = 1 - \frac{p_i^{opt}(\bar{B}_u)}{p_i^{(1)}}, \forall i : \{p_i^{opt}(\bar{B}_u) \leq p_i^{(1)}\} \quad (34)$$

which expresses the required hovering efficiency for data items $i = 1 \dots N$ and, assuming locality of demand, at site $i = 1 \dots N$ which minimizes the client mean serving time, given a desired strain \bar{B}_u .

As a closing remark, Lemma 1 states that the normalized strain \bar{B}_u affects the $p_i(t)$ quantities of equation (22) according to the O_i^* ordering in a serial manner. Starting from STATE

1, the initial increase of \bar{B}_u will affect the first (e.g. k -indexed) item of the ordering until $p_k(\bar{B}_u) = p_k^{(2)}$, then the second item, and so on. From equations (17) and (18) it is easily seen that, for any item k that is the sole one currently affected in the ordering, it holds:

$$\frac{\partial \bar{B}}{\partial (p_k h_k)} = l_k, \frac{\partial \bar{B}_u}{\partial (p_k h_k)} = 1 \Rightarrow \frac{\partial \bar{B}}{\partial \bar{B}_u} = l_k > 0 \quad (35)$$

Equation (35) states that the relation between \bar{B} and \bar{B}_u is strictly increasing, and therefore 1-1. Any value of \bar{B} can thus be converted to one unique, corresponding \bar{B}_u value, and vice versa. The conversion is straightforward, since the quantities p_i, l_i of equations (17) and (18) are static and known, and h_i are calculated through equation (34). Therefore, the real strain, \bar{B} , and its normalized version, \bar{B}_u , are interchangeable in all equations. For the sake of proper presentation we will assume the \bar{B}_u format, since it is always bounded in $[0, 1]$ regardless of the data item sizes.

B. Collaboration between subsystems

The preceding analysis considered the generic case of hybrid (pull/push) data dissemination. The conclusions of Lemma 1 and equation (34) are adapted to the case of collaboration between broadcasting and hovering as follows.

When initiating operation, the broadcast scheduling authority is unaware of the per item preferences. The request probability distribution of the items is approximated through an adaptation procedure, which has been analyzed in detail in [67, 86]. As discussed in [86], upon reception of a desired item, a client emits a single pulse designating content approval. The learning automaton [91] of [86] processes the aggregate feedback from all clients in real-time and re-approximates the popularity distribution of the data items. There are no distinct cycles of adaptation and subsequent rescheduling. The employed scheme of [86]

schedules the broadcast of the items based on their attributes (size l_i , approximate popularity p_i) and the time interval passed since their last broadcast (T_i). Before each broadcast the following array is constructed:

$$R_i = T_i^2 \cdot \frac{p_i}{l_i}, i = 1 \dots N \quad (36)$$

The item i that produces the maximum R_i entry is broadcasted and the current timeslot is marked as its last broadcast time moment. The p_i values are renewed by the learning automaton upon the reception of new feedback. Notice that neither the number of items nor their content are static. As in [86], when new items are added their corresponding T_i values are set to the current time slot, while their initially assumed request probability p_i is set to the median of the existing probability distribution. Normalization is applied to ensure that the sum of the approximate probability distribution is equal to one. The scheduling authority may also update the content referring to the geographical site i , changing its size l_i , or remove an item entirely. The procedure is repeated and the automaton converges to the real probability distribution.

The number of users and the influence of its variations on the convergence of the adaptation has been studied in [117]. It was proven that for large user sets all clients can be collectively handled as a simple Gaussian process. This outcome, which is expected due to the central limit theorem, typically holds in VANETs, since they typically contain a large number of users. Consequently, [117] showed that their actual number is irrelevant, as long as the popularity of each data item becomes known. The latter is the case of the present study.

The scheduling authority initiates or terminates information hovering through broadcast signals. These broadcast signals indicate the coordinates of the center of the Hovering

Area (X,Y), the radius, and also the TTL. Upon successful convergence of the adaptation process, the scheduling authority proceeds to calculate the $h_i^{opt}(\bar{B}_u)$ and $p_i^{opt}(\bar{B}_u)$ values via equation (25) and Lemma 1, for a given strain \bar{B}_u . The value $h_i^{opt}(\bar{B}_u)$ expresses the efficiency with which the hovering subsystem should handle queries for item i . As shown in [129] however, this efficiency is a function of the percentage of clients (H_i) contributing in the hovering process:

$$h_i = f(H_i) \quad (37)$$

As in [129], the form of the function $f(\cdot)$ depends on the covered area, the hovering range, the employed routing protocol and the mean density of vehicles. Since these parameters can be assumed to be known, the scheduling authority can calculate the optimal percentage of clients, H_i^{opt} that should contribute to the hovering of item i , in order to achieve an efficiency of $h_i^{opt}(\bar{B}_u)$. The values H_i^{opt} , $i = 1 \dots N$ are then broadcast in a single package to all clients. For each entry i , a client *flips a weighted coin* with H_i^{opt} probability of contributing to the hovering of item i . The successful clients initiate the hovering process. Hovering termination occurs with the broadcast of a specific message from the central authority, at any appropriate situation (e.g. item replenishment or invalidation). Finally, the scheduling authority re-adapts the broadcast schedule to comply with the $p_i^{opt}(\bar{B}_u)$ probabilities.

Notice that precise knowledge of equation (37) is not crucial to the system operation. Hovering can easily achieve 100% efficiency [129] in almost any case. If (37) is imprecise, the learning automaton will converge to a point that is either below or above the expected value of the mean waiting time (equation (16)). The scheduling authority can then simply increase or decrease the strain in order to get the desired result.

4.3 Simulation

The procedure described in Section 4.2 is validated through simulation in a VANET environment. The proposed scheme is compared to the CodeON [66] and CodeTorrent [63] approaches, in terms of mean client waiting time and user satisfaction [134]. The transient phenomena of P2P network establishment and efficiency, as well as the adaptive capabilities of the feedback mechanism are addressed. All simulations were performed using the VISSIM software package. The implementation of the network coding technique of [63, 66] was based on the freely available source code of [35].

4.3.1 A. Simulation Setup and Results

The simulated VANETs refers to the Bellevue and Redmond area in Washington, depicted in figure 31 (same area as in figure 12). The dimensions of the area are 4.5 x 2.5 Km. At any given moment, approximately 3,000 vehicles (clients) roam the arterial streets and the freeway, their exact number being time-variant. The speed of each vehicle varies indicatively from 20 km/h to 80 Km/h. The relation between the H_i and h_i parameters is obtained statistically from [129] as:

$$h_i = \begin{cases} -2556 \cdot H_i^2 + 101.8 \cdot H_i, & H_i \in [0\%, 2.2\%] \\ 1, & H_i \in [2.2\%, 100\%] \end{cases} \quad (38)$$

In all tests, a central broadcasting system is assumed to cover the studied area, enabling every vehicle to retrieve data from the broadcast stream with a guaranteed rate of 25 KBps. Notice that typical choices for physical implementation of the broadcast system (3G, DVB/H) supply bitrates well beyond 2-5 *Mbps* [60]. However, we choose this particularly low value to make up for the global coverage and bitrate guarantee assumptions. Finally, the probability of successful retrieval of data from the broadcast stream and correct client feedback reception is set to 0.95.

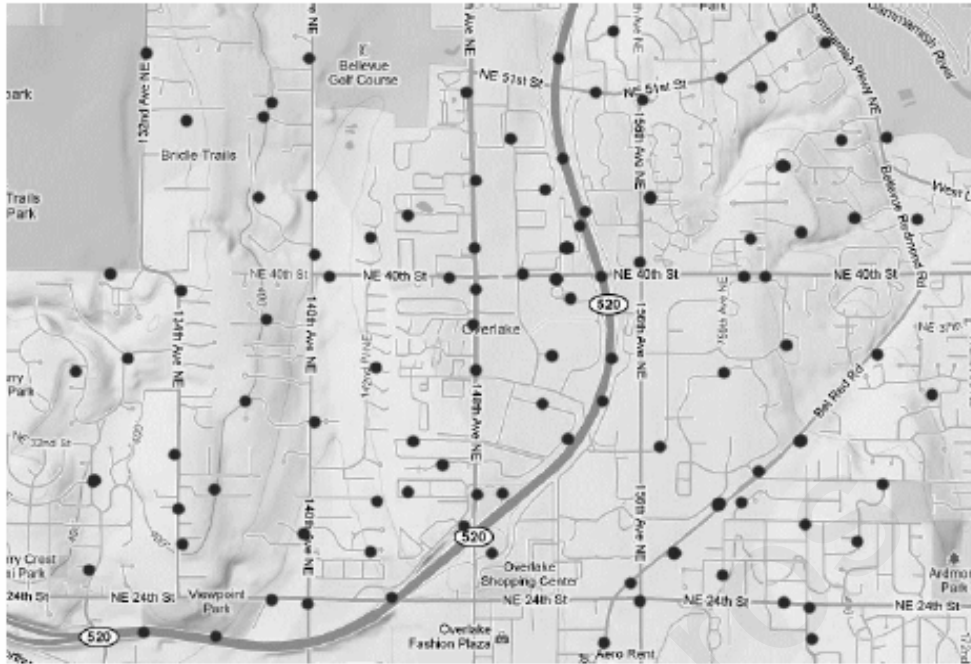


Figure 31: The simulated VANET area of Bellevue and Redmond, Washington, as depicted in Google Maps (Coordinates: 47.626413, -122.149944). A number of events (dots) regarding traffic, road condition and parking availability are selected. Information on these events is disseminated via broadcasting (global coverage) and hovering around the corresponding locations.

Concerning the physical attributes of the P2P networking, we assume that each vehicle has a varying transmission range of $180 \pm 40\text{m}$. The range is randomized prior to any transmission. Inside this range, and assuming MAC-layer clearance, the probability of successful data reception is 0.95. In all cases we adopt the MAC protocol of the 802.11 specification. Carrier sensing simulation is based on the two - ray ground propagation model. The reader is directed to the NS-2 [111] implementation of the 802.11 specification for the definition of additional attributes. The Rx/Tx rate of each client is 11Mbps .

Regarding the nature of the data items and their geographical sites, we assume N random anchoring points on the terrain of figure 32, P2P networking occurs inside a radius of 500m around each point. Each site $i = 1 \dots N$ (arbitrarily indexed) defined in this way is associated with a piece of information with size $l_i \in [1, 50] \text{Kbytes}$. This emulates a typical web page that contains

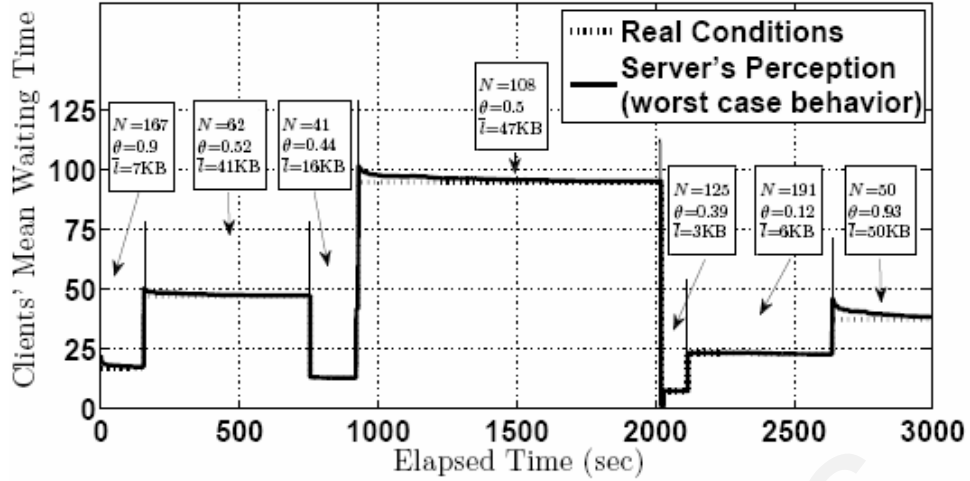


Figure 32: Performance of the learning automaton-based adaptation technique, which is used for approximating the time-variant popularity of a dynamically formed data set. Random changes are set to occur in the cardinality of the studied data set and the popularity and size of the contained items. Even in the worst-case scenario (no content correlation between successive data set updates), the automaton adapts optimally in 50 - 250 sec.

tweets on all local events of interest. Each site i , and therefore each associated data item, has a time-variant popularity among the clients, $p_i \in [0, 1]$, which is defined as the portion of the total clients interested in site i . For simplicity, each client may have only one request pending at any given time. Therefore, $\sum_{i=1}^N p_i = 1$.

In order to emulate locality of demand conditions, the data requests are assigned to the clients in a way that follows a Gaussian distribution around the anchoring point. When $p_i \rightarrow 1$, clients on the terrain are interested in site i , regardless of their current position. When $p_i \rightarrow 0$, 99.7% of the interested clients (three standard deviations) are already inside the P2P site. Intermediate values are set proportionally. In order to generate requests and assign them to clients, we set the site popularity distribution p_i according to the ZIPF distribution [94]. The skewness of this distribution is regulated through a single parameter, $\theta \in [0, \infty)$. When $\theta = 0$, the distribution is flat and $p_i = \frac{1}{N} \approx 0, \forall i, N \geq 100$. Therefore, $\theta = 0$ corresponds to high locality of demand. As θ increases, the popularity of several sites becomes higher, and the locality of their demand drops.

Consequently, the manipulation of θ is used for testing various cases of locality of demand. Finally, each simulation typically considers 600,000 client queries generated according to the distribution in use.

The popularity distribution p_i is initially not known to the broadcast scheduling authority. In the experiment of figure 32 we examine the efficiency of the employed adaptation technique. In random time moments designated by the thin vertical lines, the data set in question changes in cardinality (N), content popularity (θ) and content size (mean item size \bar{l} given in *Kbytes*). We assume the worst case, with the data set being completely renewed each time. Likewise, the server has no initial, rough info on the form of the p_i distribution, and assumes a flat one in any case. The dotted line expresses the time-variant lower bound of mean waiting time (equation (16)), should the p_i distribution be known. The solid line represents the time-variant product of the automaton, based on the current p_i approximation. Even in this worst case approach, the automaton proves its efficiency, as in [86, 91]. The convergence time for the current setup is shown to be 50 - 250sec. Note that this effort refers to measuring the popularity of *geographical sites* in a city, since cities are known to change insignificantly over time (months or years are required for the construction of new main roads, malls, parking points and other significant points of interest). Therefore, for the remainder of the simulations it will be assumed that the 1-5 minutes required for convergence have passed, and the study will focus on the behavior of the compared approaches in the steady state. The number of sites/data items will be kept constant at $N = 100$.

In figure 33 we examine the validity of the preceding analysis regarding the pull-push load balancing optimization. Starting with zero client strain (no hovering initiated yet) the broadcasting authority begins to signal the clients to initiate the hovering process, gradually increasing their load. The achieved mean waiting time in the steady state is then logged and plotted versus the corresponding strain. Three observations can be made: firstly, the experimental results follow the

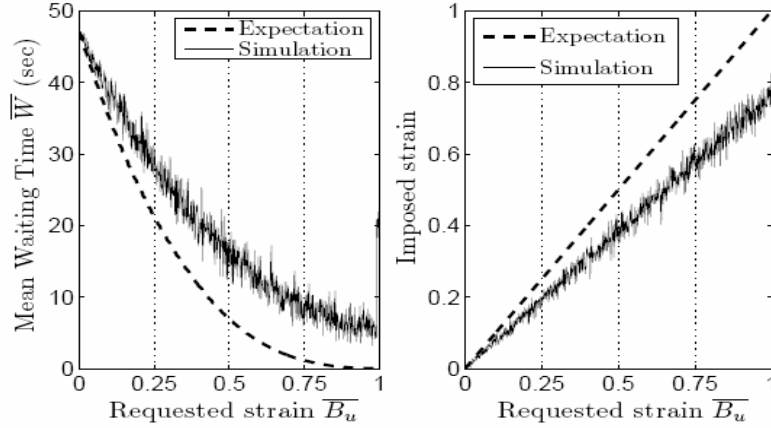


Figure 33: Plots of the requested strain \bar{B}_u , versus the achieved mean waiting time \bar{W} and the actual imposed strain. The plots reflect the case of $\theta = 0.1$, i.e. a slightly skewed ZIPF distribution of the items' popularity. The waiting time is reduced by 10 times when compared to the case of stand-alone broadcasting ($\bar{B}_u = 0$). While the reduction follows the general theoretical form, some degree of divergence is present. Results are identical for other values of θ .

general form of theoretical expectations. Secondly, the average waiting time is never nullified, even for maximum strain. This is expected, since the analysis assumes *ability* to handle the total load. However, the hovering process has a 500m of activity radius. Consequently, there will always exist clients that require an item but are outside the corresponding P2P zone. Thirdly, the requested strain is not upheld. The finally incurred strain is less than the expectation. This phenomenon is attributed to the imprecision of equation (38), which is proven to be optimistic: a requested strain value causes less real strain. This issue also highlights the fact that the performance of the system does not depend on the precision of the formula (38). The central scheduling authority is aware of both the theoretical curves (through direct definition) and the experimental curves (through the client feedback mechanism) of figure 33. Therefore, the divergence can be directly quantified, and the central authority may request less or more strain than the expectation of (38), achieving the desired result. In other words, equation (38) simply serves as an initial guide.

The proposed approach offers tunability in terms of imposed client strain but CodeOn and CodeTorrent do not. Therefore, it is required to define a standard strain for fair comparison.

Thus, for the remainder of the simulations, 95 out of 100 items (by ascending $\frac{p_i}{l_i}$ ratio) will be deterministically assigned for full dissemination via P2P networking, while the remaining 5 items will be handled explicitly via the central broadcasting.

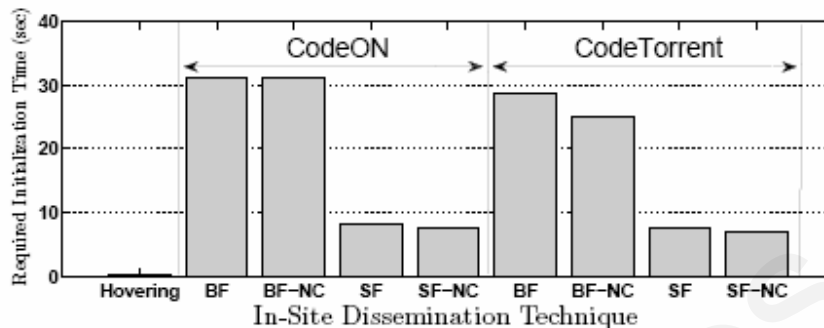


Figure 34: Time required for the initialization of the compared P2P schemes in an area of interest. Hovering requires a trivial amount of time before being established. CodeOn and CodeTorrent may be viable, requiring approximately 5sec for establishment in the best case.

The proper initiation of the P2P networking is of crucial importance, since the highly dynamic VANETs environment may hinder the employed schemes from operating efficiently. In figure 34 we consider 210 vehicles (typically present in a site) in a rectangular formation. We assume immobility, and that a randomly selected client initiates the P2P networking, seeking to disseminate a single item to the whole client set. The Hovering, CodeON and CodeTorrent P2P schemes are compared in terms of time required for the task. The hovering technique, relying simply on probabilistic flooding, achieves minimal times. The number of seeding users increases exponentially, yielding overall completion times of even less than 1sec in the employed setup. The CodeON and CodeTorrent approaches on the other hand, fall back due to the requirement for explicit packet routing. The tested approaches include blind forwarding (BF) and selective forwarding (SF) [35]. Network coding (NC) is used for maximizing the P2P network throughput, at the expense of added computational load on the clients. Both approaches present viable results in the case of selective forwarding. It is notable, however, that the use of network coding does not provide a significant

boost to the overall performance. This is a concern due to the high computational cost, especially for CodeON, which employs symbol-layer network coding, producing large amounts of data segments to be processed per single packet.

