Universidade Estadual de Campinas
Instituto de Computação

Ademar Takeo Akabane

Collaborative and Infrastructure-less Vehicular Traffic Rerouting for Intelligent Transportation Systems

Roteamento de Tráfego Veicular Colaborativo e sem Infraestrutura para Sistemas de Transporte Inteligentes

CAMPINAS
2019

Ademar Takeo Akabane

Collaborative and Infrastructure-less Vehicular Traffic Rerouting for Intelligent Transportation Systems

Roteamento de Tráfego Veicular Colaborativo e sem Infraestrutura para Sistemas de Transporte Inteligentes

Tese apresentada ao Instituto de Computação da Universidade Estadual de Campinas como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação.

Dissertation presented to the Institute of Computing of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Computer Science.

Supervisor/Orientador: Prof. Dr. Leandro Aparecido Villas
Co-supervisor/Coorientador: Prof. Dr. Edmundo Roberto Mauro Madeira

Este exemplar corresponde à versão final da Tese defendida por Ademar Takeo Akabane e orientada pelo Prof. Dr. Leandro Aparecido Villas.

CAMPINAS
2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Silvania Renata de Jesus Ribeiro - CRB 8/6592

Ak12c Akabane, Ademar Takeo, 1976-
Collaborative and infrastructure-less vehicular traffic rerouting for intelligent transportation systems / Ademar Takeo Akabane. – Campinas, SP : [s.n.], 2019.

Orientador: Leandro Aparecido Villas.
Coorientador: Edmundo Roberto Mauro Madeira.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Computação.

1. Sistemas inteligentes de veículos rodoviários. 2. Sistemas inteligentes de controle. 3. Sistemas de comunicação móvel. 4. Trânsito - Congestionamento. I. Villas, Leandro Aparecido, 1983-. II. Madeira, Edmundo Roberto Mauro, 1958-. III. Universidade Estadual de Campinas. Instituto de Computação. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Roteamento de tráfego veicular colaborativo e sem infraestrutura para sistemas de transportes inteligentes
Palavras-chave em inglês:
Intelligent transportation systems
Intelligent control systems
Mobile communication systems
Traffic congestion
Área de concentração: Ciência da Computação
Titulação: Doutor em Ciência da Computação
Banca examinadora:
Leandro Aparecido Villas [Orientador]
Antonio Alfredo Ferreira Loureiro
Luís Henrique Maciel Kosmalski Costa
Juliana Freitag Borin
Islene Calciolari Garcia
Data de defesa: 09-12-2019
Programa de Pós-Graduação: Ciência da Computação

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-4930-182X
- Currículo Lattes do autor: http://lattes.cnpq.br/6781874728187325

Universidade Estadual de Campinas
Instituto de Computação

Ademar Takeo Akabane

Collaborative and Infrastructure-less Vehicular Traffic Rerouting for Intelligent Transportation Systems

Roteamento de Tráfego Veicular Colaborativo e sem Infraestrutura para Sistemas de Transporte Inteligentes

Banca Examinadora:

- Prof. Dr. Leandro Aparecido Villas
  Instituto de Computação - UNICAMP
- Prof. Dr. Antonio Alfredo Ferreira Loureiro
  Departamento de Ciência da Computação - UFMG
- Prof. Dr. Luís Henrique Maciel Kosmalski Costa
  Grupo de Teleinformática e Automação - UFRJ
- Profa. Dra. Juliana Freitag Borin
  Instituto de Computação - UNICAMP
- Profa. Dra. Islene Calciolari Garcia
  Instituto de Computação - UNICAMP

A ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Campinas, 09 de dezembro de 2019

Acknowledgements

First of all, I thank my parents for educating me in the best way possible. Without a strong base, I certainly would not be able to get where I got. I also would like to thank my loving wife, Elza Akabane. Her support has been invaluable in these past four years. I could not have finished this thesis if it weren't for her encouragement during all the long and difficult nights of writing and coding.

My sincere thanks to my advisors: Prof. Leandro Aparecido Villas, Prof. Edmundo Roberto Mauro Madeira, and Prof. Richard Pazzi for their encouragement, guidance, and support throughout my Ph.D. period. Both of you gave me valuable pieces of advice not only about research and the process of accomplishing it but also about life.

The staff and faculty members of the Institute of Computing at the University of Campinas and the Ontario Tech University, undoubtedly, were cooperative and supported me to go throughout this process. I would like to thank everyone there for their support.

I thank my friends and colleagues of the Computer Networks Laboratory (Brazil) and Advanced Networking Technology and Security Research Lab (Canada) for their help whenever I needed and willingness to make the Lab an enjoyable and fun work environment.

Lastly, I would like to thank FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo (Process No. 2015/25588-6) for the essential financial support to carry out this research.

Resumo

Devido à atual tendência mundial de urbanização, a sociedade moderna enfrenta, cada vez mais, sérios problemas de mobilidade urbana. Além disso, com o aumento constante do fluxo de tráfego veicular, as atuais soluções existentes para gerenciamento de tráfego se tornaram ineficientes. Com isso, para atender às crescentes necessidades dos sistemas de transporte, é necessário sistemas de transporte inteligentes (ITS). O desenvolvimento de ITS sustentável requer integração e interoperabilidade contínuas com tecnologias emergentes, tais como as redes veiculares (VANETs). As VANETs são consideradas uma tecnologia promissora que provê aplicações críticas de segurança e serviços de entretenimento, consequentemente melhorando a experiência de viagem do motorista e dos passageiros.

Esta tese propõe um sistema de gerenciamento de tráfego de veículos sem a necessidade de uma infraestrutura de apoio. Para alcançar o sistema desejado foi necessário propor soluções intermediárias que contribuíram nesta tese. A primeira contribuição reside em uma solução que emprega conhecimento histórico dos padrões de mobilidade dos motoristas para obter uma visão global da situação da rede viária. Diferentemente de outras abordagens que precisam de troca constante de informações entre os veículos e o servidor central, nossa solução utiliza informações espaciais e temporais sobre padrões de mobilidade, além das informações específicas da infraestrutura viária, a fim de identificar congestionamentos no tráfego, permitindo, assim, o planejamento de roteamento de veículos. Como segunda contribuição, foi proposta uma solução distribuída para calcular a intermediação egocêntrica nas VANETs. Por meio da métrica egocêntrica foi proposto um mecanismo inovador de ranqueamento de veículos em redes altamente dinâmicas. As principais vantagens desse mecanismo para aplicações de VANETs são: *(i)* a redução do consumo de largura de banda e *(ii)* a superação do problema de topologias altamente dinâmicas. A terceira contribuição é uma solução de planejamento colaborativo das rotas com intuito de melhorar o gerenciamento do tráfego de veículos em cenários urbanos. Como última contribuição, esta tese integra as soluções descritas acima, propondo um sistema eficiente de gerenciamento de tráfego de veículos.

As soluções propostas foram amplamente comparadas com outras soluções da literatura em diferentes métricas de avaliação de desempenho. Os resultados mostram que o sistema de gerenciamento de tráfego de veículos proposto é eficiente e escalável, no qual pode ser uma boa alternativa para mitigar os problemas de mobilidade urbana.

Abstract

Due to the current global trend of urbanization, modern society is facing severe urban mobility problems. In addition, considering the constant increase in vehicular traffic on roads, existing traffic management solutions have become inefficient. In order to assist the increasing needs of transport systems today, there is a need for intelligent transportation systems (ITS). Developing a sustainable ITS requires seamless integration and interoperability with emerging technologies such as vehicular ad-hoc networks (VANETs). VANETs are considered to be a promising technology providing access to critical life-safety applications and infotainment services, consequently improving drivers' and passengers' on-road experiences.

This thesis proposes an infrastructure-less vehicular traffic management system. To achieve such a system, intermediate solutions that contributed to this thesis were proposed. The first contribution lies in a solution that employs historical knowledge of driver mobility patterns to gain an overall view of the road network situation. Unlike other approaches that need constant information exchange between vehicles and the central server, our solution uses space and temporal information about mobility patterns, as well as road infrastructure information, in order to identify traffic congestion, thus allowing for vehicle routing planning. Secondly, a distributed solution to calculate egocentric betweenness in VANETs was proposed. Through the egocentric metric, an innovative vehicle ranking mechanism in highly dynamic networks was proposed. The main advantages of this mechanism for VANETs applications are *(i)* reduced bandwidth consumption and *(ii)* overcoming the problem of highly dynamic topologies. The third contribution is a collaborative route planning solution designed to improve vehicle traffic management in urban settings. As the last contribution, this thesis integrates the solutions described above, proposing an efficient vehicle traffic management system.