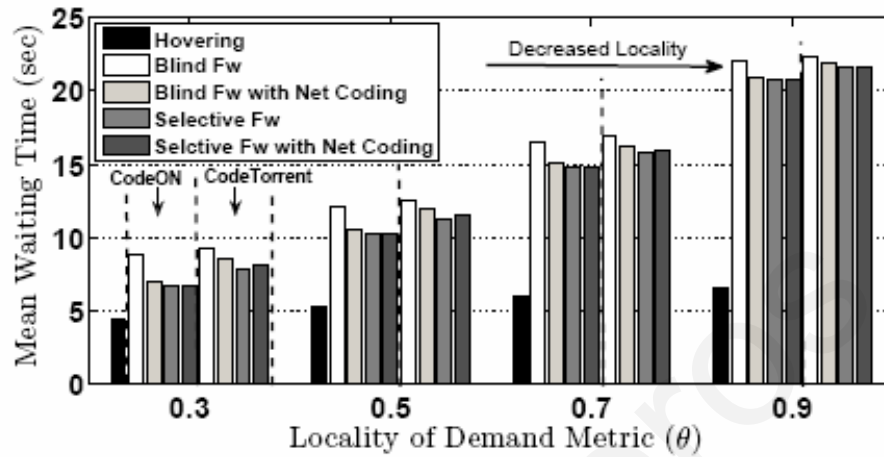


Figure 35: The mean waiting time achieved by all compared approaches in a full simulation. Steady state is assumed, with the initialization and adaptation phases having been completed. The hovering-broadcasting combination yields the best results in the tested cases of locality of demand.

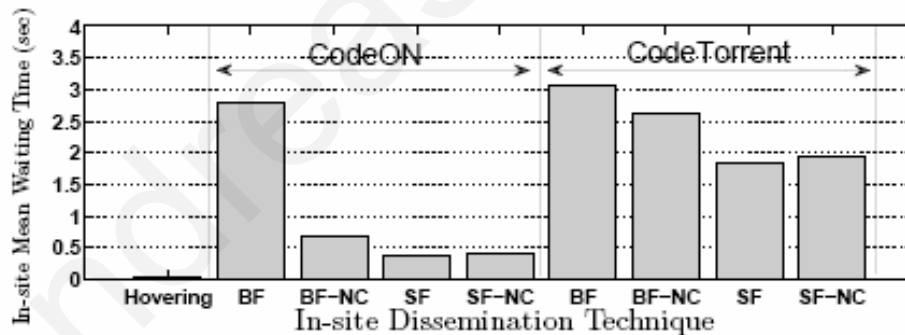


Figure 36: Average time required to obtain a data item, if the client is inside the corresponding area of interest. Hovering achieves minimal time, with CodeON and CodeTorrent following. The results also imply that network coding may not be more beneficial than an efficient routing algorithm.

Based on the requirements for initialization, we allow for a 30sec interval to pass before posing client queries. In the full-scale simulation of figure 35 the achieved mean waiting time over 600,000 client queries is presented. The locality in the demands of the clients is gradually

decreased, and the performance of each opposing scheme is logged. The broadcasting-hovering combination achieves the best results in all cases. All schemes are affected by the loss of demand locality, since a user may be far beyond the reach of the P2P networking around the area of interest. CodeON performs better than CodeTorrent, since it enforces smaller data segment sizes, limiting the impact of packet losses. However, it is once again shown that network coding is not actually more beneficial than an efficient routing algorithm in a highly dynamic VANET. This fact is more apparent in figure 36, where we measure the mean waiting time for clients inside their area of interest. Hovering-achieved times are tantamount to data pulling, as assumed. Network coding however, produces benefits only compared to blind packet forwarding. Selective forwarding is not positively affected by the added data segmentation overhead imposed by the network coding technique.

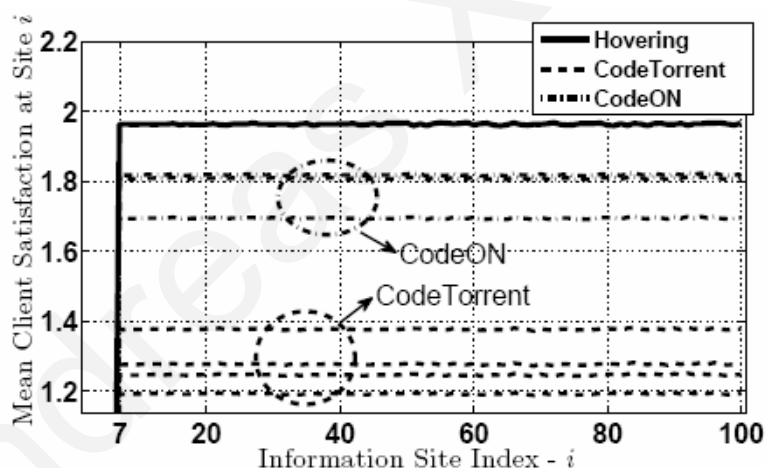


Figure 37: Alternative representation of the performance of the opposing schemes in terms of user satisfaction (try-best metric of [134]). The hovering-broadcasting combination achieves the best results, for all items $i = 6 - 100$ chosen for dissemination through P2P networking.

The results of figure 35 can also be expressed through the metric of user satisfaction [134]. The satisfaction metrics are simple transformations that receive the mean waiting time as an input and magnify the importance of certain ranges of values. Figure 37 refers to the try-best metric,

which is valid for items that should be disseminated preferably within 5-10sec. It is defined as $s = e^{-w/2}$, w being the waiting time and s the user satisfaction. We examine the average satisfaction for each data item (and thus each site) separately. As expected from the preceding results, the proposed hovering-broadcasting combination achieves higher user satisfaction than any version of the opposing schemes. This conclusion holds for all data items ($i = 6-100$) that were chosen to be disseminated through P2P schemes.

4.4 Conclusions

A novel collaborative scheme between data broadcasting and information hovering has been proposed. The clients in a realistic VANET had their mean waiting times lowered, provided that they participated in the hovering of information in the vicinity of a given area. This result demonstrates that lightweight user collaboration is sufficient for significant improvement of the system's overall performance. It also demonstrates the universality of the information hovering protocol in the sense that it can be successfully integrated in applications beyond the ones that it was originally designed for.

Chapter 5

Information Propagation Probability on Intersections in VANETs

5.1 Introduction

The majority of the applications in VANETs have messages exchanged by vehicles that are characterized by a finite lifetime period, after which their level of usefulness is greatly reduced. Thus, an important problem in VANETs is to find efficient ways to disseminate information on the target areas before this deadline expires. The problem of information dissemination has been extensively studied in the literature and a number of solutions have been proposed. However, no previous work has addressed the problem of calculating the probability to propagate information within a certain amount of time among vehicles on intersecting roads where no static infrastructure, such as repeaters, is used. In our work in [130], we concentrate on the problem of information propagation on intersections, considering only one-way traffic, and we derive a formula which gives a lower bound on this probability. We show that the propagation probability is strongly related to the traffic conditions of the road where the information is to be transmitted. We use the derived formula to estimate, via simulations, the minimal conditions required to ensure that information propagation occurs with high probability on intersections.

5.2 Problem Formulation

We first introduce the formulation of the problem which we study, and then we present the basic notations, definitions and assumptions.

We consider a roadway network which consists of a set of intersections $I = I_1, I_2, \dots, I_w$, where I_j denotes the j^{th} intersection and w is the total number of intersections that exist in the network. These intersections are interconnected by a set of straight line roads. The road connecting intersection I_j with intersection I_k is denoted by h_{jk} . The roadway network accommodates a number of vehicles. Vehicle i is denoted by veh_i . Note that even though they may be intersections with more than two roads, this is very uncommon and we will ignore. In any case, if more than two roads intersect the propagation probability will improve, and thus will not affect the lower bound.

We assume one way traffic along the roads and vehicles on road h_{jk} travel from intersection I_j to intersection I_k . We study the information propagation to the direction of the traffic. When we refer to vehicles we refer to instrumented vehicles able to participate in VANETs. Also, we assume that all vehicles have constant transmission range denoted by r , which is the same for all vehicles. *Informed vehicles* are vehicles that have the information while *uninformed vehicles* are vehicles that do not. On each road h_{jk} , a vehicle travels with a constant speed that is selected uniformly and independently from the interval $[v_{min}(h_{jk}), v_{max}(h_{jk})]$. Vehicles move independently at their chosen velocity.

The number of vehicles entering a road h_{jk} is assumed to be a stochastic variable and the corresponding stochastic process is modeled as a Poisson process. Several experiments have shown that the outcomes of such a model are in good agreement with real measurements obtained in practise [93] and is also adapted by simulation tools.

The probability density function of the arrival process at road h_{jk} is thus given by the following formula:

$$P_z^{h_{jk}}(t) = \frac{(\lambda_{jk}t)^z}{z!} e^{-(\lambda_{jk}t)} \quad (39)$$

where λ_{jk} denotes the mean arrival rate at road h_{jk} and z denotes the number of arrivals in the time interval 0 to t . The equation describes the probability of seeing exactly z arrivals in the period of time from 0 to t .

Without loss of generality, for the remainder of our analysis we consider a segment of the roadway network with one way roads, as shown in figure 38.

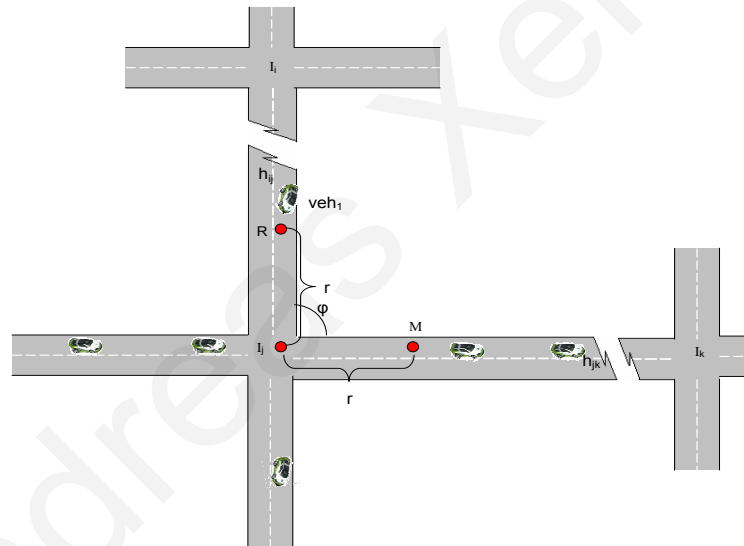


Figure 38: Representation of an intersection in a road network presenting the basic notations

The segment presented in figure 38 includes intersection I_j and the roads h_{ij} and h_{jk} interconnect intersections I_i with I_j and I_j with I_k respectively. The angle between roads h_{ij} and h_{jk} is denoted by φ . R is the point on h_{ij} that is r apart from intersection I_j and M the corresponding point of road h_{jk} . Finally, veh_1 is the *head of the information* (meaning that there are no other informed vehicles ahead of it) on road h_{ij} and is traveling with speed V_1 . We start counting time,

$t = 0$, at the point where veh_1 is at a distance less than r from intersection I_j which means that it is able to transmit the information to road h_{jk} . This can happen by either having veh_1 get the information before passing point R ($t = 0$ when veh_1 is at point R) or having the information transmitted to veh_1 by a following vehicle on h_{ij} , after passing point R and before reaching intersection I_j ($t = 0$ is when veh_1 receives the information).

Since we assume one-way vehicle traffic, we are interested in intersections where veh_1 has the opportunity to choose among two or more roads. Otherwise, if h_{jk} was the only choice, the probability of message propagation would be equal to 1 since veh_1 will definitely enter road h_{jk} .

There are two ways to propagate information from vehicles of road h_{ij} to vehicles of road h_{jk} at their intersection. The first way is by transmitting the information to a vehicle on h_{jk} . We call this probability $p_{tr}^{h_{ij}h_{jk}}$. The second way is the driving way and we call the probability $p_{dr}^{h_{ij}h_{jk}}$, where an informed vehicle from h_{ij} turns into h_{jk} . The probability $p_{tr}^{h_{ij}h_{jk}}$ is strongly related to the traffic characteristics of the road h_{jk} where the information is to be transmitted. Probability $p_{dr}^{h_{ij}h_{jk}}$ depends on what portion of the arrival rate road h_{ij} has, compared to the total arrival rate of the roads attached to intersection I_j . The calculation of p_{dr} is better explained in subsection 5.3.2.

5.3 Theoretical Analysis of Message Propagation Probability on Intersections

In this paragraph we provide a lower bound of the probability to propagate information from an informed vehicle of road h_{ij} to a vehicle in road h_{jk} when these two vehicles are close to the intersection I_j . The reason we concentrate on a lower bound and not on the actual probability is because the calculation of the actual probability is extremely complicated, as it needs to consider all possible propagation scenarios, even if their contribution to the overall probability is very small. In this work we concentrate on the two basic scenarios to propagate information:

- 1) by transmitting the information from vehicles on road h_{ij} directly to vehicles on road h_{jk} and
- 2) by having the veh_1 driving into road h_{jk}

The following equation gives a lower bound of the probability to propagate the information combining the two aforementioned propagation ways:

$$p_{h_{ij}h_{jk}} = p_{tr}^{h_{ij}h_{jk}} + (1 - p_{tr}^{h_{ij}h_{jk}}) * p_{dr}^{h_{ij}h_{jk}} \quad (40)$$

In the following subsections we derive the formulas of calculating the probabilities $p_{tr}^{h_{ij}h_{jk}}$ and $p_{dr}^{h_{ij}h_{jk}}$.

5.3.1 Probability of Transmission of Information among vehicles on intersecting roads

$$(p_{tr}^{h_{ij}h_{jk}})$$

Firstly, we study the different scenarios of transmitting information from vehicles on road h_{ij} to vehicles on road h_{jk} . We assume that there are no buildings to block signal transmission. By this assumption, information can be passed from vehicles at any point on road h_{ij} to vehicles on road h_{jk} which are of distance smaller than r .

As shown in figure 38, veh_1 is the head of information on road h_{ij} . We start counting time ($t = 0$) at the point where information enters road segment RI_j on veh_1 . There are two different cases that we need to consider and we study them separately. In *case 1*, veh_1 was already informed before passing from point R . In *case 2*, veh_1 has passed from point R without the information and before reaching intersection I_j , a following vehicle transmitted the information to veh_1 .

For our theoretical analysis, we compute the probability to propagate information from a vehicle traveling on road h_{ij} to a vehicle traveling on road h_{jk} in the time period $[0, y]$, where $y \leq \frac{r}{V_{max}(h_{ij})}$, which is the time needed by the fastest moving vehicle on h_{ij} to travel distance r .

By choosing this specific bound we make the analysis less complicated since veh_1 will be able to cover a distance less or equal to r . In addition, a greater value for y , even though it will increase the propagation probability, it will also increase the theoretical time that a message needs to reach its target. This is because we need to consider all the intersections that the message has to pass and add the time y of each one in the overall time up to the target area.

For each different case we need to consider two possible ways to transmit information to vehicles on road h_{jk} . The first way is to have a vehicle entering h_{jk} during time interval $[0, y]$. This vehicle will definitely catch the information from veh_1 , since it is going to be of distance smaller than r from veh_1 . The second way is to have, during the interval $[0, y]$, veh_1 's transmission range to catch up with a vehicle that has entered h_{jk} before $t = 0$. This second way can appear when the vehicles that have entered h_{jk} , are moving so slowly that, at some point, the transmission range of veh_1 (which must be moving fast enough) catches up to them. We call the probability of the first way *probability of entering* and denote it by p_e , and the probability of the second way *probability of catching up* and denote it by p_c .

Case 1. veh_1 , has the information when it passes from point R

It can be easily seen that $y \leq \frac{r}{V_{max}(h_{ij})} \leq \frac{r}{V_1}$, where $\frac{r}{V_1}$ is the time that veh_1 needs to cover distance r on road h_{ij} and V_1 is the speed of veh_1 . Thus, we need to calculate the probability to have a vehicle entering road h_{jk} (from intersection I_j) during the time interval $[0, y]$, since it is definitely going to receive the information from veh_1 . This gives us the *probability of entering* and can be determined by the following equation.

$$p_e = 1 - P_0^{h_{jk}}(y) \quad (41)$$

where $P_0^{h_{jk}}(y)$ is the probability of having zero vehicles entering h_{jk} during period y as defined in equation (39).

The calculation of the *probability of catching up* (p_c) is more complicated than p_e . We define $d(t)$ as the furthest away point from intersection I_j on road h_{jk} where the transmission range of veh_1 can cover in time t . It can be easily seen that any vehicle in between I_j and $d(t)$ is going to receive the message. Using basic trigonometric rules we provide the relation between $d(t)$, transmission range r , angle φ and V_1 , which is given by:

$$d(t) = (r - V_1 t) \cos \varphi + \sqrt{r^2 - (r - V_1 t)^2 \sin^2 \varphi} \quad (42)$$

In order to gain insights on how the value of the angle φ affects $d(t)$, in figure 39 we show a plot $d(t)$ versus φ , which is based on equation (42). The plot is obtained by fixing the variables r , V_1 and t to the values $r : 250m$, $V_1 : 11m/s$, $t : 7$ sec.

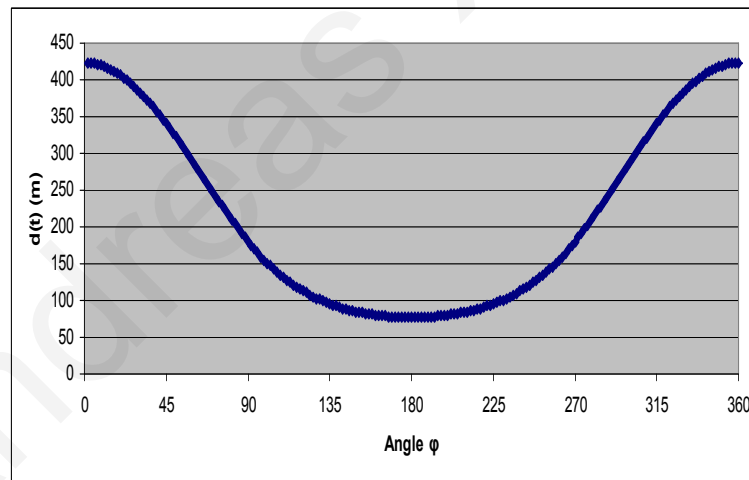


Figure 39: Influence of angle φ to $d(t)$ where the rest of the variables are constant. Their values are $r : 250m$, $V_1 : 11m/s$, $t : 7$ sec. and $X(t)$ vs time

It is obvious from the figure 39 that the transmission range of Veh_1 gives the vehicle the potential to transmit the critical message to an uniformed vehicle that has passed point M . This can happen when angle φ is less than 90° . Veh_1 can also transmit the message to a vehicle moving

on road h_{jk} before passing point R . This makes the analysis even more complicated, since for the cases when veh_1 does not pass point R (having φ less than 90°) the *probability of entering* is 0. In addition, we need to reconsider the definition of $t = 0$, since in these cases the information can be transmitted before veh_1 passes point R . For simplicity, and considering that in practise there are very few intersections with angle less than 90° , in our analysis we concentrate on angles greater or equal to 90° .

In the case where ϕ is a right angle then $d(t) = \sqrt{r^2 - (r - V_1 t)^2}$. Also, we define $X(t)$ as the distance from I_j on road h_{jk} that the vehicle closest to I_j has on time t , compared to all vehicles that have entered h_{jk} before $t = 0$.

$$X(t) = \min(V'_i * (t - T_i)) \quad i = 1, 2, \dots, Z(\gamma) \quad (43)$$

where V'_i is the speed that vehicle i has on road h_{jk} and it is uniformly distributed in the range $[v_{min}(h_{jk}), v_{max}(h_{jk})]$, T_i is the time when vehicle i passed intersection point I_j ($T_i < 0$), γ is the period of time before $t = 0$ where there is a chance the transmission range of veh_1 to catch up with the vehicles that have entered road h_{jk} . This period is equal to $\frac{r}{v_{min}(h_{jk})}$ which is the time that the slowest vehicle moving on h_{jk} needs to cover distance r . Finally, $Z(\gamma)$ is the number of vehicles that have entered h_{jk} during period $[-\gamma, 0)$.

The function distribution of $X(t)$ is given by the equation:

$$F_{X(t)}(d(t)) = \sum_{z=0}^{\infty} P[X(t) < d(t) | Z(\gamma) = z] * P[Z(\gamma) = z] \quad (44)$$

The points of time (T_i) at which $Z(\gamma)$ vehicles have entered road h_{jk} , are considered as random variables and are distributed independently and uniformly in the interval $[-\gamma, 0)$. This leads to $T_i \sim \text{uniform}(-\gamma, 0)$.

Since T_1, T_2, \dots, T_z and V'_1, V'_2, \dots, V'_z are independent and identical distributed (i.i.d), we can drop the subscripts.

Following from (44), we have:

$$P[X(t) \leq d(t)|Z(\gamma) = z] = 1 - P[X(t) > d(t)|Z(\gamma) = z] = 1 - P[V' * (t - T) > d(t)]^z \quad (45)$$

In figure 40, we plot an example of functions $d(t)$ and $X(t)$ versus time. In this example, we assume that only one vehicle, veh_q , has entered h_{jk} during period $[-\gamma, 0)$. Also, we set $r = 250$ m, $V_1 = 11$ m/s (39.6 Km/h), $\varphi = 90^\circ$ and the vehicle veh_q the position of which is represented with $X(t)$, moves with speed 10 m/s (36 Km/h) and enters h_{jk} from intersection I_j at $T_q = -5$.