The proposed solutions were widely compared with other literature solutions on different performance evaluation metrics. The evaluation results show that the proposed vehicle traffic management system is efficient, scalable, and cost-effective, which may be a good alternative to mitigate urban mobility problems.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

This chapter presents the motivation, objectives, and main contributions of the thesis, as well as the thesis outline.

1.1 Motivation

Urbanization is a worldwide phenomenon describing a movement of the countryside's population into urban areas. According to the United Nations' report [132], for the first time in human history, in 2007, more than half of the world's population was living in urban areas, as can be seen in Figure 1.1. Besides, the report also forecasts that two-thirds of the world's population will be living in urban areas by the year 2050.

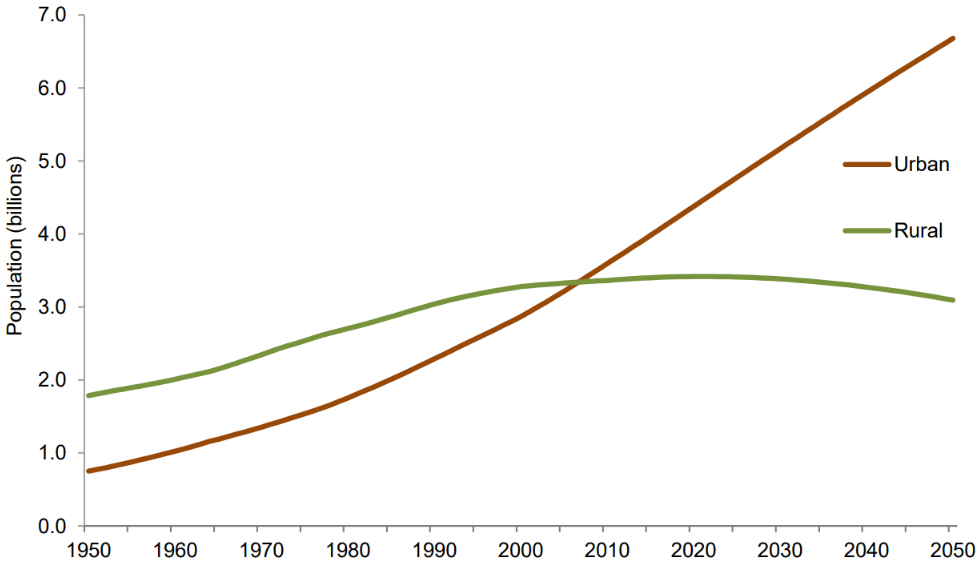


Figure 1.1: Urban and rural populations of the world [132].

Figure 1.2 shows the microscopic view of the world's urbanization. According to a study carried out by the United Nations, only fourteen countries still have low levels of urbanization, i.e., less than 20% of their population living in urban areas [132].

Rapid urbanization has greatly accelerated the economic and social development of citizens. On the other hand, it has also created serious challenges in urban administration for public