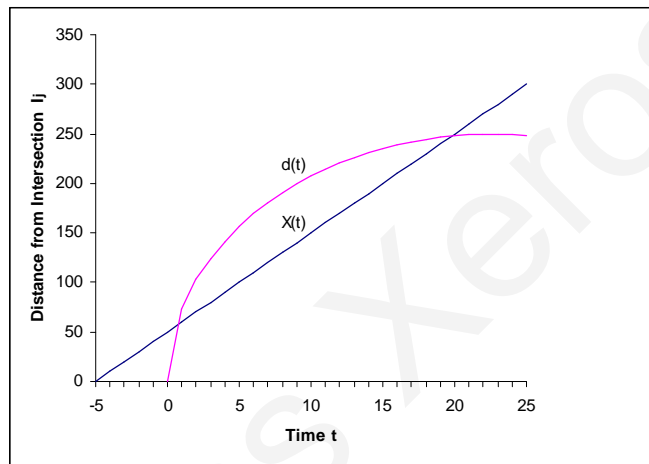


Figure 40: Example of catching up case. Plot of $d(t)$ and $X(t)$ vs time

We observe that $d(t)$ is smaller than $X(t)$ at the beginning of plot whereas, later on, at some point around $t = 0.5$, it catches-up with $X(t)$ and afterwards it becomes greater than $X(t)$ until around $t = 20$. After that, $X(t)$ is again greater than $d(t)$. The catching up period is $[0.5, 20]$. By keeping this in mind, we see that it is not enough to study the possibility that $X(t)$ is smaller than $d(t)$ just on time y (which is when $d(t)$ gets its greatest value) but it is necessary to see if, at any time during the period $[0, y]$, $X(t)$ becomes smaller than $d(t)$.

So, the probability of transmitting the information in the catching up scenario is given by the following equation:

$$p_c = \int_0^y F_{X(t)}(d(t)) dt \quad (46)$$

Case 2. veh_1 passed point R without carrying the information and it got informed before reaching intersection I_j

This case is more complicated than *case 1*. veh_1 did not have the information when passing from point R . The information was transmitted to it by a following vehicle on h_{ij} before reaching point I_j . We consider $t = 0$ the time when veh_1 gets the information. We call S the position of veh_1 on segment of road RI_j , on time $t = 0$ and s the distance that point S has from R . Figure 41 shows the notations of *case 2*.

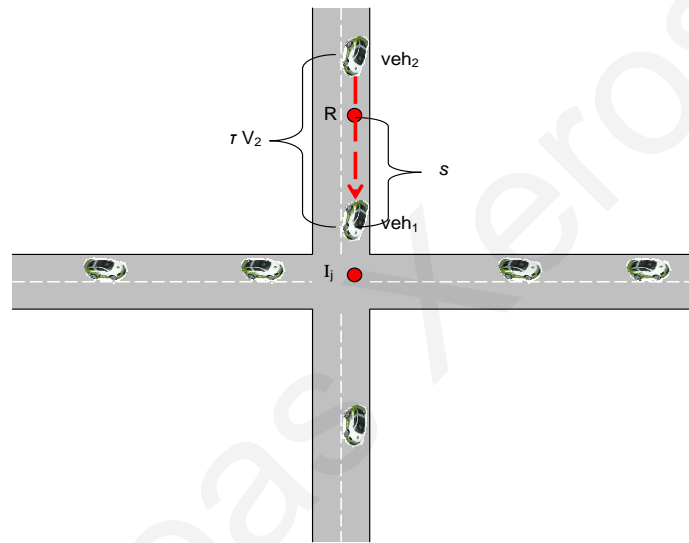
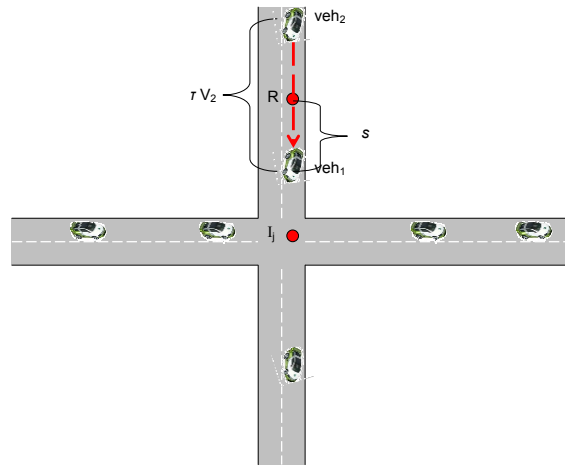
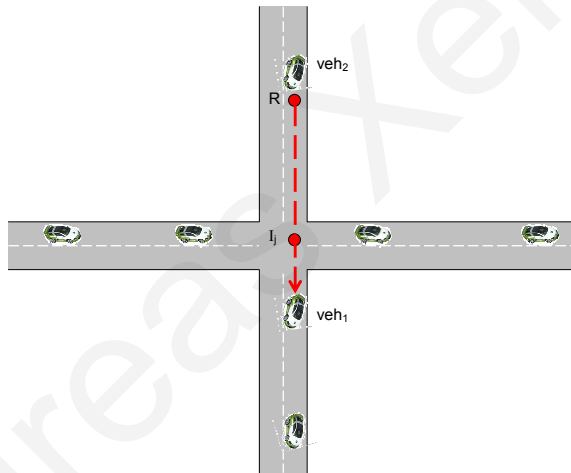


Figure 41: Case 2 scenario where veh_1 received the information at some point after passing from point R

In case when the time needed by veh_1 to reach I_j , which is $\frac{r-s}{V_1}$, is greater than or equal to y , then this scenario is very similar to case 1. If $y > \frac{r-s}{V_1}$, we need to consider the probability that the vehicle veh_2 which travels with speed V_2 , following veh_1 , will pass point R by time $t = \frac{r-s}{V_1}$ after which, veh_1 will not be on h_{ij} . If veh_2 passes from R before veh_1 passes from intersection I_j , then it can also transmit the information to vehicles entering road h_{jk} until time y expires. In order to not have an informed vehicle in RI_j during some time in period $[0, y]$ the following must hold:

- veh_1 passed point R without having the information and a vehicle veh_2 transmits the information to veh_1 at time $t = 0$. This means that the distance between veh_1 and veh_2 is smaller than r on time $t = 0$.
- on $t = \frac{r-s}{V_1}$, which is the time needed by veh_1 to pass from intersection I_j , veh_2 must not have passed from point R . In order for this to happen V_1 must have been greater than V_2 .

In other words, since V_1 is greater than V_2 (as indicated in the point 2 above), veh_1 must have either entered h_{ij} before veh_2 or (since veh_1 is the head of information) passed veh_2 at some time just before passing from point R . In the former case, since the distance between the two vehicles remained smaller than r from the beginning until the end of the road where veh_2 transmitted the information to veh_1 , we can conclude that it is highly unlikely that their distance became greater than r just before veh_1 exited road h_{ij} and veh_2 did not pass point R . In the latter case, since veh_1 passed from veh_2 and at that time neither of the vehicles was informed, veh_2 must have received the information from following vehicles just after the passing of veh_1 and before the distance of the two vehicles became greater than r . The only possible way that there were no informed vehicles between R and I_j during period of y is in the extreme case where veh_1 was very close to intersection I_j at the time when veh_2 received the message and immediately propagated it to veh_1 . Then, the time needed by veh_1 to exit road h_{ij} should not be enough for veh_2 to cover the distance from R given that the distance between the two vehicles on the time of transmission was smaller than or equal to r . The scenario is explained graphically in figure 42.

(a) veh_1 receives the Information

(b) veh_1 passed from intersection I_j and at the same time, veh_2 did not pass from point R . No informed vehicles in the area RI_j

Figure 42: The rare scenario where there is no informed vehicle in the area RI_j during the period of $[0, y]$. (a) veh_1 received the information message from veh_2 while being in the area RI_j . (b) veh_1 passed the intersection I_j before veh_2 had passed from point R resulting in no informed vehicles in the specific area.

It is obvious from the previous discussion that the circumstances under which there are no informed vehicles in the area RI_j during the period of $[0, y]$ are extremely rare to occur. Therefore, we reach the reasonable assumption that in the scenario of *case 2* during the period $[0, y]$ there is at least one informed vehicle in the area RI_j to propagate the information to any vehicle entering road h_{jk} . This means that the *probability of entering* is given by the same formula as in *case 1*.

$$p_e = 1 - P_0^{h_{jk}}(y) \quad (47)$$

Now we need to calculate the *probability of catching up* (p_c). For this probability we are going to work the same way as in the previous case, only now we have $Q(t, s)$ instead of $d(t)$, which is a random variable depending on the time and the initial value of s , given that s is uniformly distributed along RI_j .

$$Q(t, s) = (r - s - V_1 t) \cos \varphi + \sqrt{r^2 - (r - s - V_1 t)^2 \sin^2 \varphi} \quad (48)$$

In the case that φ forms a right angle then $Q(t, s) = \sqrt{r^2 - (r - s - V_1 t)^2}$. The function distribution that we are interested in is given by $F_{Q(t,s)-X(t)}(0)$. So,

$$F_{Q(t,s)-X(t)}(0) = \int_0^r \sum_{z=0}^{\infty} P[Q(t, s) - X(t) < 0 | Z(\gamma) = z] * P[Z(\gamma) = z] * f_s(s) ds \quad (49)$$

Finally, as in *case 1*, we want to see if during period $[0, y]$, $Q(t, s)$ gets bigger than $X(t)$. So, the *probability of catching up* is given by the following equation:

$$p_c = \int_0^y F_{Q(t,s)-X(t)}(0) dt \quad (50)$$

Now that we know the probability to transmit the information for both cases we need to combine them in order to calculate the overall probability $p_{tr}^{h_{ij}h_{jk}}$ to transmit the information from road h_{ij} to road h_{jk} . To do so, we need to find the probability for each case to happen separately.

In order for *case 1* to happen, veh_1 must pass point R carrying the information without any other vehicle in RI_j to transmit it. We call veh_2 , the vehicle that is in front of veh_1 on h_{ij} . The probability that veh_2 does not have the information is equal to the probability that veh_1 and veh_2 are of distance greater than r . If we call τ the time gap between these two vehicles, then their distance, $dist_{veh_1,veh_2}$ is equal to $\tau * V_1$.

In [98] it is given that the time gaps between vehicles are distributed according to the following pdf and PDF,

$$p_\tau(\tau) = \lambda e^{-\lambda\tau} \text{ and } P_\tau(\tau > T) = e^{-\lambda T}, \quad (51)$$

respectively. So, the probability that *case 1* happens is:

$$P_{cs_1} = P_\tau(\tau > \frac{r}{V_1}) = e^{-\lambda \frac{r}{V_1}} \quad (52)$$

Regarding *case 2*, it is sufficient to see that it is the compliment of *case 1* since either veh_1 is in the road segment RI_j when it gets the information or, it passes R and enters RI_j with the information. So,

$$P_{cs_2} = 1 - P_{cs_1} \quad (53)$$

So, the overall probability to have the information transmitted to road h_{jk} from road h_{ij} during a period of time y is

$$p_{tr}^{h_{ij}h_{jk}} = P_{cs_1} * (p_e^{cs_1} + (1 - p_e^{cs_1}) * p_c^{cs_1}) + P_{cs_2} * (p_e^{cs_2} + (1 - p_e^{cs_2}) * p_c^{cs_2}) \quad (54)$$

5.3.2 Driving Probability where veh_1 turns into road h_{jk} ($p_{dr}^{h_{ij}h_{jk}}$)

In this paragraph we are going to derive the probability $p_{dr}^{h_{ij}h_{jk}}$, which is the probability that vehicle veh_1 drives on road h_{jk} . In the previous subsection, where we calculated the $p_{tr}^{h_{ij}h_{jk}}$, the

time needed by veh_1 to reach intersection I_j is definitely less than or equal to the time period y , which is the period we study the probability of transmission to road h_{jk} . However, for simplicity reasons, we ignore the time gap between y and the time that veh_1 needs to reach intersection and make the choice of which road to drive on.

Probability $p_{dr}^{h_{jk}}$ is strongly related to the intersection we are studying, since one needs to consider all the other possible roads that can be chosen by a vehicle approaching intersection h_{jk} . For example, if we consider figure 43, there are two roads that veh_1 can follow, h_{jk} and h_{jq} . By assuming that the road a vehicle follows when arriving to an intersection is independent from the road it uses to get to the intersection, the probability $p_{tr}^{h_{ij}h_{jk}}$ is equal to $\frac{\lambda_{jk}}{\lambda_{jk} + \lambda_{jq}}$.

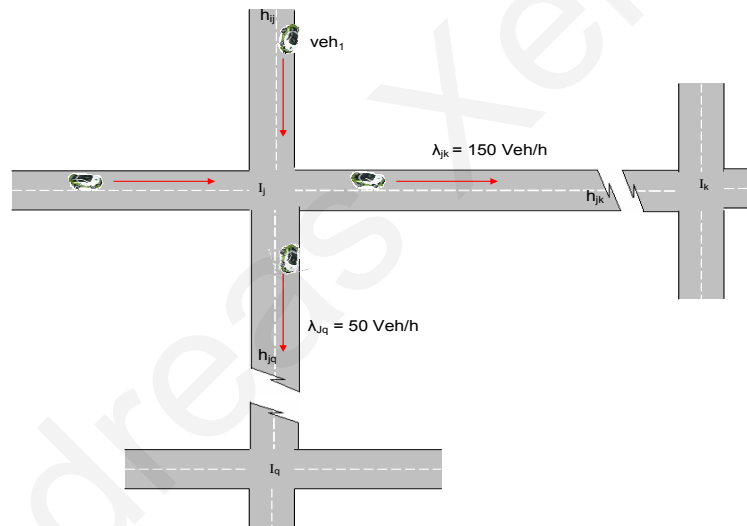


Figure 43: Example showing how to calculate the probability of information propagation with the driving way

In general cases, there are more complicated intersections, where the driver has more roads to choose from. The general equation covering all cases is:

$$p_{dr}^{h_{ij}h_{jk}} = \frac{\lambda(h_{jk})}{\sum \lambda(h_{jb})}, \text{ for all } b \text{ such that road } h_{jb} \text{ exists.} \quad (55)$$

Now, we have everything needed to calculate all the terms of the equation (40), which gives the lower bound of the probability of information propagation $p_{h_{ij}h_{jk}}$, from informed vehicle on road h_{ij} to uninformed vehicles on road h_{jk} .

5.4 Simulation Validation

In this paragraph we validate our theoretical findings with simulation results. We conduct our simulations on VISSIM. In all of our simulations we used VISSIM to model the setup shown in figure 44. For simplicity we set angle φ equal to 90° . Since the probability of information propagation with the *driving way* depends only on the average arrival rates of the roads that the vehicle reaching the intersection may follow, there is no need to simulate this case because we assume that the arrival rates are given. Hence, we concentrate on the validation of the information transmission probability $p_{tr}^{h_{ij}h_{jk}}$. Information transmission probability, as it is shown in equation (54), is only related to the traffic characteristics of the road intended to receive the information.

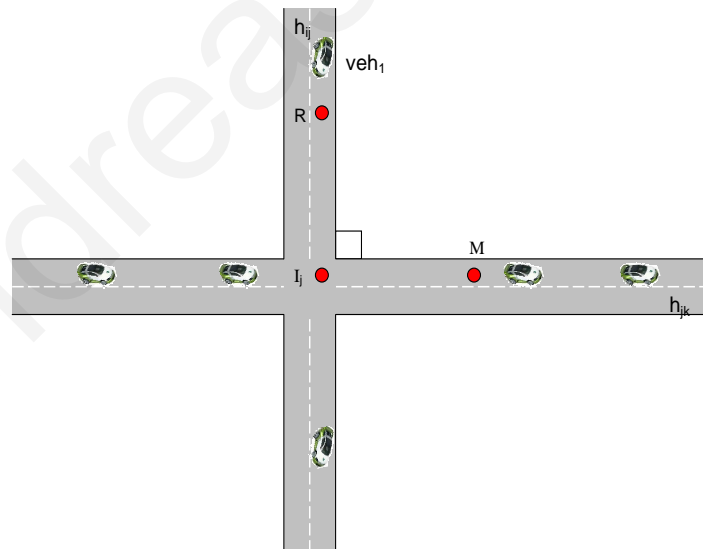


Figure 44: Intersection setup used in VISSIM

The setup parameters are the vehicle arrival rates on road h_{jk} , the range of speeds attained by the vehicles on road h_{jk} and the speed of veh_1 moving on road h_{ij} and driving towards intersection I_j . Each simulation generates ascii files that include the position coordinates of the vehicles at each simulation step. The simulation step is set to 100 ms. We have also developed an application on C++ to process the simulator's output in order to generate information with which we can infer whether any vehicle on road h_{jk} eventually receives the desired information from veh_1 , as well as at what time. The transmission range of each vehicle is constant and set to 250 m for all vehicles. We consider that vehicles that are of smaller distance than the transmission range can exchange messages. Since the transmission speed is much bigger (250 m in 6 ms) in comparison to the vehicle speed, we consider the transmission time as zero. Therefore, we do not use a specific wireless communication model for the measurements. For each set of parameters we repeat the simulation one hundred times and we calculate the frequency with which the information is successfully transmitted. The probability of information transmission is estimated by dividing the number of successes by the number of times we have repeated the simulation. The estimated probability is then compared to the theoretical probability obtained using the derived equation (54).

At first we give insights of how the information transmission probability varies with respect to different parameters. We calculate the probability using the equation (54) we derived in the previous section. In all calculations we keep the speed range of road h_{jk} constant in the interval (60-80 Km/h). In figure 45 we plot the probability versus the time window for different arrival rates. The time window is the time which elapses from the instant veh_1 passes from point R . As expected, the probability increases in a concave fashion with increasing time window. In addition, as the arrival rate increases so does the probability.

In figure 45 we can see that, as time goes by the probability increases for all the arrival rates. This is because as time increases while also increases the chances to have a vehicle entering h_{jk}

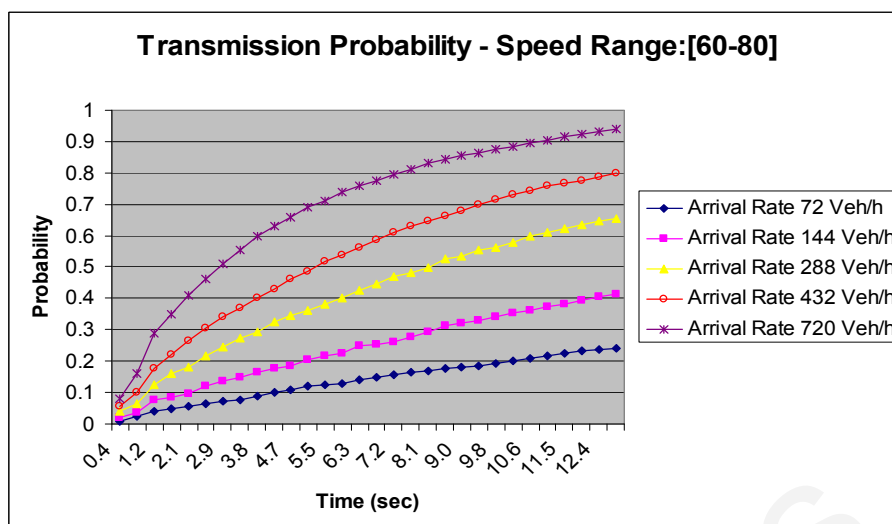
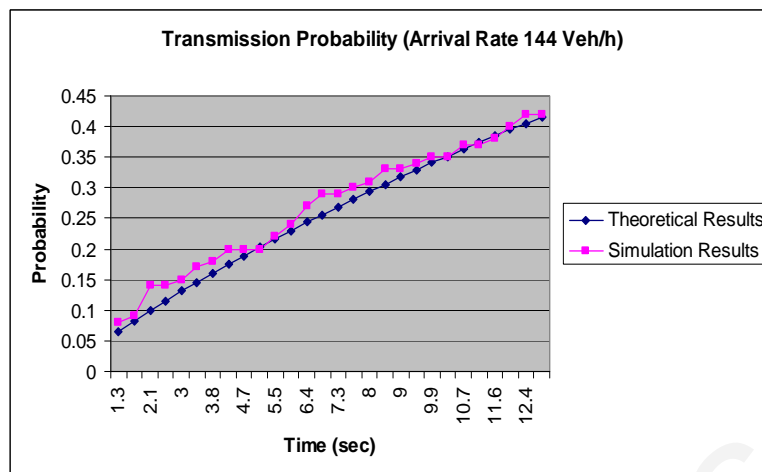


Figure 45: Theoretical results of Transmission Probability for different arrival rates

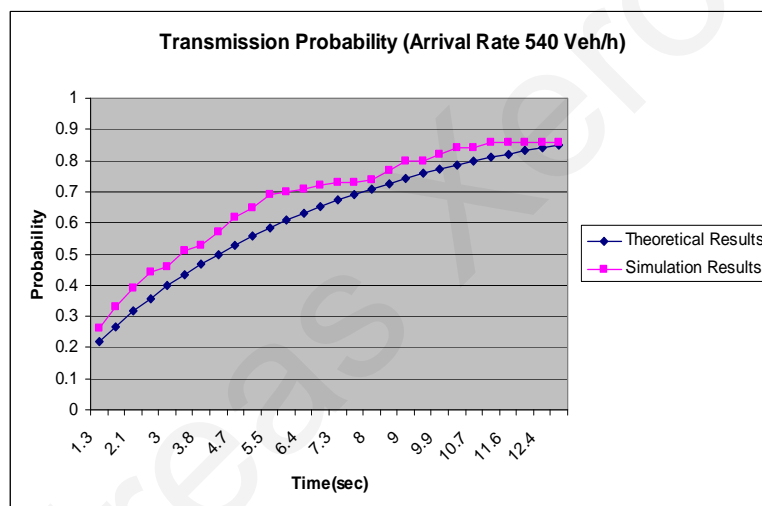
(probability of entering) as well as the chances to have the transmission range of veh_1 catching up with a vehicle that has entered h_{jk} before $t = 0$ (probability of catching up). Also, we see that the probability increases when the arrival rate increases, since there are more vehicles entering road h_{jk} .