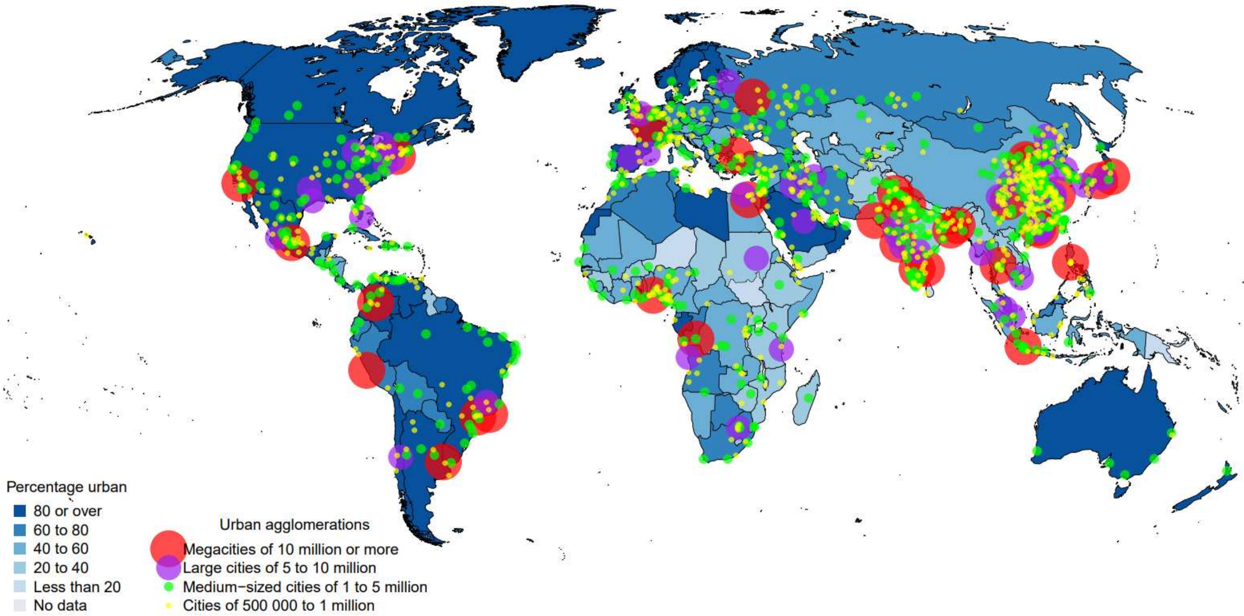


Figure 1.2: Urban agglomerations and urban percentage with over 500,000 inhabitants in 2018 [132].

authorities, for example, those related to vehicle traffic management.

A global leader of connected car services and mobility analytics, INRIX[1], published the 2018 Global Traffic Scorecard that identified and ranked vehicular traffic congestion and mobility trends in more than 200 cities, across 38 countries [74]. Figure 1.3 presents the top 25 most congested cities in the world. The figure shows that Moscow, Istanbul, Bogota, Mexico City, and Sao Paulo represent the top 5 in the Global Congestion Impact ranking[2].

According to research firm INRIX, in 2018, Americans lost an average of 97 hours a year due to congestion, costing them nearly \$87 billion, reaching an average of \$1,348 per driver [74]. For a microscopic view of this scenario, Figure 1.4 presents the top 25 most congested cities in the U.S. As can be seen, Boston and Washington D.C. are the top two most congested cities, and drivers in each city waste up to 164 and 155 hours in congestion, respectively. Drivers from these two cities spend more than 15 hours per year compared to the next two worst cities in terms of total hours: Chicago (138 hours) and Seattle (138 hours). This lost time has an annual cost of \$2,291, \$2,161, \$1,920, and \$1,859, for drivers in Boston, Washington D.C., Chicago, and Seattle, respectively.

A straightforward way to alleviate vehicular traffic congestion is to decrease the absolute number of vehicles in circulation. To this end, a public policy well known as *end-number license plate policy* was elaborated and implemented in most large cities. In this policy, all registered vehicles are classified into five groups according to the last digit of the plate number. Thus, each group of vehicles is prohibited from being driven in a particular region of the city during the rush-hour of a certain business day. The work of Li and Guo [87] has demonstrated a reduction in almost 40% of the daily emission and nearly 20% of traffic volume on public roads after implementing the traffic restriction policy in Beijing city. The other typical policy is *road*

[1]http://inrix.com/

[2]Impact rank is a calculated commute based upon a city's population and the delay attributable to congestion [74]

2018 IMPACT RANK (2017)	URBAN AREA	COUNTRY	REGION	HOURS LOST IN CONGESTION (RANK 2018)	YEAR OVER YEAR CHANGE	INNER CITY LAST MILE TRAVEL TIME (MINUTES)	INNER CITY LAST MILE SPEED (MPH)
1 (1)	Moscow	Russia	Europe	210 (10)	-12%	5	11
2 (3)	Istanbul	Turkey	Europe	157 (32)	6%	6	10
3 (2)	Bogota	Colombia	South America	272 (1)	-5%	8	7
4 (4)	Mexico City	Mexico	South America	218 (9)	3%	7	9
5 (5)	São Paulo	Brazil	South America	154 (39)	-1%	6	10
6 (6)	London	United Kingdom	Europe	227 (6)	1%	8	7
7 (8)	Rio de Janeiro	Brazil	South America	199 (13)	15%	5	13
8 (7)	Boston, MA	United States	North America	164 (25)	-10%	6	11
9 (9)	Saint Petersburg	Russia	Europe	200 (12)	-5%	6	11
10 (13)	Rome	Italy	Europe	254 (2)	16%	8	8
11 (10)	Ankara	Turkey	Europe	128 (75)	-5%	5	12
12 (11)	Izmir	Turkey	Europe	154 (38)	1%	6	10
13 (12)	Sydney	Australia	Oceania	138 (63)	-1%	6	10
14 (14)	Singapore	Singapore	Oceania	105 (106)	-2%	4	15
15 (16)	Berlin	Germany	Europe	154 (40)	-5%	5	11
16 (18)	Paris	France	Europe	237 (5)	7%	7	8
17 (15)	Melbourne	Australia	Oceania	118 (87)	-13%	6	11
18 (22)	Belo Horizonte	Brazil	South America	202 (11)	12%	8	8
19 (20)	Washington D.C.	United States	North America	155 (36)	-3%	5	11
20 (19)	Toronto, ON	Canada	North America	164 (26)	-4%	6	10
21 (23)	Guayaquil	Ecuador	South America	167 (24)	2%	5	12
22 (24)	Madrid	Spain	Europe	129 (74)	3%	7	8
23 (25)	Chicago, IL	United States	North America	138 (64)	4%	5	12
24 (26)	Brisbane	Australia	Oceania	157 (33)	3%	6	11
25 (26)	Medellin	Colombia	South America	138 (62)	12%	6	10

* The Impact Ranking includes a weighting based on city population

Figure 1.3: Top 25 most congested cities in the world [74].

pricing, which requires each driver to contribute to the costs of roads according to their level of use. Rather than prohibiting, as the plate policy does, *road pricing* charges drivers who are driving into the congested area during the business days [104].

Another way to minimize vehicular traffic congestion is by using information and communication technologies (ICT) through advanced traffic management systems (ATMS). SCOOT [72] and SCATS [127] were two of the first systems that employed ICT for traffic management. The SCOOT and SCATS systems need a traffic operation center (TOC) that manages all traffic lights and optimizes the traffic light timings. Such systems, basically, during fixed time intervals, collect the real-time traffic information by induction loops that are installed underground of major urban roads. Using such information, the system can identify the vehicle flow on the induction loop area, thus adjusting the time cycle of traffic lights. To optimize the traffic light timings and control vehicles queuing in front of junctions, both systems need a TOC that manages all traffic lights. Another ICT service commonly used for congestion control is the vehicle navigation system. This type of system collects traffic information through the user's mobile devices, and

2018 IMPACT RANK (2017)	URBAN AREA	HOURS LOST IN CONGESTION (RANK 2018)	YEAR OVER YEAR CHANGE	INNER CITY LAST-MILE TRAVEL TIME (MINUTES)	INNER CITY LAST-MILE SPEED (MPH)	COST OF CONGESTION (PER DRIVER)	COST OF CONGESTION (PER CITY)
1 (1)	Boston, MA	164 (1)	-10%	6	11	$2,291	$ 4.1B
2 (2)	Washington, DC	155 (2)	-3%	5	11	$2,161	$ 4.6B
3 (5)	Chicago, IL	138 (4)	4%	5	12	$1,920	$ 6.2B
4 (3)	New York City, NY	133 (5)	-4%	7	9	$1,859	$ 9.5B
5 (4)	Los Angeles, CA	128 (6)	0%	4	14	$1,788	$ 9.3B
6 (6)	Seattle, WA	138 (3)	0%	6	10	$1,932	$ 2.9B
7 (11)	Pittsburgh, PA	127 (7)	5%	5	13	$1,776	$ 1.2B
8 (7)	San Francisco, CA	116 (9)	-5%	6	10	$1,624	$ 3.4B
9 (10)	Philadelphia, PA	112 (10)	0%	6	10	$1,568	$ 3.3B
10 (8)	Portland, OR	116 (8)	-9%	5	13	$1,625	$ 1.4B
11 (13)	Atlanta, GA	108 (11)	10%	4	14	$1,505	$ 3.5B
12 (9)	Miami, FL	105 (12)	-5%	5	12	$1,470	$ 4.0B
13 (14)	Houston, TX	98 (14)	6%	4	15	$1,365	$ 3.8B
14 (12)	Austin, TX	104 (13)	-2%	5	13	$1,452	$ 1.2B
15 (16)	Baltimore, MD	94 (16)	3%	6	10	$1,315	$ 1.3B
16 (15)	Charlotte, NC	95 (15)	0%	5	12	$1,332	$ 953.8M
17 (19)	Tampa, FL	87 (19)	11%	5	13	$1,216	$ 1.5B
18 (17)	Honolulu, HI	92 (17)	-4%	5	12	$1,282	$ 432.0M
19 (18)	Denver, CO	83 (20)	-3%	5	13	$1,152	$ 1.5B
20 (23)	Nashville, TN	87 (18)	20%	4	16	$1,221	$ 694.7M
21 (20)	Dallas, TX	76 (22)	6%	4	17	$1,065	$ 3.1B
22 (21)	Phoenix, AZ	73 (25)	3%	4	17	$1,013	$ 1.8B
23 (31)	Orlando, FL	74 (23)	16%	4	15	$1,037	$ 900.1
24 (24)	Minneapolis, MN	70 (28)	4%	4	14	$971	$ 1.3B
25 (26)	Columbus, OH	71 (27)	6%	4	14	$990	$ 734.9M