In figures 46 (a and b) we present our theoretical results in comparison with the simulations' results. They show how the probability increases as time increases for specific values of the arrival rates. We choose arrival rates that are low enough (144 Veh/h - 1 vehicle in 25 seconds and 540 Veh/h - 1 vehicle in 7 seconds) to keep the probability from getting its highest value (close to 1) very soon. In this way we can better observe the relation of the different results during a longer period of time.

In figures 46 (a and b) we observe that the simulation probability, for most of the cases, is slightly higher than the theoretical one, whereas there are some cases that the results are equal. Also, we see that the trends that the plots have are similar for theoretical and simulation probabilities for both values of arrival rates. This is a good support of the estimation of the actual probability that our analysis provides.



(a)

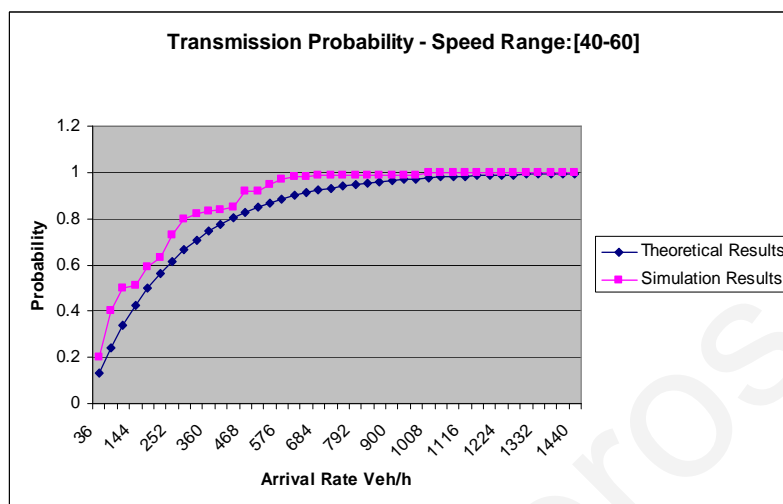


(b)

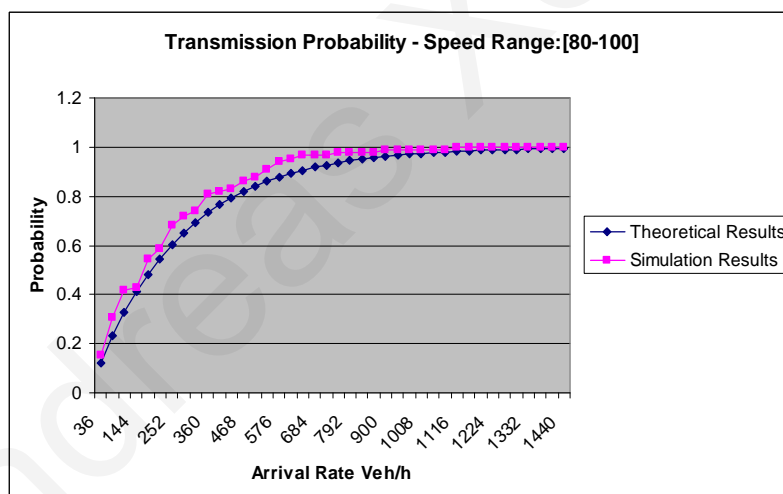
Figure 46: Comparing Theoretical with Simulation results of Transmission Probability for arrival rates 144 Veh/h and 540 Veh/h

Finally, figures 47 (a and b) show the relation of theoretical and simulation results regarding the increase of probability as the arrival rate increases. We start with very low traffic density of 36 veh/h (1 vehicle every 100 seconds) and we increase the rate up to a very heavy traffic density of

1440 veh/h (1 vehicle every 2.5 seconds). The time is fixed and is equal to the time needed by a vehicle with speed 65 Km/h to cover distance r .



(a)



(b)

Figure 47: Comparing Theoretical with Simulation results of Transmission Probability after 14 seconds

As in the previous figures (47 (a and b)), our theoretical results follow the same trend as the simulation ones, thus supporting the validation of our analysis. Moreover, figures 47 show that the transmission probability for high arrival rates (1100 Veh/h - 1 vehicle every 3 seconds) is very

close to 1. This is because it is highly unlikely for such arrival rates not to have any vehicle entering road h_{ij} after 14 seconds. Another useful observation is that for arrival rates greater than 500 Veh/h (1 vehicle every 7 seconds) the information transmission probability is greater than 0.8, after passing time of 14 seconds.

5.5 Conclusions / Sum up

In this chapter we study the problem of information dissemination in VANETs and we provide a measure of the probability to propagate information on intersecting roads where no static infrastructure is used. We present a lower bound on the probability to propagate information between vehicles of two roads close to their point of intersection. We show that this probability is strongly related to the arrival rate of the vehicles entering the road where the information is to be transmitted. We also show that, as the time allowed for propagation increases, so does the probability. We validate our results with simulation evaluation using VISSIM. The derived probability function can prove to be a valuable tool in the solution of a number of information dissemination problems. It can, for example, be used to derive appropriate selection criteria for the installation of static infrastructure at roadway cross-sections. It may also be used to calculate links costs for routing protocols whose objective is to maximize the probability of successful information delivery. The latter approach is further explored in the subsequent chapter.

Chapter 6

Time Constrained Message Delivery Probability in VANETs

6.1 Introduction

In this chapter we investigate the problem of information delivery based on the probability to deliver messages to a specific area in a certain amount of time and in a setting where no static infrastructure is used [127]. We adopt a graph theoretical approach where we model the road map and the traffic characteristics using a directed weighted graph, which we refer to as the Road-Graph. We use the Road-Graph to calculate the probability of disseminating information along any given path, in a specific amount of time. So, for any two points on the Road-Graph we can find the path with the maximum probability of successful information delivery in the chosen time interval. This information is of great significance, since it can be utilized by routing protocols to optimally route packets in the vehicular network. We validate our analytical findings by comparing them with data obtained using VISSIM.

6.2 Problem Formulation - Assumptions

We consider that we are given a road map of an area that consists of the main roads and highways of the area and the corresponding intersections. We don't take into consideration small roads

with insignificant traffic or roads in city centers since buildings constrain the signal propagation. Thus, it is more difficult to establish wireless connectivity in urban areas and the network efficiency is decreased. On the other hand, since vehicle speed is low and the car density very high, the information delivery probability can approach the value 1. Furthermore, the complexity of the vehicle movements in cities which include traffic lights, stop signs and different speed limits makes the theoretical analysis of information propagation more complex which is out of the scope of this research work. We also assume that we are given the traffic conditions of each main road and highway with the arrival rate of vehicles and their speed range on the specific road. We assume that a vehicle travels with a constant speed, selected uniformly and independently from the speed range of the road. We also assume that the number of vehicles entering a road is a random variable that follows the Poisson process, since several experiments have shown that the outcomes of such a model are in good agreement with real measurements obtained in practice [93]. In the case of highways, where there are many exits and entrances along them, we consider the number of vehicles entering each highway independently, since vehicles may enter or leave during their driving. For simplicity reasons, we assume that traffic conditions remain constant along the highway, at least on the parameters that affect our analysis, given that vehicles can enter or exit from smaller roads on the side. This assumption influences the traffic integrity in the intersections of the highways. However, the analysis of the traffic flow and the calculation of the arrival rate of vehicles in each highway is not considered in this paper. We consider that each highway follows an independent Poisson process.

We also assume one-way and one-lane traffic along the roads. In the case where the highway is multi-lane, we only consider the lane with the highest density. For two-way highways we consider that they consist of one-lane in each direction. In the Road-Graph that we construct for

modeling the road map we add a different directed edge for each direction. We study the information propagation to the direction of the traffic using vehicles moving on this direction. Of course, considering vehicles moving on the other direction will definitely improve both information speed and information delivery probability.

Also, we consider that all vehicles have a constant transmission range denoted by r , which is the same for all vehicles.

Finally, since traffic conditions vary depending on the period of the day that are measured (i.e. morning, noon, afternoon or night), the same road map has different traffic characteristics during different hours of a given day. Therefore, the same road map is modeled by different Road-Graphs for different periods of the day.

6.3 Road-Graph Construction

In this section we outline the steps followed to create the Road-Graph. The resulting Road-Graph contains all the necessary data to study the information delivery probability in VANETs and decide on the path that gives the highest probability.

We consider that we are given the road map of the area that we study, which includes the roads and their intersections, the arrival rate and the vehicle speed range of each road during a specific time period in a day. We begin by presenting the basic notations used in our analysis.

I_i : the i^{th} intersection in the road map

n_j : the node of the Road-Graph that corresponds to the i^{th} intersection of the road map

h_{ij} : the road connecting intersection I_i with intersection I_j . The traffic direction is from I_i to I_j

L_{ij} : the length of the road h_{ij}

$e(n_i, n_j)$: the directed weighted edge of Road-Graph that joins n_i to n_j

$w(n_i, n_j)$: the weight of the edge $e(n_i, n_j)$ in the Road-Graph that represents the average time needed for a message to travel the distance $L_{ij} - r$

v_{ij} : the average message propagation speed on the road h_{ij} in the direction of vehicles drive

$prob_{j ik}(t)$: the probability that a message will propagate from road h_{ji} to road h_{ik} in a time period t

We now provide an algorithmic representation of the steps that must be followed in order to generate the Road-Graph.

Algorithm: Construction of the Road-Graph

Input: a road map (roads and their intersections), the arrival rate and the vehicle's speed range for each road

Output: the corresponding Road-Graph (V, E)

BEGIN

1. **FOR ALL** intersections I_j in the road map **do**

- Create a node n_j and add it in the Road-Graph (i.e., $V = V \cup \{n_j\}$)

2. **FOR ALL** roads h_{ij} in the road map **do**

- add the directed weighted edge $e(n_i, n_j)$ in the Road-Graph

3. **FOR ALL** roads h_{ij} in the road map **do**

- calculate the average message propagation speed v_{ij} in the direction of vehicles drive (using equation (2) of [125] which is explained briefly in section 6.3.1)

- $w(n_i, n_j) = \frac{(L_{ij} - r)}{v_{ij}}$

4. **FOR ALL** nodes n_i in the Road-Graph **do**

FOR ALL pairs of edges $e(n_j, n_i)$ and $e(n_i, n_k)$ **do**

- $prob_{jik}(t) = f_{jik}(t)$ (using equation (40) of chapter 5)

END

6.3.1 Analysis of the Algorithm: “Construction of the Road-Graph”

The first two loops are straightforward and do not need any further analysis. In the third loop we must calculate the average message propagation speed on a single road. In order to do this we use the analytical model proposed in [125].

In the research work of [125] authors study information propagation in one direction along a straight line road. In particular, they have developed a formula that can be used to calculate the information propagation speed along the straight line road. Figure 48 shows an example of such information propagation. The arrow represents the road and the circles below (above) the arrow are positive (negative) vehicle traffic flow. Only one-dimensional distance along the road is considered; the road’s width is neglected. Position coordinates of vehicles increase to the right. A message is propagated in the positive direction from a vehicle at location H at time 0. The “message head” at time t refers to the informed vehicle with the largest position coordinate. The primarily concern is the movement of the message head. Note that the message reaches the front-most vehicle in its current partition through multi-hop forwarding, and begins to travel with the message head. At time t_2 , the message head reaches the partition that lies ahead, and the prior process repeats.

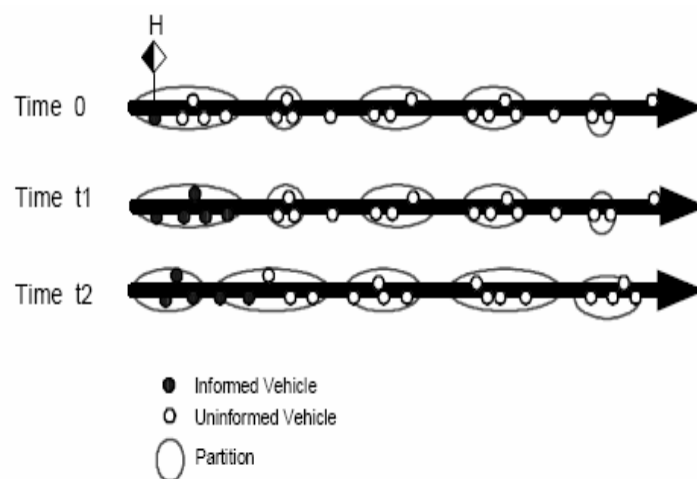


Figure 48: An example of message propagation

The data propagation models are based on the following observations:

- The V2V network is a partitioned network. A snapshot of the positions of vehicles at any time instant will typically yield a network with many partitions.
- The configuration of partitions is dynamic. Partitions may split or merge as the relative positions of vehicles within the partition or between partitions change.
- A message propagates in either one of two processes termed the forward process and the catch-up process. The forward process involves the propagation of the message within a partition via multi-hop forwarding. In the catch-up process the message moves along with its carrying vehicle until it comes within the radio range of the last uninformed vehicle in the partition ahead of it. The propagation speed in the catch-up process will normally be much slower than that in the forward process. When information propagates along the road, it alternates between the forward and catch-up phase, resulting in a cyclic process.

The analysis is based on the scenario where a message propagates in the direction in which the vehicle traffic is traveling in one-way vehicle traffic.

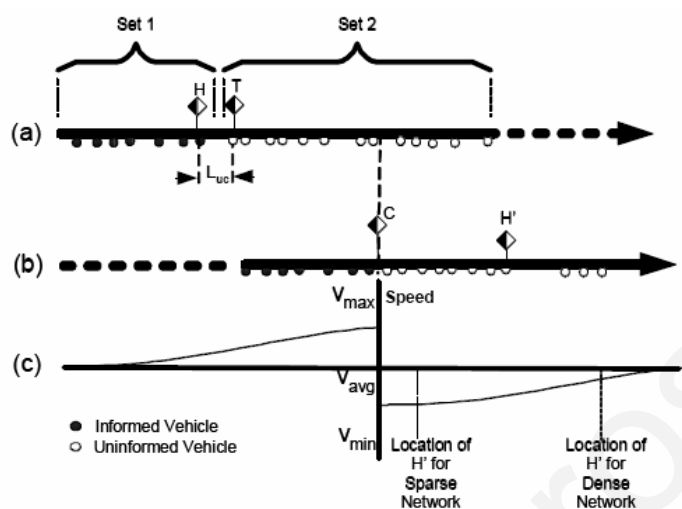


Figure 49: Catch-up Process

Figure 49 illustrates the catch-up process. Figure 49(a) shows the beginning of the catch-up process. H is the location of the message head. Vehicles are in either one of two sets: informed vehicles are in set 1 and uninformed vehicles are in set 2. Figure 49(b) shows the end of the catch-up process, when the message head just enters the radio range of the last uninformed vehicle in set 2. Location C is the catch-up point. H' will be the location of the message head in the beginning of the next catchup process if we ignore the time spent in intra-partition forwarding. Figure 49(c) shows the temporary non-uniform spatial distribution of average vehicle speed around C in the end of the catch-up process. During the catch-up process, the fast informed vehicles in set 1 are likely to pass the slow informed vehicles and concentrate near C , and the slow uninformed vehicles in set 2 tend to lag behind and stay close to C .

It worth notice that the catch-up process can be improved by using vehicles moving on the opposite direction to the direction of information traveling. The information message can reach

the last uninformed vehicle in set 2 using vehicles moving on the opposite road. In addition, the two-way vehicle traffic can extend the partitioned areas and thus shorten the inter-partition distance by increasing the vehicle density. The effect of vehicles moving on the opposite direction is more evident when the traffic density of vehicles moving in the same direction as the information is very small and the traffic density on the opposite direction is very high. In this case, the information message will utilize multi-hop propagation in the opposite direction to move forward and reach the vehicle tail in the direction of information propagation. In this thesis we will not account for the two-way traffic due to the overly complex mathematics required.

For a given time t , $X(t)$ is defined as the traveling distance of the message head during $[0, t]$, $V_p(t)$ the average message propagation speed during the interval $[0, t]$. We have $V_p(t) = X(t)/t$. The definition of the longterm average message propagation speed v_p is

$$v_p = \lim_{t \rightarrow \infty} V_p(t) = \lim_{t \rightarrow \infty} \frac{X(t)}{t} \quad (56)$$

The message propagation process can be modeled as a renewal reward process with a new cycle beginning each time the catch-up process begins and the reward is the message propagation distance during each cycle. Each cycle consists of a catch-up phase followed by a forward phase. It can be shown that

$$v_p = \frac{E[X_c + X_f]}{E[T_c + T_f]} = \frac{E[X_c] + E[X_f]}{E[T_c] + E[T_f]} \quad (57)$$

where X_c (X_f) is the distance traveled during the catch-up (forward) phase and T_c (T_f) is the time spent in the catch-up (forward) phase. The details of the computation of the $E[T_c]$ and $E[X_c]$ of the catch up phase and the computation of the $E[T_f]$ and $E[X_f]$ are explained in [125].

Using simulation authors have validated the theoretical analysis showing how information propagation speed increases with the increase of traffic density for a specific road. Figure 50 shows their theoretical and simulation results. The vehicle speed range is from 40 Km/h to 70 Km/h and the transmission range is 50 m.

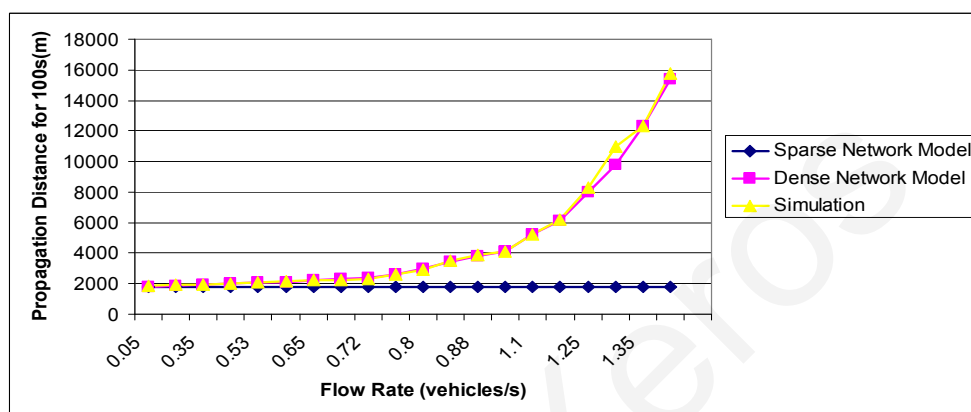


Figure 50: Theoretical and Simulation results of Information Propagation speed on a single lane road

The reason we divide the length L_{ij} of road h_{ij} minus the transmission range r , with the average message propagation speed is because in distance r from the intersection I_j information can propagate to a road intersecting h_{ij} on intersection I_j . The probability that a message will propagate to the next road in a given time period is calculated in the last loop. This calculation requires the lower bound on the probability of information propagation between two intersecting roads in a given time period. For this calculation we use the transmission probability given by equation (54), which is provided in our research work in chapter 5.

A simple example

For clarity of presentation we provide a simple example, which helps the reader get a better understanding of the method used to derive the Road-Graph and how it looks like given a specific road map. We assume that we are given the road map as depicted in figure 51. The road map comprises of 3 roads and 4 intersections. Road h_{12} starts from intersection I_1 and ends on intersection I_2 . Vehicles arrive on the road from intersection I_1 with arrival rate 1200 Veh/h and drive with a constant speed chosen randomly from the range 70 Km/h to 90 Km/h. The length of the road is 18 Km. The other two roads are road h_{23} and road h_{24} . The arrival rate, vehicle speed range and length of roads are shown in figure 51. The direction of car movement is defined by the direction of the arrows. Since we consider highways with adjacent roads at their side that can be used by vehicles in order to enter or leave the highway, the sum of the arrival rates of the roads h_{23} and h_{24} does not need to be equal to the arrival rate of h_{12} .

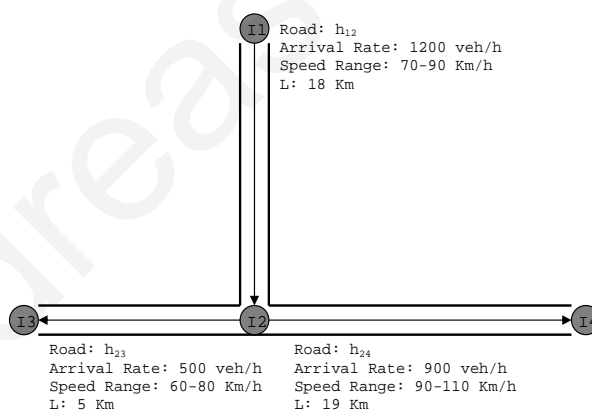


Figure 51: Representation of a simple road map with 3 roads and 4 intersections

Following the aforementioned “Construction of the Road-Graph” algorithm, we create the Road-Graph from the road map provided in figure 51. We first create the four nodes, n_1, n_2, n_3, n_4 representing the four intersections I_1, I_2, I_3, I_4 respectively. Then, for each road in the map

we create an edge in the Road-Graph with the direction as defined by the direction of traffic flow. Thus, we have $e(n_1, n_2)$, $e(n_2, n_3)$ and $e(n_2, n_4)$. According to the algorithm, we proceed to the calculation of the average time needed for the information to propagate along the road based on [125] and we assign this value to the weight of the respective road. Finally, for node n_2 , which is the only one that has input and output edges, we provide the two functions of time for the computation of a lower bound of the information dissemination probability, from road $h_{1,2}$ to road $h_{2,3}$ and to road $h_{2,4}$. These functions are denoted by $f_{n_1n_2n_3}(t)$ and $f_{n_1n_2n_4}(t)$, respectively, which can be calculated using the equation (40). The analysis and the derivation of equation (40) is presented in section 5. The resulting Road-Graph is shown in figure 52.

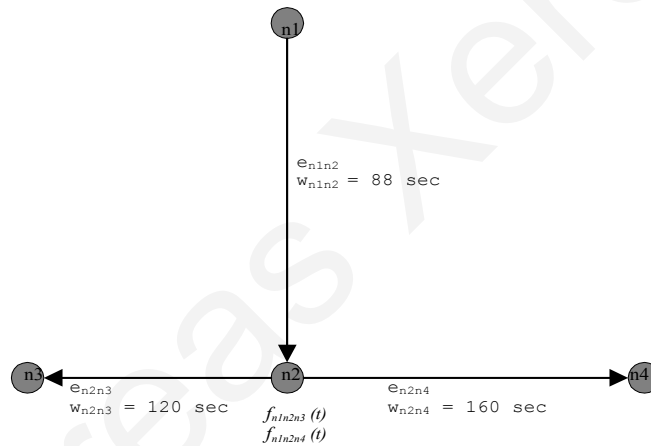


Figure 52: The Road-Graph that corresponds to the road map of figure 51

6.4 Message Delivery Probability

In this paragraph we explain how to calculate the probability that a message, starting from the source area, will be delivered to the target area in a given time period T , using the Road-Graph. For simplicity and without loss of generality, we use intersections as source and target areas. We denote by n_s the node of the Road-Graph that corresponds to the source intersection and n_d as the node corresponding to the destination intersection.

Firstly, we need to find the most probable paths from source to destination that the information will follow in order to be delivered. To do so, we find the 10 shortest paths with respect to the weight of the edges, ignoring the time needed for information propagation on intersections. The algorithm we use is described in [38] which enumerates the k shortest simple (loopless) paths in a directed Graph with n nodes and m edges, with time complexity $O(kn(m + n \log n))$ in worst case. Alternatively we could use the algorithm presented in [76] where the authors develop a new algorithm to find the k -best (disjoint) paths with complexity $O(n^2)$. The results of such an analysis could be useful when studying disaster scenarios, which involve total loss of a particular path. In such cases a different path with no common edges should be used. Such analysis will be considered in future work.

For the enumeration we only consider the time needed by the message to travel the distance of each road, ignoring the time spent on the intersections. The reason that we choose the 10 shortest paths is that experimental results, which we have run, have shown that investigating the probability of delivery using 10 shortest paths is more than enough to find the path with the highest probability. The paths are sorted starting from the shortest one.

We call $p_q = (n_{q1}, \dots, n_{qm})$ the q^{th} shortest path that joins the source node of the message to the destination node, i.e., n_{q1} is the starting (source) node n_s and n_{qm} is the destination node n_d . The number of nodes in the path is m and the number of edges is $m - 1$. We define as $c(p_q)$ the cost of path p_q , which is equal to the sum of the weights of the edges of the path. The calculation of the cost of the path does not account for the time spent by the information message to travel between roads on their intersection. In other words, $c(p_q)$ is the average time needed by the information message to travel along the specific path assuming that there are no delays (zero time) in information propagation on intersections. So,

$$c(p_q) = \sum_{j=1}^{m-1} w(n_{qj}, n_{qj+1}) \quad (58)$$

In order to be able to calculate the probability that the message will be delivered using a specific path in the given time period T , the cost of path $c(p_q)$ must be less than or equal to T . This is because, by definition, $c(p_q)$ is the average time that a message will spend in order to be delivered, ignoring the time spent on intersections. Otherwise, the average message delivery time will be more than the given time T , even when ignoring intersections that will give a delivery probability equals to zero.

We define as t_{p_q} the extra time that we can distribute to the $m_q - 2$ internal intersections of the path p_q and calculate the probability of each one. The reason we are considering $m_q - 2$ intersections is because we do not consider the nodes n_1 and n_{m_q} , since the former, as source of information, starts to provide the information as soon as there is a vehicle there, and the latter one is the destination node, where information will not need to travel any more. The equation giving t_{p_q} is:

$$t_{p_q} = T - c(p_q) \quad (59)$$

In other words, t_{p_q} is the average spare time available to the message in order to reach the destination node, considering that the average time spent on the roads of the path is $c(p_q)$.

The next step is to find the probability that the information will pass from all internal intersections in the given time t_{p_q} . To do so, we must first find the probability of each intersection independently, given the portion of time that we can spend on each one, which is $\frac{t_{p_q}}{m_q - 2}$. Then, the overall probability of path p_q is given by the product of all the probabilities of the internal intersections of the path. So, the message delivery probability using path p_q denote by $P_{Del}(p_q)$,

is given by equation:

$$P_{Del}(p_q) = \prod_{j=2}^{m_q-1} f_{j-1,j,j+1}(t_{p_q}/(m_q - 2)) \quad (60)$$

Finally, once we find the delivery probabilities of the 10 shortest paths, we select the path with the highest delivery probability. So, equation (59) gives the message delivery probability P_{max} from source to destination in the given amount of time T .

$$P_{max} = \max_{1 \leq q \leq k} (P_{Del}(p_q)) \quad (61)$$

6.5 Real World Example

In this section we apply the proposed methodologies and the obtained theoretical results to a real world example. We consider a section of the highway network in the Los Angeles area, which includes 210 highways. The road map, as obtained from Google Earth [29], is depicted in figure 53.

To characterize the traffic conditions on each highway we use Google Earth to obtain approximations of the traffic density and SIGALERT [105] to obtain average vehicle speeds. We assume that the vehicles can attain speeds between 10 kilometers below the average value and 10 kilometers above the average value.

We follow the algorithm described previously to obtain the desired Road-Graph. For each intersection in the road map we initially create a node on the Road-Graph. Then, for each highway we create a directed, weighted edge following the vehicle driving direction as shown in figure 54. By following loop 3 of the algorithm, for all the edges on the Road-Graph we find the time needed for the information to travel along the corresponding highway and we assign this value to the weight of the edge. Finally, as described in loop 4, for each node n_i and for all pairs of edges

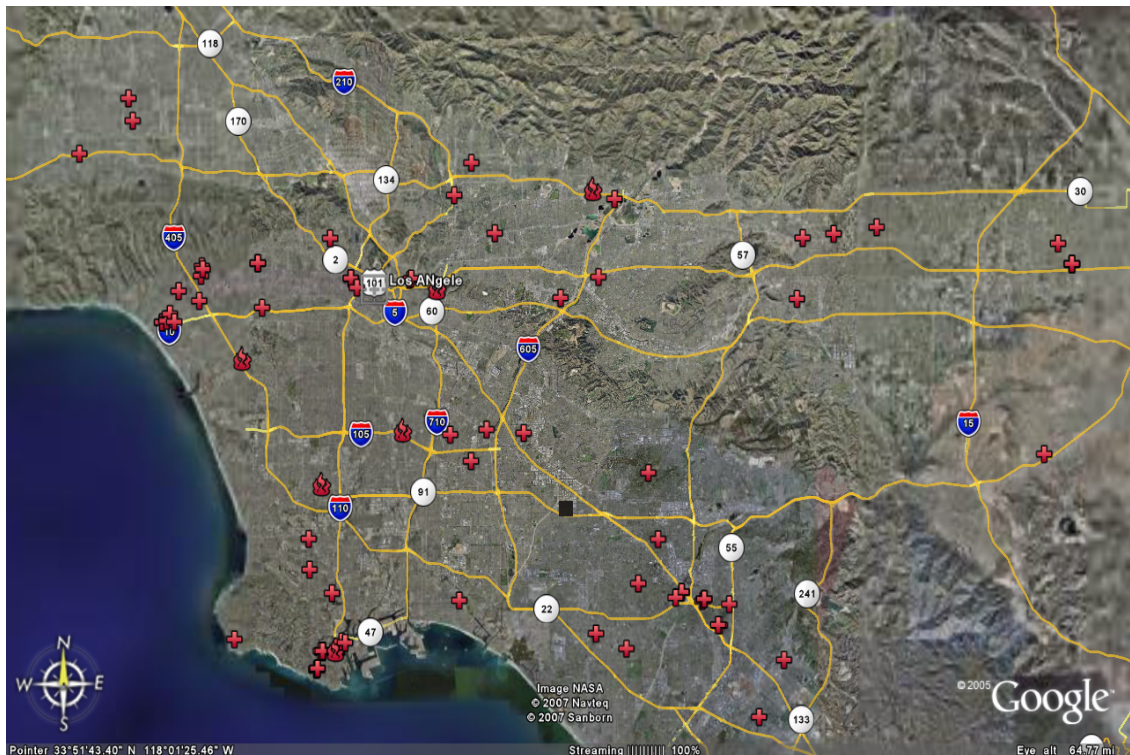


Figure 53: Major highways around the city of Los Angeles

$e(n_j, n_i)$ and $e(n_i, n_k)$ we derive the formula $f_{jik}(t)$, which, given a time period t , calculates the probability for the information to be transmitted from highway h_{ji} to highway h_{ik} .

Having obtained the Road-Graph (see figure 54), we can now calculate the information delivery probability versus time for any combination of source and destination intersections, using the method described in Section 6.4. We consider five scenarios, which involve different combinations of source and destination nodes that the message has to travel, as shown in figure 54. Three of the cases ($I_{13}(S_1) - I_{53}(D_1)$, $I_1(S_2) - I_{51}(D_2)$, $I_{54}(S_3) - I_{12}(D_3)$), which are studied only theoretically, have many intermediate nodes in the paths from source to destination. In addition, the distances of these combinations are comparable to the length of the entire Los Angeles area, which implies that the information traveling time is significant. The other two cases ($I_{15}(S_4) - I_{41}(D_4)$, $I_3(S_5) - I_{22}(D_5)$) are studied both theoretically and via simulations, and the results are compared. Due to the high number of traveling vehicles and the long simulation time, the output

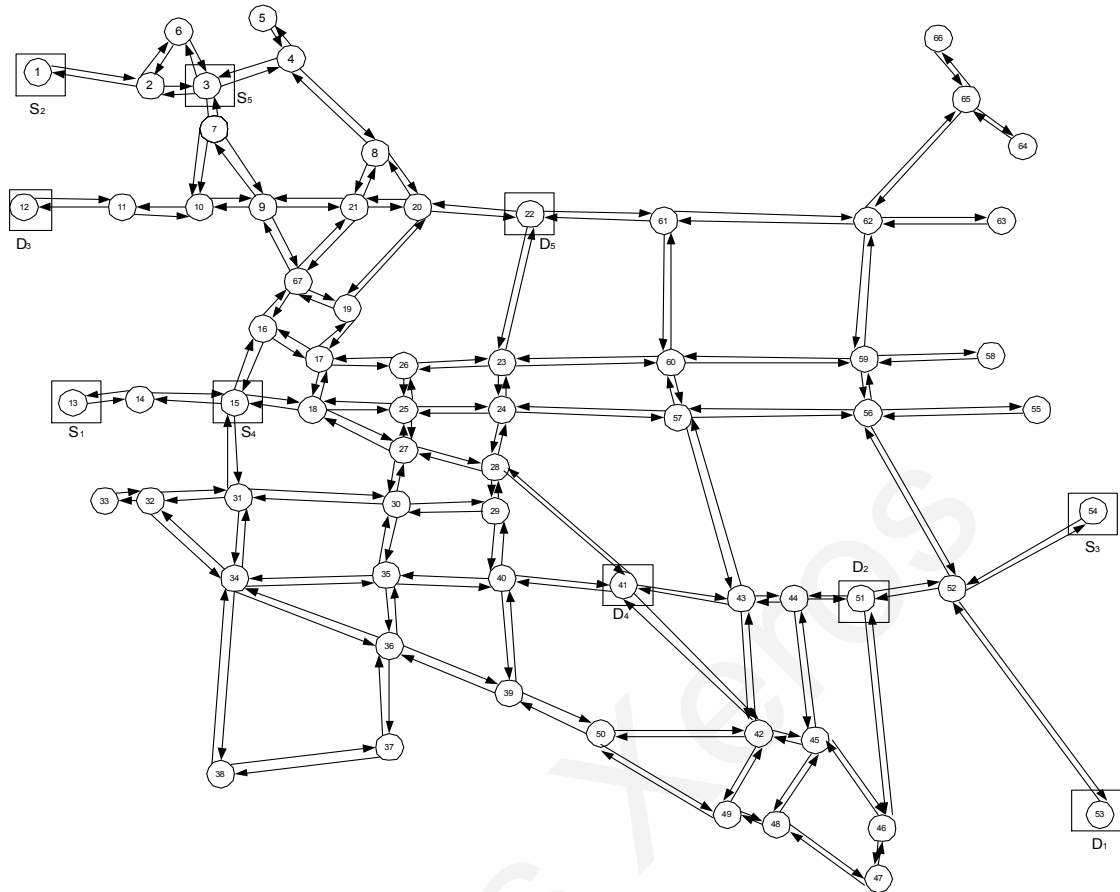


Figure 54: The Road-Graph of major highways in the Los Angeles Area. S_i and D_i indicate the scenarios that we study (Source-Destination)

file of the simulator (VISSIM) is huge. In order to make simulation runs manageable we choose combinations of source-destination nodes that are close enough (but not very close) and we filter out the traffic of roads that do not contribute towards information propagation.

6.5.1 Analytical Results

In figure 55 we present the probability of message delivery for the first Source-Destination pair, from intersection I_{13} to the destination intersection I_{53} .

In this case, the highest delivery probability is provided by the shortest path for all the values of time period T . The shortest path is $p_1 : n_{13} - n_{14} - n_{15} - n_{16} - n_{17} - n_{18} - n_{27} - n_{28} -$

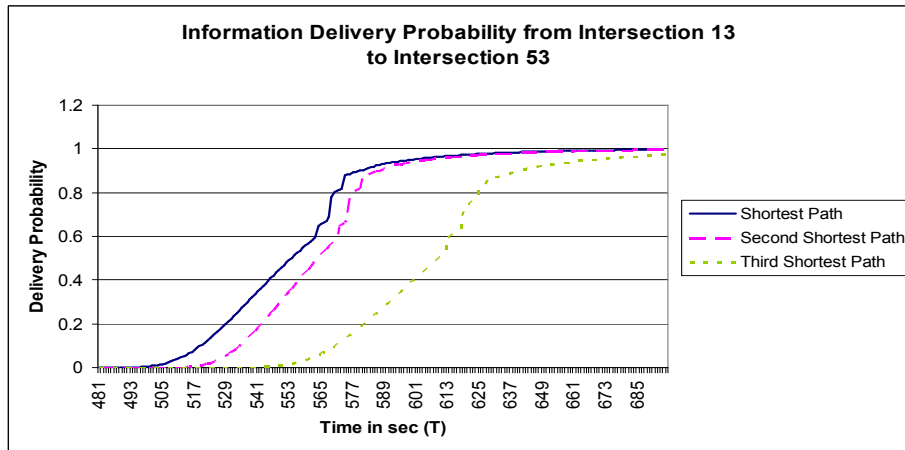


Figure 55: Information Delivery Probability versus Time from intersection I_{13} to intersection I_{53}

$n_{41} - n_{43} - n_{44} - n_{51} - n_{52} - n_{53}$, with cost $c(p_1)$: 473 sec. The total distance of the path is 180 Km, which implies that the best case average information speed along this path is 1370 Km/h. In figure 55 we also show the message delivery probability along the second and third shortest paths. The figure indicates that for all values of T the highest probability is offered by the shortest path.

The second Source - Destination pair that we consider is from intersection I_1 to intersection I_{51} . The information delivery probability versus time is presented in figure 56.

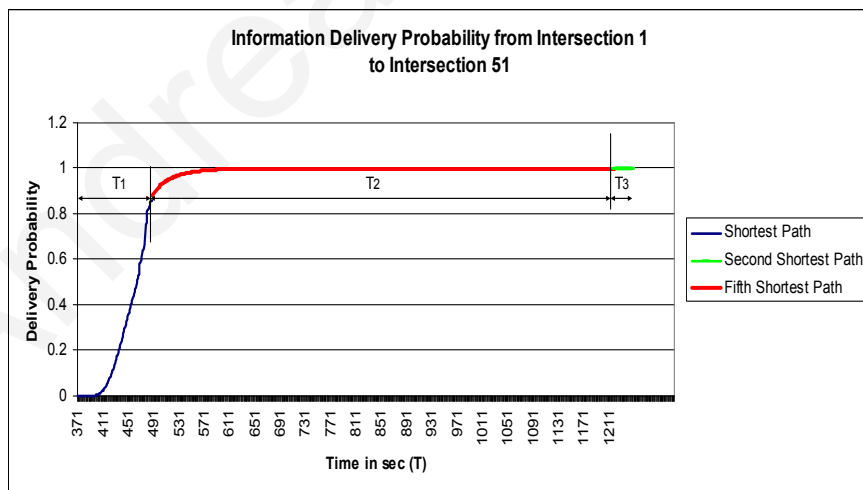


Figure 56: Information Delivery Probability versus Time from intersection I_1 to intersection I_{51}

In this case, the highest probability is not always given by the shortest path. We observe that the highest probability is given by the shortest path only during simulation time starting from 371 seconds to 490 seconds (period T1). Between 490 seconds and 1200 seconds (period T2) the highest probability is offered by the fifth shortest path and after 1200 (period T3) seconds the highest delivery probability is given by the second shortest path. This is possible because the formula that gives the information propagation probability on intersections does not increase in the same way for all intersections. The probability strongly depends on traffic conditions of the intersected highways, especially on the arrival rate of the highway that the information needs to be transmitted. So, different paths with different intersections will have different increasing modes of information delivery probability as time period T increases. The first five shortest paths for this combination of source and destination intersections with their cost, total distance and best case average message propagation speed are listed below.

1. $p_1 : n_1 - n_2 - n_3 - n_7 - n_9 - n_{67} - n_{16} - n_{17} - n_{18} - n_{27} - n_{28} - n_{41} - n_{43} - n_{44} - n_{51}$. $c(p_1)$:
365 sec, Total distance: 156 Km, Best case average message propagation speed: 1538 Km/h
2. $p_2 : n_1 - n_2 - n_{11} - n_{14} - n_{15} - n_{16} - n_{17} - n_{18} - n_{27} - n_{28} - n_{41} - n_{43} - n_{44} - n_{51}$. $c(p_2)$:
383 sec, Total distance: 180 Km, Best case average message propagation speed: 1691 Km/h
3. $p_3 : n_1 - n_2 - n_3 - n_7 - n_9 - n_{67} - n_{19} - n_{17} - n_{18} - n_{27} - n_{28} - n_{41} - n_{43} - n_{44} - n_{51}$.
 $c(p_3)$: 389 sec, Total distance: 155 Km, , Best case average message propagation speed:
1434 Km/h.
4. $p_4 : n_1 - n_2 - n_3 - n_7 - n_9 - n_{67} - n_{16} - n_{15} - n_{18} - n_{27} - n_{28} - n_{41} - n_{43} - n_{44} - n_{51}$.
 $c(p_4)$: 390 sec, Total distance: 162 Km, , Best case average message propagation speed:
1495 Km/h.

5. $p_5 : n_1 - n_2 - n_{11} - n_{14} - n_{15} - n_{18} - n_{27} - n_{28} - n_{41} - n_{43} - n_{44} - n_{51}$. $c(p_5)$: 398 sec, Total distance: 169 Km, , Best case average message propagation speed: 1528 Km/h.

What is interesting to observe in the quantitative results found above is that an information message may need 6-7 minutes to travel distances of around 150 Km. Even though this may sound a large figure compared to a call using a cell phone, it can be proven crucial in cases of a disaster, where static infrastructure may be destroyed and the mobile telecommunication providers may not be available. In addition, in case of an accident or a catastrophic event, VANETs can send information in the form of video, photos or audio files, offering critical information to the receiver, which might be a police station, a hospital or a fire station.

Even though the field of Vehicular Ad Hoc Networks is widely studied and the concept of proof tests is becoming widely accepted, many years will pass until all vehicles are suitably equipped to participate in VANETs. So, it is very interesting to see how the information delivery probability will behave in more realistic scenarios, where only a portion of vehicles are appropriately equipped to participate in VANETs. In figure 57 we show the information propagation probability versus time for three different percentages of equipped vehicles out of all vehicles traveling on the highways.

We can observe from the figure that the higher the percentage of vehicles that are able to participate in VANETs, the greater the probability of information delivery. That is because having more equipped vehicles in the highways increases the opportunities of information transmission and then increases the speed of message traveling.

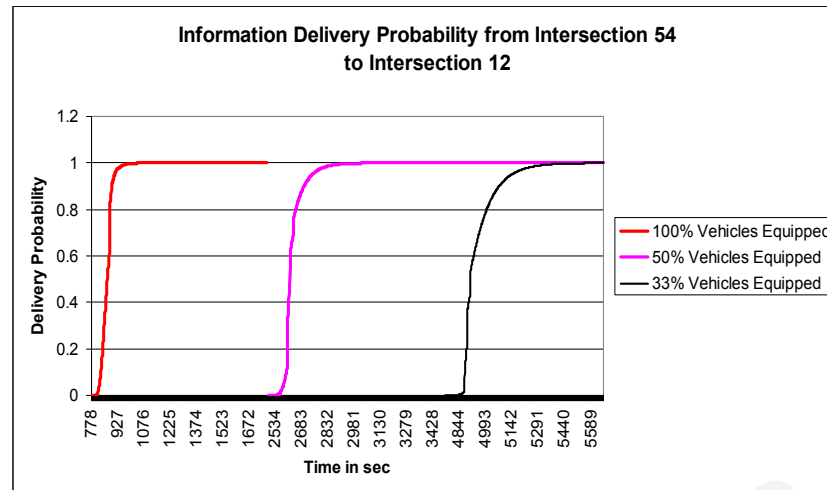


Figure 57: Information Delivery Probability versus Time for three different portions of equipped vehicles

6.6 Simulation Validation

In this section we validate our theoretical findings with simulation results obtained using VISSIM for the last two Source-Destination pair cases. The topology of the reference model that we use in simulations is shown in figure 54. We assume identical traffic conditions (vehicle arrival rate and vehicle speed range) to the ones used in the previous section. In order to make the simulation more realistic, we assume that not all vehicles are equipped to participate in VANETs. More specifically, we assume that 50% of the vehicles have the appropriate equipment that allows them to participate in VANETs, a situation that is expected to hold in the near future. Since we would like the simulation setup to match as closely as possible the theoretical setup considered in the previous section, in the reference model we ignore the on ramp entries and exits of the considered highways. We assume that all vehicles enter each highway at its beginning and exit the highway at its end. This property can easily be implemented on VISSIM, since VISSIM provides the option of creating highways with their own “Vehicle Input”, independent from other roads.

For each scenario that we consider, we calculate the information delivery probability using frequency measurements over 20 simulation runs with different seed numbers. After 1800 seconds, which are more than enough to allow the system to reach an equilibrium state, we drive the output of the simulator to ascii files. The files generated contain the coordinates of each vehicle in every simulation step. The simulation step is set to 200 ms and we consider data for 1200 sec. Due to the fact that the resulting files are huge (around 2GB each), it is difficult to study the information propagation along long distances involving a big number of vehicles and a number of different roads, as we did in the previous section.

In order to process the resulting files we have developed an application in C++ which calculates the time required by a message to travel from the starting intersection until the destination intersection. The transmission range of each vehicle is constant and set to 250 m for each vehicle. Vehicles which lie within the transmission range can exchange messages. Since the transmission speed is much larger than the vehicle speed (approximately 250 m in 6 ms), we consider the transmission time to be negligible. Furthermore, we assume that any retransmissions do not add significantly to the total times. Therefore we do not use any specific wireless communication model.

To find the information delivery probability in a specific time period we consider the frequency with which the information is successfully delivered over 20 simulation runs. The probability is simply estimated by dividing the number of successes (information to be delivered to the destination) with the number of times we have conducted the simulation given a specific time period T .

Next we compare the results obtained using simulations with the theoretical results obtained in the previous section. We consider the scenario where information is delivered from intersections I_{15} to intersection I_{41} . In figure 58 we show plots of the maximum information delivery

probability versus the time period when the former is obtained using simulations and when it is found using equation (61).

The selected path in both cases is the following: $n_{15} - n_{16} - n_{17} - n_{18} - n_{27} - n_{28} - n_{41}$ which is the shortest path. The total distance of the path is 60 Km and the cost of the path is 295 seconds.

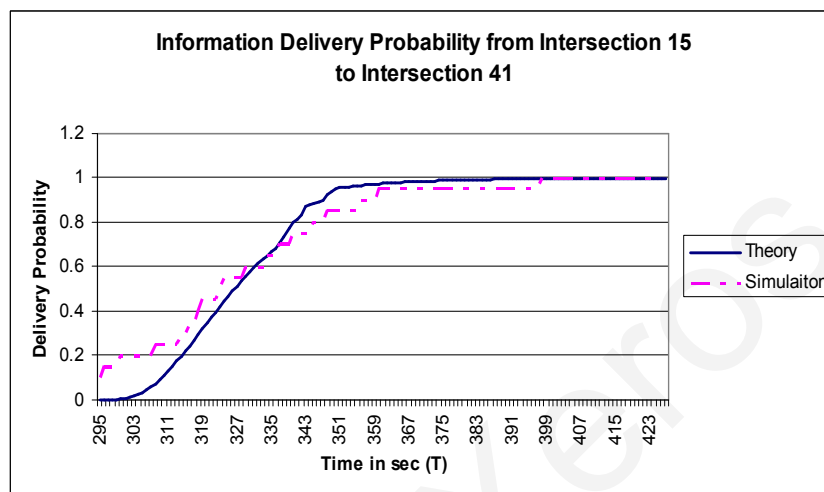


Figure 58: Information Delivery Probability versus Time from intersection I_{15} to intersection I_{41}

Figure 58 indicates that the theoretical results are very similar to the ones obtained using simulations. The small deviations observed are probably due to the relatively small number of simulation runs used to calculate the delivery probability. We observe that the delivery probability calculated using simulations is nonzero for a subset of the range of time values that the theoretical probability is equal to zero. This is due to the fact that the weight of each edge represents the average message propagation time, which implies that in some simulation experiments the reported propagation time was less than the average.

In figure 59 we compare the theoretical and simulation for a different scenario. We consider message delivery from intersection I_3 to intersection I_{22} . In a way similar to the previous case, the maximum probability of information delivery is provided by the shortest path at all time intervals

in both cases. The shortest path is the following: $n_3 - n_7 - n_9 - n_{21} - n_{20} - n_{22}$. The total distance of the path is 76 Km and the cost of the path is 931 seconds.

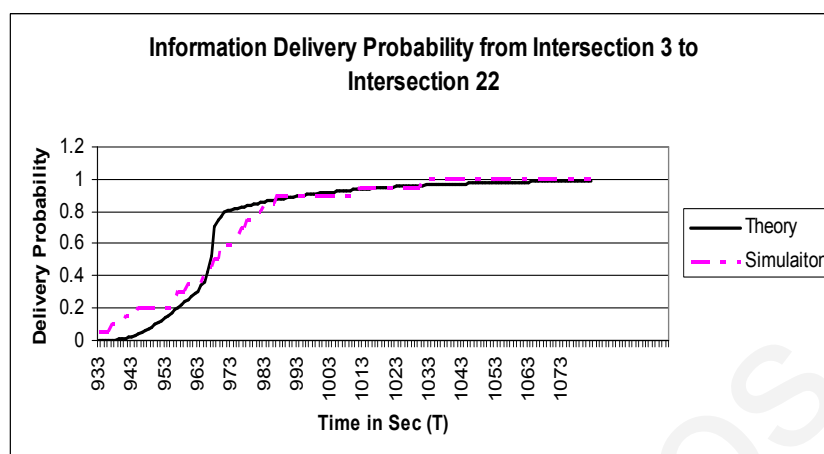


Figure 59: Information Delivery Probability versus Time from intersection I_3 to intersection I_{22}

Again, we observe in figure 59 reasonable agreement between the theoretical and simulation results. Also for the same reasons outlined above, the simulation delivery probability has a nonzero value (between 0 and 0.1) in the time period where the theoretical value is equal to zero.

6.7 Conclusions and Sum Up

In this chapter we study the problem of information dissemination in VANETs and we provide a measure of the probability to have the information be delivered to the destination in a given amount of time where no static infrastructure is used. We present a road map modeling that results in the corresponding Road-Graph given the traffic conditions of the roads. Using the resulting Road-Graph we provide the theoretical analysis, calculate the information delivery probability given a period of time on any source destination pair, and we then select the shortest path based on a cost function which assumes the above probability. We validate our results with simulation evaluation using VISSIM. We use close-to-reality traffic conditions from the highways around the

city of Los Angeles and we study the propagation of information between intersections of the highways.

Andreas Xeros

Chapter 7

Conclusions and Future Work

In this chapter we summarize our conclusions and the main contributions of our research. We also provide possible extensions of our work and future research directions in the general area of information dissemination in VANETs.

VANETs constitute an emerging network technology with expected high returns. VANETs can be expected to harness the potential of information and communication technologies to create a safer, smarter and more efficient transportation network. Their gradual deployment is expected to improve the quality of life by improving safety and comfort in driving, while it is also expected to create a new technology market, which promises to attract significant funding and create new job opportunities contributing in economic and social growth. The recently published 802.11p standard supports both vehicle-to-vehicle and vehicle-to-infrastructure communications, allowing the formation of vehicular ad hoc networks, which are envisioned to accommodate the new generation of cooperative safety applications. The range of applications of VANETs goes beyond the safety related ones to include traffic monitoring, platooning, text messaging, distributed passenger teleconferencing, music downloading, roadside e-advertisements etc.

Many of the aforementioned applications in VANETs, especially the safety related ones, set up requirements for information dissemination that are different from conventional networks and are thus difficult to fulfil with existing strategies. Safety applications pose stringent delay requirements on emergency message delivery and address geographical areas in which data needs to be cooperatively collected, distributed and maintained. Design challenges are then posed by the variable node density along the transportation network, the high mobility, the confined but often unpredictable movement, and the unreliable radio channel. Variations in road traffic density are of particular importance, as low vehicle traffic densities cause the network to become intermittently connected, whereas high vehicle traffic densities lead to excessive contention. These phenomena significantly degrade the performance of data dissemination strategies whether these are routing protocols or broadcast-based schemes.

In this thesis we address the information dissemination problem in VANETs with particular interest in information hovering, broadcast schemes and routing without the necessity for roadside units. We propose new protocols and methods that overcome the design challenges outlined above. Since the system is highly stochastic in nature, our approaches are based on probabilistic methods.

We first consider the information hovering problem in VANETs and propose a traffic aware information hovering protocol that is successful in achieving a high percentage of vehicles receiving the relevant message while at the same time reducing the number of exchanged messages. Information Hovering is a new concept of information dissemination over a mobile set of peers. It naturally applies in many applications in VANETs, where useful information needs to be made available to all vehicles within a confined geographical area for a specific time interval. A straightforward approach is to have all vehicles within the hovering area exchange messages with each other. However, this method does not guarantee that all vehicles within the hovering area will receive the message, because of potential partitioning of the network, especially in areas with

low traffic density and/or low market penetration rate. To alleviate this problem in this work we propose a scheme based on the application of epidemic routing within the hovering area and probabilistic flooding outside the hovering area. Informed vehicles outside the area can serve as information bridges towards partitioned uninformed areas, thus leading to high reachability.

The design methodology has been mainly simulative, and a major challenge in the overall procedure has been the design of the rebroadcast probability function for the probabilistic flooding scheme. Among a number of candidate functions, we choose the one yielding superior performance and we tune its parameters taking advantage of phase transition phenomena, which are typical in probabilistic flooding schemes [61]. A unique feature of the proposed protocol is that it is adaptive, in the sense that the rebroadcast probability outside the hovering area is adaptively regulated based on estimates of the vehicle density within the hovering area. Estimates of the vehicle density within the hovering area are obtained using measurements of the number of neighbors of each vehicle. The formula relating the two quantities is derived using a simple model of the transportation network within the hovering area. We demonstrate through simulations that the proposed scheme is successful in satisfying the design objectives and outperforms other candidate hovering protocols such as epidemic routing in the entire network, epidemic routing in the hovering area only and the scheme proposed in [37]. We also demonstrate that if we utilize the same design methodology to render the protocol proposed in [37] adaptive, its performance becomes comparable to the performance of the scheme proposed in this work. Thus, our contribution goes beyond the proposal of a specific information hovering protocol and extends to the introduction of a design procedure that can be used to design a class of density adaptive hovering protocols.

Next, we demonstrate the universality of the proposed information hovering protocol and its applicability in areas beyond the Vehicular Networking field, by integrating it in a Push based wireless broadcast system as a load offloading tool. The combination of wireless broadcasting

and information hovering is shown to significantly decrease the mean waiting times of clients in various wireless data broadcasting scenarios. In this setup the system's architecture typically considers a large number of clients dispersed over an area covered by a single wireless system. The clients are interested in a common set of data, albeit with varying individual preferences. A central authority creates a common data broadcast schedule and transmits it over the covered area. The central authority is more or less initially unaware of the popularity of each data item. This knowledge is acquired gradually, typically through a lightweight, client feedback mechanism, and the schedule is adapted accordingly. This technique is typically applied to Instant Messaging Services, Tweets and audio-video broadcasting (e.g. DVB/H, DAB+).

In order to be able to reduce the mean waiting time, we present a novel, MANET-oriented merging of wireless broadcasting and information hovering. The logic of the merge follows the fact that any wireless broadcast system is typically assigned bandwidth that is inversely proportional to its area of coverage. Thus, in order for it to achieve adequate mean waiting times, a portion of the data dissemination task can be offloaded to a local hovering scheme. The number of items to be offloaded is analytically evaluated. Simulation under realistic conditions yielded performance superior to individual solution-exclusive approaches, in all cases. It is worth noting that addressing the last problem using the proposed information hovering scheme highlights the generality of the proposed methodology.

Finally, we address the information routing problem in VANETs and utilize a simple mathematical model to find analytically lower bounds on the probability of information propagation on intersections. The probability, as expected, is found to depend heavily on the traffic state in the vicinity of the intersection. The derived result is interesting and significant in its own respect, but may also constitute a first step in developing a traffic-aware routing protocol. The probability of successful information delivery along a straight line road has been derived analytically in the

literature. Combining the latter with the result obtained in this thesis, one can find the probability of successful information delivery along any route in a city map. Such a probability can be used as a cost measure for the specific route in a graph theoretic formulation of the routing problem. Solutions of the routing problem are readily available in the literature, and one can then use such solutions as a baseline to develop traffic aware routing protocols for VANETs. In this thesis we adopt the aforementioned approach and examine whether the information propagation probability on intersections derived in this work combined with the information propagation probability along a straight line road can yield a formula that gives the probability to deliver messages to a specific area in a certain amount of time and in a setting where no static infrastructure is used. We use graph theoretical approach where we model the road map and the traffic characteristics using a directed weighted graph, which we refer to as the Road-Graph. We use the Road-Graph to calculate the probability of disseminating information along any given path and in a specific amount of time. So, for any two points on the Road-Graph we can find the path with the maximum probability of successful information delivery in the chosen time interval. This information is of great significance, since it can be utilized by routing protocols to optimally route packets in the vehicular network. We validate our analytical findings by comparing them with data obtained using VISSIM.

As indicated above, in this thesis we address several aspects of the information dissemination problem in VANETs and we propose a number of solutions based on probabilistic methods. Our work is significant in that it offers practical solutions to a number of open problems in VANETs, but it also demonstrates the potential of these probabilistic methods in developing effective solutions for a variety of problems. Below we list our main contributions:

1. We develop a new hovering scheme which is successful in accomplishing its design objectives as it achieves high reachability, while at the same time alleviating the broadcast

storm problem by reducing the number of exchanged messages. In order to allow message delivery in case of partitioned hovering areas we allow message dissemination outside the hovering area. For alleviating the broadcast storm problem, the protocol employs epidemic routing within the hovering area and probabilistic flooding outside the hovering area. A unique feature of the protocol is that it is traffic-adaptive, in the sense that the rebroadcast probability is calculated based on estimates of the vehicle density inside the hovering area, which are in turn based on estimates of the average number of neighbors.

2. We utilize a simple model of the roadway topology to develop a formula that estimates the average vehicle density within the hovering area based on measurements of the average number of neighbors readily available to each node using a beaconing method. The theoretical result is validated using simulations.
3. The methodology used to design the rebroadcast probability function is based on the existence of phase transition phenomena, which are typical when using probabilistic flooding outside the hovering area. In this thesis we verify the existence of phase transition phenomena using simulations. When utilizing probabilistic flooding, we observe that, as the rebroadcast probability increases, the reachability remains low until a critical value is reached beyond which the reachability attains a relatively high value. We thus identify a phase of low reachability and a phase of high reachability. The probability at which the transition takes place depends on the traffic state in the area of interest.
4. For comparison purposes, in this thesis we design an alternative traffic adaptive hovering scheme, which attempts to alleviate the low reachability problem in cases of low traffic density, by extending the area in which epidemic routing is allowed beyond the hovering area. The vehicles outside the hovering that can participate in the message exchange process

serve as information bridges towards the partitioned uninformed areas. The size of the extended hovering area is adaptive, in the sense that it changes based on estimates of the traffic density. In addition, the design procedure has been the same with the procedure adopted to design the proposed hovering scheme that applies probabilistic flooding outside the hovering area. Comparing the extended area approach and the probabilistic approach we observe very similar behavior. So, our contribution goes beyond the proposition of a new hovering scheme and includes the proposition of a new design methodology, which can yield a class of information hovering schemes with similar characteristics.

5. We present a novel MANET-oriented merging of wireless broadcasting and information hovering. The logic of the merge follows the fact that any wireless broadcast system is typically assigned bandwidth that is inversely proportional to its area of coverage. Thus, in order for the system to achieve adequate mean waiting times, a portion of the data dissemination task can be offloaded to a local hovering scheme. The number of items to be offloaded is defined heuristically. Two cooperation schemes are proposed, targeting solely mean waiting time (MWT) minimization and MWT-hovering bandwidth conservation balancing respectively. Simulation under realistic conditions yielded performance superior to stand-alone approaches, in all cases. It is worth noting that addressing the last problem using the proposed information hovering scheme highlights the generality of the proposed methodology.
6. We use a simple model of the roadway network to develop a lower bound on the probability to propagate information along an intersection. This lower bound depends heavily on the traffic state and is validated using simulations. The result is important as it can be used in the development of traffic aware routing protocols.

7. We formulate the routing problem in VANETs using a graph theoretic approach and for a specific route we find analytically the probability of successful information delivery in a specific amount of time. We use this probability to find the optimal route which maximizes the probability of information delivery. Our analytical findings are again validated using simulations. The simulations are conducted on the VISSIM simulator and the reference model constitutes a section of the freeway system in the Los Angeles Area.

The aforementioned contributions open new avenues of research activity pertinent to the analyzed subjects. Our future plans include:

1. The approach adopted to design the information hovering protocol has been mainly simulative. Despite the successful application of this approach, there is still a great need to support these findings with analytical results. The design of protocols with analytically verifiable properties prior to implementation is in many cases a big engineering challenge. Thus, our objective is to consider simple mathematical models of the roadway topology to derive analytical results which support our simulative findings. We aim at formulating the reachability problem as a connectivity problem in a graph theoretic framework and calculating the probability of all vehicles receiving the relevant message as a function of the rebroadcast probability. Based on this analysis we aim at identifying phase transition phenomena typical when using probabilistic flooding.
2. For the information hovering protocol we calculate the rebroadcast probability function outside the hovering area based on only two criteria: traffic density and distance from the hovering area. By adding more input variables to the rebroadcast probability function, we can improve both the reachability and the number of messages exchanged. An example of

such an input variable is the driving direction of the vehicle. Vehicles moving towards the hovering area should have higher rebroadcast probability than the ones moving away.

Another input variable can be the road topology around the hovering area. The road topology can be used to build rules to decide the rebroadcast probability. For example, if there are few number of road sections around the hovering area then the probability should be larger than the case of many road sections around the area.

The hovering protocol can be further enhanced by more simple and efficient estimates of the traffic density. For example, in highway settings, the speed can be used to estimate the traffic density [81]. Such approaches can be tested in the future and different estimation schemes for in-city streets can also be investigated.

In this thesis we develop a solution to the problem of low reachability in cases of low traffic density in the hovering area. The protocol does not account for the cases of high traffic density within the hovering area. If epidemic routing is used then the broadcast storm problem can lead to low latency of message delivery. A number of techniques have been proposed in literature to alleviate this problem. Such techniques can be integrated in the proposed solution and the evaluation of the resulting protocol will be the topic of future research.

3. In an environment with a Push Server we can utilize VANETs to improve mean waiting times. We can also use VANETs to have informed vehicles outside the transmission range of the Push Server. Moreover, message exchange outside the range can be intelligently selected, so that further reductions in mean waiting times can be achieved. Several intelligent rules can be considered and their impact on the overall system performance can be evaluated.

4. Lower bounds on the information propagation probability on intersections are important since they can be combined with information propagation probability results along a straight line road to derive cost metrics in a graph theoretic formulation of the routing problem in VANETs. Such metrics are vital in the development of a traffic adaptive routing scheme. Despite having obtained such cost metrics in this thesis, our next objective is to improve these metrics by improving the underlying models on which the analytical results are based. By reducing the underlying assumptions and making these models more realistic, one can develop routing protocols with improved properties. Towards this end we aim at extending the roadway models to include two-way roads with reverse traffic. Such model extensions can significantly affect the complexity of the analysis, but will lead to more realistic results. Our initial objective is to pursue formulas that can be used to calculate the probability of successful information delivery along a straight line road when this straight line road consists of two lines and accounts for reverse traffic on one lane.
5. The analysis presented in this thesis and the extensions mentioned above aim at deriving link cost metrics in a graph theoretic formulation of the routing problem in VANETs. The next step is to finalize this formulation and derive routing algorithms that solve the considered problem. In addition, the derived routing algorithms can be used as a baseline to develop new routing algorithms for VANETs. Our objective is to derive such routing algorithms and protocols. We aim at deriving routing protocols that are traffic adaptive and take routing decisions that maximize the probability of successful information delivery from source to destination. Both distance vector and link state algorithms will be investigated, though distance vector algorithms are more attractive due to their decentralized nature and the smaller communication burden that they impose in the network. In such a distance vector approach

intersections will serve as the nodes, and the roads connecting these intersections will serve as the links. Intersections will maintain states regarding their traffic and that of their neighbors. This information will be exchanged periodically between the nodes. These states will be maintained at each intersection either using static infrastructure or using the proposed information hovering scheme. The information exchange protocol between the intersections will also be investigated. These are some early ideas which will constitute the basis of future research pertinent to routing protocol design.

Having demonstrated the applicability of the proposed information hovering protocol in areas beyond the Vehicular networking field by successfully integrating it in a wireless broadcast scheme, our future work will investigate further applications of the information hovering protocol in other areas such as cloud computing, mobile cloud and the Boomerang Protocol.

In recent years, “cloud computing” has emerged as an important trend in information technology. The US National Institute of Standards and Technology (NIST) defines cloud computing as a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud computing is still an evolving paradigm and covers a large ecosystem of many models, vendors, and market niches, according to NIST [39].

Mobile cloud computing is the usage of cloud computing in combination with mobile devices. Cloud computing exists when tasks and data are kept on the Internet rather than on individual devices, providing on-demand access. Applications are run on a remote server and then sent to the user. Because of the advanced improvement in mobile browsers over the past couple of years, nearly every mobile should have a suitable browser. With cloud computing, developers with innovative ideas for new Internet applications and services are no longer required to have large

capital outlays in hardware to deploy their service or, more importantly, the human expense to operate it. They also do need to be concerned about over-provisioning for services with popularity that does not meet their predictions and the market analysis, thus wasting costly resources, or under-provisioning for one that becomes wildly popular missing potential customers and revenue.

The applicability of mobile cloud computing in VANETs is a new research area presented in [89] and described in more details in [88]. The huge fleet of energy-sufficient vehicles that criss-cross our roadways, airways, and waterways, featuring substantial on-board computational, storage, and sensing capabilities can be thought of as a huge farm of computers on the move. These attributes make vehicles ideal candidates for nodes in a cloud. Indeed, the owner of a vehicle may decide to rent out their in-vehicle capabilities on demand, or a per instance, or a per-day, per-week or per-month basis, just as owners of large computing facilities find it economically appealing to rent out excess capacity to seek pecuniary advantages.

Furthermore, vehicles perform increasingly more complex data collection through a variety of sensors and data processing such as: road alarms (pedestrian crossing, electronic brake lights, etc), cooperative content downloading via P2P car-torrent, urban surveillance (video, mechanical, chemical sensors), road mapping via “crowd sourcing”, accident and crime witnessing (for forensic investigations). Uploading this vast amount of data on the Internet for further processing is extremely expensive due to scarce resources. This necessitates the deployment of a new network and computing paradigm widely known as vehicular cloud computing. Instead of uploading the data and on the Internet it is kept and processed in a local vehicle cloud. As opposed to Internet clouds, vehicular clouds are mainly location relevant [27].

Authors believe it is only a matter of time before the huge vehicular fleets on our roadways, streets and parking lots will be recognized as an abundant and underutilized computational resource that can be tapped into for the purpose of providing third-party or community services.

However, what distinguishes vehicles from standard nodes in a conventional cloud is autonomy and mobility. Indeed, large numbers of vehicles spend substantial amounts of time on the road and may be involved in dynamically changing situations. Thus, vehicles have the potential to cooperatively solve problems that would take a centralized system an inordinate amount of time, rendering the solution useless. We aim at addressing these issues in the future and propose solutions guided by the methods and algorithms proposed in this thesis.

Another research area that we aim to investigate is related to the Boomerang Protocol. In 2011 the Boomerang Protocol [107] was introduced, which can efficiently retain information at a particular geographic location in a sparse network to highly mobile nodes without using infrastructure. The protocol first allows a mobile node to carry packets away from their location of origin and periodically returns them to the anchor location. A unique feature of this protocol is that it records the geographical trajectory while moving away from the origin and exploits the recorded trajectory to optimize the return path. This feature can be utilized by our proposed hovering protocol to improve performance both by increasing reachability and decreasing the number of message exchanged, even with a small tradeoff in increasing the computational complexity.

Appendix A

VISSIM is a widely used microscopic multi-modal traffic flow simulation software. It can analyze private and public transport operations under constraints such as lane configuration, traffic composition, traffic signals, PT stops, etc., thus making it a useful tool for the evaluation of various alternatives based on transportation engineering and planning measures of effectiveness. VISSIM can be applied as a useful tool in a variety of transportation problem settings. The following list provides a selective overview of previous applications of VISSIM:

- Development, evaluation and fine-tuning of signal priority logic: VISSIM can use various types of signal control logic. In addition to the built-in fixed-time functionality there are several vehicle-actuated signal controls identical to signal control software packages installed in the field. In VISSIM some of them are built-in, some can be docked using add-ons and others can be simulated through the external signal state generator (VAP) that allows the design of user-defined signal control logic. Thus virtually every signal control can be modeled and simulated within VISSIM if either the controller details are available or there is a direct VISSIM interface available (e.g. VS-PLUS).
- Evaluation and optimization of traffic operations in a combined network of coordinated and actuated traffic signals.
- Feasibility and traffic impact studies of integrating light rail into urban street networks.
- Analysis of slow speed weaving and merging areas.
- Easy comparison of design alternatives including signalized and stop sign controlled intersections, roundabouts and grade separated interchanges.
- Capacity and operations analyses of complex station layouts for light rail and bus systems have been analyzed with VISSIM.
- Preferential treatment solutions for buses (e.g. queue jumps, curb extensions, bus-only lanes) have been evaluated with VISSIM.
- With its built-in Dynamic Assignment model, VISSIM can answer route choice dependent questions such as the impacts of variable message signs or the potential for traffic diversion into neighborhoods for networks up to the size of medium sized cities.
- Modeling and simulating flows of pedestrians - in streets and buildings - allow for a wide range of new applications. VISSIM can also simulate and visualize the interactions between road traffic and pedestrians.

The accuracy of a traffic simulation model is mainly dependent on the quality of the vehicle modeling, e.g. the methodology of moving vehicles through the network. In contrast to less complex models using constant speeds and deterministic car following logic, VISSIM uses the psycho-physical driver behavior model developed by WIEDEMANN (1974). The basic concept of this model is that the driver of a faster moving vehicle starts to decelerate as he/she reaches

his/her individual perception threshold to a slower moving vehicle. Since he/she cannot exactly determine the speed of that vehicle, his/her speed will fall below that vehicles speed until he/she starts to slightly accelerate again after reaching another perception threshold. This results in an iterative process of acceleration and deceleration. Stochastic distributions of speed and spacing thresholds replicate individual driver behavior characteristics. The model has been calibrated through multiple field measurements at the Technical University of Karlsruhe, Germany. Periodical field measurements and their resulting updates of model parameters ensure that changes in driver behavior and vehicle improvements are accounted for. VISSIMs traffic simulator not only allows drivers on multiple lane roadways to react to preceding vehicles (4 by default), but also neighboring vehicles on the adjacent travel lanes are taken into account. Furthermore, approaching a traffic signal results in a higher alertness for drivers at a distance of 100 meters in front of the stop line. VISSIM simulates the traffic flow by moving “driver-vehicle-units” through a network. Every driver with his/her specific behavior characteristics is assigned to a specific vehicle. As a consequence, the driving behavior corresponds to the technical capabilities of his/her vehicle. Attributes characterizing each driver-vehicle unit can be discriminated into three categories:

- Technical specification of the vehicle, for example:
 - Length
 - Maximum speed
 - Potential acceleration
 - Actual position in the network
 - Actual speed and acceleration
- Behavior of driver-vehicle units, for example:
 - Psycho-physical sensitivity thresholds of the driver (ability to estimate, aggressiveness)
 - Memory of the driver
 - Acceleration based on current speed and the drivers desired speed
- Interdependence of driver-vehicle units, such as:
 - Reference to leading and following vehicles on own and adjacent travel lanes
 - Reference to current link and next intersection
 - Reference to next traffic signal

Bibliography

- [1] A. Abbas, and O. Kure, "A Quality of Service in Mobile Ad hoc Networks: A Survey", in *International Journal of Ad hoc and Ubiquitous Computing (IJAHUC)*, vol. 6, no. 2, pp. 75-98, 2010.
- [2] S. Acharya, R. Alonso, M. Franklin, and S. Zdonik, "Broadcast disks: data management for asymmetric communication environments", in *Proc. of the ACM SIGMOD International Conference on Management of data SIGMOD*, vol. 24, no. 2, pp. 199-210, 1995.
- [3] K. Alzoubi, P. Wan, and O. Frieder, "New Distributed Algorithm for Connected Dominating Set in Wireless Ad Hoc Networks", in *Proc. of the 35th Annual Hawaii International Conference on System Sciences*, vol. 9, pp. 297, 2002.
- [4] H. Arbabi, and M. C. Weigle, "Monitoring Free-Flow Traffic using Vehicular Networks," in *Consumer Communications and Network Conference (CCNC)*, pp. 272-2760, 2011.
- [5] J. Arunadevi, and V. Rajamani, "Optimized routing in Mobile Ad Hoc Networks using Evolutionary Location Intelligence", *IJCA Special Issue on "Mobile Ad-hoc Networks" MANETs*, pp. 120-123, 2010.
- [6] L. Briesemeister, and G. Hommel, "Role-based multicast in Highly Mobile but Sparsely Connected Ad-Hoc Networks", in *Proc. of the 1st ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 45-50, 2000.
- [7] I. Broustis, and M. Faloutsos, "Routing in Vehicular Networks: Feasibility, Security and Modeling Issues", in *Technical Report No. UCRCS- 2006-05219, Dept of Computer Science and Engineering*, University of California, Riverside, 2006.
- [8] K.V.R.H Bharadwaj and V. Sunitha, "Application of random graph theory and percolation theory in MANETs". In *Technical Report*, 2004.
- [9] G. Calinescu, I. Mandoiu, P. J. Wan, and A. Zelikovsky, "Selecting Forwarding Neighbors in Wireless Ad Hoc Networks", in *Proc. ACM Int'l Workshop Discrete Algorithms and Methods for Mobile Computing*, pp. 34-43, 2001.
- [10] Car-to-Car Communication Consortium (C2C-CC), [Online], Available: <http://www.car-2-car.org>.
- [11] P. C. Cheng, J. T. Weng, L. C. Tung, K. C. Lee, M. Gerla, and J. Haerri, "GeoDTN+NAV: A Hybrid Geographic and DTN Routing with Navigation Assistance in Urban Vehicular Networks", in *Proc. of the 1st International Symposium on Vehicular Computing Systems (ISVCS'08)*, Dublin, Ireland, 2008.

- [12] I. Chisalita, and N. Shahmehri, "A context based vehicular communication protocol", in *IEEE Personal, Indoor and Mobile Radio Communication Symposium*, pp. 5-8, 2004.
- [13] J. Dargay, and H. Huntington, "Vehicle ownership and income growth, worldwide: 1960-2030", in *The Energy Journal*, vol. 28, No. 4, 2007.
- [14] S. Das, A. Nandan, and G. Pau, "Spawn: a swarming protocol for vehicular ad-hoc wireless networks", in *VANET'04*, pp. 93-94, 2004.
- [15] M. Dikaiakos, A. Florides, T. Nadeem and L. Iftode, "Location-Aware Services over Vehicular Ad-Hoc Networks using Car-to-Car Communication", in *IEEE Journal on Selected Issues in Communications (JSAC) - Special Issue on Vehicular Networks*, Vol. 25, No. 8, pp. 1590-1602, 2007.
- [16] N. Dragos, and B. Nath, "Trajectory based Forwarding and its Applications", in *Proc. of the ACM International Conference on Mobile Computing and Networking (MOBICOM'03)*, 2003.
- [17] IEEE Standards Association. [Online]. Available: <http://standards.ieee.org/>. [Accessed July, 2009].
- [18] R. Eberhardt, and C. Maihofer, "Virtual warning signs: A geocast enabled service for ad hoc networks", in *Proc. of the 3rd Workshop on Applications and Services in Wireless Networks (ASWN 2003)*, pp. 135-147, Switzerland, 2003.
- [19] P. Erdos and A. Renyi. On the evolution of random graphs. In *Publ. Math. Inst. Hung. Acad. Sci*, 5:1761, 1960.
- [20] ETSI, [Online], Available: <http://www.etsi.org>.
- [21] M. U. Farooq, J. A. Shadab, K. U. R. Khan, and J. C. MuffaKham, "An effective way of broadcasting alert messages in vehicular Ad hoc Network", in *Communication and Computational Intelligence (INCOCCI), 2010 International Conference*, pp. 132-137, 2010.
- [22] K. Fawaz, A. Ghandour, M. Olleik, and H. Artail, "Improving Reliability of Safety Applications in Vehicle Ad hoc Networks through the Implementation of a Cognitive Network", *17th IEEE International Conference on Telecommunications 2010 (ICT2010)*, pp. 798-805, Doha, Qatar, 2010.
- [23] H. Fissler, M. Mauve, H. Hartenstein, M. Ksemann, and D. Vollmer, "A comparison of routing strategies for vehicular ad-hoc networks", (*poster*), *ACM MobiCom*, 2002.
- [24] H. Fubler, H. Hannes, W. Jorg, M. Martin, and E. Wolfgang, "Contention-Based Forwarding for Street Scenarios", in *Proc. of the 1st International Workshop in Intelligent Transportation (WIT 2004)*, pp. 155-160, Hamburg, Germany, 2004.
- [25] V. Gau, and J. Hwang, "Adaptive Probabilistic Broadcasting over Dense Wireless Ad Hoc Networks", in *Proc. of Int. J. Digital Multimedia Broadcasting*, 2010.
- [26] J. Gecsei, "The Architecture of Videotex Systems", *Prentice-Hall*, Englewood Cliffs, NJ, 1983.

- [27] M. Gerla, "Vehicular Cloud Computing", *Ad Hoc Networking Workshop (Med-Hoc-Net)*, pp. 152-155, Cyprus 2012.
- [28] S. Ghandeharizadeh, S. Kapadia, and B. Krishnamachari, "Comparison of replication strategies for content availability in C2P2 networks", in *Proc. of the 6th international conference on Mobile*, pp. 107-115, Cyprus, 2005.
- [29] Google Earth, [Online], Available: <http://earth.google.com/>, 2008.
- [30] Q. Guan, F.R. Yu, S. Jiang, and G. Wei, "Prediction-Based Topology Control and Routing in Cognitive Radio Mobile Ad Hoc Networks", in *IEEE Trans. Vehicular Technology*, vol. 59, no. 9, pp. 4443-4452, 2010.
- [31] Y. Guo, S. K. Das, and C. M. Pinotti, "A new hybrid broadcast scheduling algorithm for asymmetric communication systems", in *Proc. of the 4th ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems (MSWIM'01), Co-located with MOBICOM'01*, pp. 123-130, Rome, Italy, 2001.
- [32] K. A. Hafeez, L. Zhao, Z. Liao, and B. N. Ma, "A New Broadcast Protocol for Vehicular Ad Hoc Networks Safety Applications", in *Proc. of GLOBECOM*, pp. 1-5, 2010.
- [33] M. Hamdy, A. Derhab, and B. Konig-Ries, "A Comparison on MANETs' Service Replication Schemes: Interest versus Topology Prediction", *The Second International Conference on Wireless and Mobile Networks (WiMo-2010)*, June 2010.
- [34] S. Hameed, and N. Vaidya, "Efficient algorithms for scheduling data broadcast," *ACM/Baltzer Wireless Networks*, vol. 5, no. 3, pp. 183-193, 1999.
- [35] A. Hamra, C. Barakat, and T. Turletti, "Network Coding for Wireless Mesh Networks: A Case Study," in *2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'06)*, IEEE, pp. 103-114, 2006.
- [36] J. Harri, F. Filali, and C. Bonnet, "Mobility models for vehicular ad hoc networks: a survey and taxonomy", *IEEE Communications Surveys and Tutorials*, pp. 1941, 2009.
- [37] S. Hermann, C. Michl, and A. Wolisz, "Time-Stable Geocast in Intermittently Connected IEEE 802.11 MANETs", in *Proc. of IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1922-1926, 2007.
- [38] J. Hersheberer, M. Maxel, and S. Suri, "Finding the k Shortest Simple Paths: A New Algorithm and its Implementation", in *ACM Transactions on Algorithms (TALG)*, Vol.3, Issue 4, Article No. 45, USA, 2007.
- [39] NIST - National Institute of Standards and Technology, [Online], Available: (<http://csrc.nist.gov/groups/SNS/cloud-computing/>)
- [40] C. Hu, Y. Hong, and J. Hou, "On mitigating the broadcast storm problem with directional antennas" in *Proc. IEEE International Conference, on Communications*, vol. 1, pp. 104-110, 2003.
- [41] IEEE 802.11 WG, Part 11, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE, 1999.

- [42] IEEE P1609.1, Trial-Use Standard for Wireless Access in Vehicular Environments (WAVE) - Resource Manager, 2006.
- [43] IEEE P1609.2, Trial-Use Standard for Wireless Access in Vehicular Environments (WAVE) - Security Services for Applications and Managements Messages, 2006.
- [44] IEEE P1609.3, Trial- Use Standard for Wireless Access in Vehicular Environments (WAVE) - Networking Services, 2007.
- [45] IEEE P1609.4, Trial- Use Standard for Wireless Access in Vehicular Environments (WAVE) - Multi-Channel Operation, 2006.
- [46] IEEE P802.11p, *Amendment 3: wireless access in vehicular environments (WAVE)* Draft D0.26, 2006.
- [47] IEEE Std 802.11, "IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements Part 11", *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2007.
- [48] IETF, [Online], Available: <http://www.ietf.org>.
- [49] "Internet Protocol Version 6 (IPv6) Specification", *IETF*, RFC 2460.
- [50] ISO, [Online], Available: <http://www.iso.org>.
- [51] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs, and N. Braynard, "Networking named content", in *Proc. of the 5th ACM Conference on Emerging Network Experiment and Technology (CoNEXT'09)*, pp. 1-12, Italy, 2009.
- [52] M. Jerbi, R. Meraihi, M. Senouci, and Y. Doudane, "GyTAR: improved Greedy Traffic Aware Routing Protocol for Vehicular Ad Hoc Networks in City Environments", in *Proc. of the 3rd International workshop on Vehicular ad hoc network*, pp. 88-89, 2006.
- [53] F. Jeremy and U. M. Charles. Phase transitions in the properties of random graphs. In *Principles and Practice of Constraint Programming (CP-95)*, pages 6269, 1995.
- [54] S. H. Kang, S. Choi, S. J. Choi, G. Lee, J. Lew, and J. Lee, "Scheduling data broadcast based on multi-frequency in mobile interactive broadcasting", *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 405-411, 2007.
- [55] M. Kasemann, "A Reactive Location Service for Mobile Ad Hoc Networks", *Technical Report TR-02-014*, Department of Computer Science, University of Mannheim, 2002.
- [56] K. Katsaros, M. Dianati, R. Tafazolli, and R. Kernchen, "CLWPR - A Novel Cross-Layer Optimized Position Based Routing Protocol for VANETs" in *IEEE Vehicular Networking Conference*, pp. 139-146, 2011.
- [57] Z. K. Kayhan, and A. B. Kamalrulnizam, "Inter-Vehicle Communication Protocols for Multimedia Transmission", in *Proc. of the International MultiConference of Engineers and Computer Scientists*, pp. 17-19, Hong Kong, China, 2010.

- [58] D. Konstantas, and A. Villalba, "Hovering Information: A paradigm for sharing location-bound information," in *Proc. of the 2006 International Conference on Intelligent Systems And Computing: Theory And Applications (ISYC'06)*, pp. 116-122, 2006.
- [59] G. Korkmaz, and E. Ekici, "An Efficient Fully Ad-Hoc Multi-Hop Broadcast Protocol for Inter-Vehicular Communication Systems", in *Proc. of IEEE International Conference on Communications*, vol. 1, pp. 104-110, 2006.
- [60] M. Kornfeld, "DVB-H - the emerging standard for mobile data communication", in *IEEE International Symposium on Consumer Electronics*, IEEE, pp. 193-198, 2004.
- [61] B. Krishnamachari, S. B. Wicker, and R. Bejar, "Phase Transition Phenomena in Wireless Ad-Hoc Networks", in *Proc. of IEEE Globecom'01*, Texas, USA, 2001.
- [62] K. C. Lee, U. Lee, and M. Gerla, "TO-GO: TOpology-assist geo-opportunistic routing in urban vehicular grids", in *Proc. of the Sixth International Conference on Wireless On-Demand Network Systems and Services (WONS 2009)*, pp. 11-18, 2009.
- [63] S. H. Lee, U. Lee, K. W. Lee, and M. Gerla, "Content Distribution in VANETs Using Network Coding: The Effect of Disk I/O and Processing O/H", in *Proc. of the 5th Annual IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON'08)*, pp. 117-125, San Francisco, CA, 2008.
- [64] B. Lee, B. Seet, G. Liu, C. Foh, and K. Wong, "A Routing Strategy for Metropolis Vehicular Communications", in *Proc. of the International Conference on Information Networking (ICOIN04)*, vol.2, pp 533-542, 2004.
- [65] M. F. Leung, and S. H. G. Chan, "Broadcast-Based Peer-to-Peer Collaborative Video Streaming Among Mobiles," *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 350-361, 2007.
- [66] M. Li, Z. Yang, and W. Lou, "CodeOn: Cooperative Popular Content Distribution for Vehicular Networks using Symbol Level Network Coding", in *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 1, pp. 223-235, 2011.
- [67] C. Liaskos, and G. Papadimitriou, "Ultra lightweight adaptation processes for scheduling servers in push-based systems", in *17th Annual Symposium of the IEEE/CVT*, pp. 1-6, 2010.
- [68] C. Liaskos, S. Petridou, and G. Papadimitriou, "Cost-aware wireless data broadcasting", *IEEE Transactions on Broadcasting*, vol. 56, no. 1, pp. 66-76, 2010.
- [69] C. Liaskos, S. Petridou, and G. Papadimitriou, "Towards realizable, low cost broadcast systems for dynamic environments", *IEEE Transactions on Networking*, vol. 19, pp.383-392, 2011.
- [70] C. Liaskos, S. Petridou, G. Papadimitriou, P. Nicopolitidis, M. Obaidat, and A. Pomportsis, "Clustering-driven Wireless Data Broadcasting", *IEEE Wireless Comm. Magazine*, vol. 16, pp. 80-87, 2009.
- [71] C. Liaskos, S. Petridou, G. Papadimitriou, P. Nicopolitidis, and A. Pomportsis, "On the Analytical Performance Optimization of Wireless Data Broadcasting," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 884-895, 2010.

- [72] C. Liaskos, A. Xeros, G. Papadimitriou, M. Lestas, and A. Pitsillides, "Balancing Wireless Data Broadcasting and Information Hovering for Efficient Information Dissemination", *IEEE Transactions on Broadcasting*, Accepted for Publication, pp. 66-76, 2011.
- [73] H. Lim, and C. Kim, "Flooding in Wireless Ad Hoc Networks" in *Computer Communications*, vol. 24, issues 3-4, pp. 353-363, 2001.
- [74] T. Little, and A. Agarwal, "An information propagation scheme for VANETs", in *MCL Technical Report 07-01-2005*, Department of Electrical and Computer Engineering, Boston University, Boston, USA, 2005.
- [75] C. Lochert, H. Hartenstein, J. Tian, H. Fussler, D. Hermann, M. Mauve, "A Routing Strategy for Vehicular Ad Hoc Networks in City Environments", in *Intelligent Vehicles Symposium, IV*, Ohio, USA, pp. 156-161, 2003.
- [76] S. Louca, A. Pitsillides, and G. Samaras, "On Network Survivability Algorithms Based on Trellis Graph Transformations", in *ISCC(1999)*, pp. 235-243, 1999.
- [77] Y. Ma, and A. Jamalipour, "A cooperative cache-based content delivery framework for intermittently connected mobile ad hoc networks", in *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 366-373, 2010.
- [78] C. Maihofer, "A survey on geocast routing protocols", in *IEEE Communications Surveys and Tutorials*, 2nd quarter issue, vol.6, 2004.
- [79] C. Maihofer, T. Leinmuller, and E. Schoch, "Abiding geocast: time-stable geocast for ad hoc networks", in *Proc. of the 2nd ACM international workshop on Vehicular ad hoc networks (VANET05)*, pp. 20-29, 2005.
- [80] N. Meghanathan, "A Location Prediction Based Routing Protocol and its Extensions for Multicast and Multi-path Routing in Mobile Ad hoc Networks", *Elsevier Ad hoc Networks*, vol. 9, no. 7, pp. 1104-1126, 2011.
- [81] Y. Mylonas, M. Lestas, A. Pitsillides, and P. Ioannou, "Speed Adaptive Probabilistic Flooding for Vehicular Ad-Hoc Networks", in *IEEE Transactions on Intelligent Transportation Systems*, pp. 719-723, 2009.
- [82] T. Nadeem, S. Dashtinezhad, C. Liao, and L. Iftode, "TrafficView: A Scalable Traffic Monitoring System", in *Proc. of the IEEE International Conference on Mobile Data Management*, pp. 13-26, 2004.
- [83] T. Nadeem, P. Shankar, and L. Iftode, "A Comparative Study of Data Dissemination Models for VANETs", in *Proc. of the 3rd Annual International Conference on Mobile and Ubiquitous Systems (MOBIQUITOUS)*, pp. 1-10, 2006.
- [84] National Highway Traffic safety Administration, [Online], Available: [http : //www.nhtsa.gov](http://www.nhtsa.gov).
- [85] D. Q. Nguyen, P. Minet, T. Kunz, and L. Lamont, "On the Selection of Cluster Heads in MANETs", in *International Journal of Computer Science Issues (IJCSI)*, vol. 8, issue 1, 2011.

- [86] P. Nicopolitidis, G. Papadimitriou, and A. Pomportsis, "Continuous Flow Wireless Data Broadcasting for High-Speed Environments", in *IEEE Transactions on Broadcasting*, vol. 55, no. 2, pp. 260-269, 2009.
- [87] P. Nicopolitidis, G. Papadimitriou, and A. Pomportsis, "Exploiting locality of demand to improve the performance of wireless data broadcasting", in *IEEE Transactions on Vehicular Technology*, vol. 55, no. 4, pp. 1347-1361, 2006.
- [88] S. Olariu, M. Eltoweissy, and Mohamed Younis, "Towards autonomous vehicular clouds", in *ICST Transactions on Mobile Communications and Applications*, vol. 11, 2011.
- [89] S. Olariu, I. Khalil, and M. Abuelela, "Taking VANET to the clouds", in *International Journal of Pervasive Computing and Communications*, vol. 7, issue 1, pp. 7 - 21, 2011.
- [90] C. Palazzi, M. Rocchetti, and S. Ferreti, "Online Games on Wheels: Fast Game Event Delivery in Vehicular Ad-hoc Networks", in *V2VCOM*, 2007.
- [91] G. Papadimitriou, M. Sklira, and A. Pomportsis, "A new class of optimal learning automata", in *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 34, no. 1, pp. 246-254, 2004.
- [92] M. Peden, R. Scurfield, and D. Sleet, "World report on road traffic injury prevention", *World Health Organization*, [http : //who.int/violence_injury_prevention/publications/road_traffic/world_report/en/](http://who.int/violence_injury_prevention/publications/road_traffic/world_report/en/), Retrieved 2008-06-24.
- [93] E. Peköz, and N. Joglekar, "Poisson traffic flow in a general feedback queue", *Journal of Applied Probability*, vol.39, no.3, pp. 630-636, 2002.
- [94] L. Pietronero, E. Tosatti, V. Tosatti, and A. Vespignani, "Explaining the uneven distribution of numbers in nature: The laws of Benford and Zipf", *Physica A: Statistical Mechanics and its Applications*, vol. 293, no. 1-2, pp. 297-304, 2001.
- [95] PRE-DRIVE C2X Project, [Online], Available: [http : //www.pre - drive - c2x.eu](http://www.pre-drive-c2x.eu).
- [96] PTV Vision, "The VISSIM traffic simulator", [Online], Available: [http : //www.vissim.de](http://www.vissim.de), 2011
- [97] M. Raissi-Dehkordi, and J. Baras, "Broadcast Scheduling for Time- Constrained Information Delivery," in *Proc. of the 2007 IEEE Global Telecommunications Conference (GLOBECOM07)*, pp. 5298-5303, USA, 2007.
- [98] M. Rudack, M. Meincke, and M. Lott, "On the Dynamics of Ad Hoc Networks for Inter Vehicle Communications (IVC)", *ICWN'02*, 2002.
- [99] H. Saleet, R. Langar, K. Naik, R. Boutaba, A. Nayak, and N. Goel, "Intersection-Based Geographical Routing Protocol for VANETs: A Proposal and Analysis", in *IEEE Trans Veh Technol*, vol. 60, issue 9, pp. 4560-4574, 2011.
- [100] Y. Sasson, D. Cavin, and A. Schiper, "Probabilistic Broadcast for Flooding in Wireless Mobile Ad Hoc Networks", in *Proc. of IEEE Wireless Communications and Networking Conference*, vol. 2, pp. 1124-1130, 2003.

- [101] G. Serugendo, A. Villalba, and D. Konstantas, "Dependable Requirements for Hovering Information", in *Supplementary Proceedings of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, (DSN'07)*, 2007.
- [102] N. Schabanel, "The data broadcast problem with preemption," in *Proc. 17th International Symp. Theoretical Computer Science (STACS)*, 2000.
- [103] S. Schnauffer, and W. Effelsberg, "Position-based unicast routing for city scenarios", in *World of Wireless, Mobile and Multimedia Networks, 2008. WoWMoM 2008. 2008 International Symposium*, pp.1-8, 2008.
- [104] E. Schoch, F. Kargl, and Leinm, "Communication Patterns in VANETs", in *IEEE Communications Magazine*, vol. 46, no. 11, pp. 2-8, 2008.
- [105] SigAlert, [http : //www.sigalert.com](http://www.sigalert.com), 2008.
- [106] "Standard Specification for Telecommunications and Information Exchange Between roadside and Vehicle Systems", *5 GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specifications, ASTM E2213-03*, 2003.
- [107] T. Sun, B. Zan, Y. Zhangand, and M. Gruteser, "The Boomerang Protocol: Tying Data to Geographic Locations in Vehicular Ad Hoc Networks", *In IEEE Transactions on Mobile Computing*, vol. 99, 2011
- [108] T. Taleb, A. Benslimane, and K. B. Letaief, "Towards an Effective Risk-conscious and Collaborative Vehicular Collision Avoidance System", in *IEEE Tran. Vehicular Technology*, vol. 59, No. 3, pp. 1474-1486, 2010.
- [109] M. Taha, "Broadcasting Protocols in Vehicular Ad-Hoc Networks (VANETs)", *A thesis submitted in partial fulfillment of the requirements for the degree Master of Science*, Department of Electrical Engineering, Assiut University, Assiut, Egypt, 2008.
- [110] "Telecommunication Industry Association", [Online], Available: www.tiaonline.org/
- [111] "The Network Simulator 2", [Online], Available: [http : //www.isi.edu/nsnam/ns](http://www.isi.edu/nsnam/ns), I. S. I. University of California, California, USA, 2011.
- [112] O. Tonguz, and M. Boban, "Multiplayer games over Vehicular Ad Hoc Networks: A new application", *Ad Hoc Networks*, pp. 531-543, 2010.
- [113] Traffic safety facts, *Tech. Rep. Report DOT HS 809 767, NHTSA*, [http : //www – nrd.nhtsa.dot.gov](http://www-nrd.nhtsa.dot.gov), 2003.
- [114] Y. Tseng, S. Ni, Y. Chen, and J. Sheu, "The broadcast storm problem in a mobile Ad Hoc network", in *Proc. of the 5th annual ACM/IEEE international conference on Mobile computing and networking In Wireless Networks*, vol. 8, pp. 153-167, 2002.
- [115] U.S. Department of Commerce, "U.S. Frequency Allocation Chart," [Online], Available, [http : //www.ntia.doc.gov/osmhome/allochrt.html](http://www.ntia.doc.gov/osmhome/allochrt.html), 2003.
- [116] A. Vahdat, and D. Becker, "Epidemic routing for partially connected ad hoc networks", *Tech. Rep. CS-200006*, Duke University, 2000.

- [117] N. Vaidya, and S. Hameed, "Scheduling data broadcast in asymmetric communication environments" *Wireless Networks*, vol. 5, no. 3, pp. 171-182, 1999.
- [118] "Vehicle Safety Communications Project (VSC)", *Task 3 Final Report: Identify intelligent vehicle safety applications*, U.S. Department of Transportation, Tech. Rep., Mar. 2005.
- [119] M. D. Venkata, M. M. Pai, R. M. Pai, and J. Mouzna, "Traffic monitoring and routing in VANETs - A cluster based approach", in *11th International Conference on ITS Telecommunications (ITST)*, pp. 27-32, 2011.
- [120] M. Venkataraman, M. Chatterjee, and K. Kwiat, "A Dynamic Reconfigurable Routing Framework for Wireless Sensor Networks", *Elsevier Journal on Ad hoc Networks*, vol. 9, Issue 7, pp. 1270-1286, 2011.
- [121] A. Villalba, and D. Konstantas, "Towards hovering information", on *Proc. of the First European Conference on Smart Sensing and Context (EuroSSC)*, pp. 161-166, 2006.
- [122] A. Villalba, G. Serugendo, and D. Konstantas, "Hovering Information - Self-Organising Information that Finds its own Storage", in *International IEEE Conference on Sensor Networks, Ubiquitous and Trustworthy Computing*, pp. 193-200, 2008.
- [123] P. M. Wightman, and M.I A. Labrador, "Topology Maintenance: Extending the Lifetime of Wireless Sensor Networks", in *IEEE Latin America Transactions*, vol. 8, no. 4, pp. 469-475, 2010.
- [124] H. Wu, R. Fujimoto, R. Guensler, and M. Hunter, "MDDV: A Mobility- Centric Data Dissemination Algorithm for Vehicular Networks", in *1st ACM Workshop on Vehicular Ad Hoc Networks*, pp. 47-56, 2004.
- [125] H. Wu, R. Fujimoto, and G. Riley, "Analytical Models for Information Propagation in Vehicle-to-Vehicle Networks", *IEEE Vehicular Technology Conference*, vol.6, pp. 4548-4552, 2004.
- [126] Y. Wu, E. Pliszka, B. Caron, P. Bouchard, and G. Chouinard, "Comparison of terrestrial DTV transmission systems: the ATSC 8-VSB, the DVB-T COFDM, and the ISDB-T BST-OFDM", *IEEE Transactions on Broadcasting*, vol. 46, no. 2, pp. 101-113, 2000.
- [127] A. Xeros, M. Andreou, M. Lestas, G. Papageorgiou, V. Vassiliou, and A.Pitsillides, "Time Constrained Message Delivery Probability in VANETs", in *Proc. of the 4th Vehicle to Vehicle Communication workshop, V2Vcom*, 2008.
- [128] A. Xeros, M. Andreou, A. Pitsillides, and M. Lestas, "Adaptive Probabilistic Flooding for Information Hovering in VANETs", in *2nd IEEE VNC 2010 Vehicular Networking Conference*, pp. 121-127, 2010.
- [129] A. Xeros, M. Andreou, A. Pitsillides, and M. Lestas, "Information Hovering in Vehicular AdHoc Networks", in *IEEE Workshop on Networking Intelligent Vehicles and Infrastructures (NiVi09)*, pp. 1-6, Honolulu, Hawaii, USA, 2009.
- [130] A. Xeros, M. Andreou, A. Pitsillides, and M. Lestas, "Information Propagation Probability on Intersections in VANETs", in *3rd International Workshop on Vehicle-to-Vehicle Communications*, Istanbul, Turkey, pp. 57-64, 2007.

- [131] J. Xu, D. Lee, Q. Hu, and W. Lee, "Data broadcast", in *New York: John Wiley and Sons Handbook of Wireless Networks and Mobile Computing, Chapter 11*, Ivan Stojmenovic, Ed., pp. 243-265, 2002.
- [132] J. Xu, X. Tang, and W. C. Lee, "Time-critical ondemand data broadcast: algorithms, analysis, and performance evaluation", *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 1, pp. 3-14, 2006.
- [133] B. Zheng, X. Wu, X. Jin, and D. Lee, "TOSA: a near-optimal scheduling algorithm for multi-channel data broadcast", in *Proc. of the 6th International Conference on Mobile Data Management (MDM05)*, pp. 29-37, Cyprus, 2005.
- [134] L. Zhou, Y. Zhang, K. Song, W. Jing, and V. Vasilakos, "Distributed Media Services in P2P-Based Vehicular Networks", *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 692-703, 2011.