Figure 1.4: Top 25 most congested cities in the U.S. [74].

the best known are Google Maps[3], TomTom[4], and Waze[5]. Thus, the users of these systems can monitor the current traffic conditions easily to plan their travel routes.

The current vehicles are equipped increasingly with a variety of computational resources, for example, sensors, cameras, and wireless communication devices to facilitate the utmost travel comfort and safety of drivers and passengers. Through the advancement of wireless communication technology, a new paradigm of wireless networks, known as vehicular ad-hoc networks (VANETs) [10, 40, 65, 68], is emerging. Thus, VANETs can collect, process and share sensed data supporting various intelligent transportation systems (ITS) applications such as ATMS and urban environment sensing. Thus, we firmly believe that VANETs can help deal with urban mobility problems. Due to the current global trend of urbanization, modern society is facing serious urban mobility problems, higher fuel prices, and an increase in CO_2 emissions. In addition to that, with the constant increase in vehicular traffic on roads, existing traffic management solutions have become inefficient. In order to serve the increasing needs of transport systems, there is a need for ITS. Developing a sustainable ITS requires seamless integration and interoperability with emerging technologies such as VANETs.

[3]https://maps.google.com
[4]https://www.tomtom.com
[5]https://www.waze.com

1.2 Objective

The main objective of this thesis is to design, implement, and evaluate a collaborative and infrastructure-less system for vehicular traffic management. To achieve this goal, we need to answer the following questions:

- Many VANET-based traffic management systems [21, 48, 100, 109, 137] were proposed to generate a global view of the road network, allowing the detection of all possible road traffic congestions. These systems need constant information exchange between the vehicle and the central server in order to obtain a global view of road traffic conditions. In this regard, it is known that if this information exchange is not well managed, it can lead to network overload.

  Research Question 1: *How can we obtain a global view of road network topology without exchanging data between vehicles and the central server for traffic management purposes?*

- High mobility of nodes is the main characteristic of VANETs. Therefore, identifying and selecting the best-located vehicles available at the right time and place for a given application task through inter-vehicle communications is a very challenging task. The best-located vehicle is defined as the importance of the car concerning the information flows that passes through it. On the other hand, once it is identified, it can be beneficial for a large number of services, such as those that spread the information flow through the network.

  Research Question 2: *How can we dynamically identify the best-located vehicle among the candidate ones, in a distributed manner, to perform a given application task?*

- It is known that the primary goal of the vehicle rerouting algorithm is to move vehicular traffic away from the congestion point. To this end, two main requirements for this type of algorithm in VANETs are expected: *(i)* to calculate alternative routes for each vehicle that can improve the vehicle's path and also maximize the global network efficacy; and *(ii)* to alert vehicles quickly so that they have enough time to compute a new route. To do this, collaborative route planning was proposed to answer the question below. It is worth mentioning that this type of planning takes into account the surrounding vehicles' routes to compute an alternative route.

  Research Question 3: *Can collaborative route planning help effectively minimize traffic congestions without compromising scalability?*

- Several systems have been proposed to deal with issues related to vehicular traffic management. Usually, their solutions include the integration of computational technologies such as vehicular networks, central servers, and roadside units. Most of them apply a hybrid approach, which means they still need a central entity (central server or roadside unit) and Internet connection to achieve their objectives. It is known that integrating different types of technologies increases the cost of developing systems and often making the implementation unfeasible.

Research Question 4: *Can infrastructure-less vehicular traffic management systems be as efficient as infrastructure approaches and also scalable and cost-effective?*

1.3 Main Contributions

The main contributions of this thesis are a people-centric approach for vehicular traffic management, a comparative study on the egocentric and sociocentric betweenness measure in VANETs, a distributed system for information management and knowledge distribution, and collaborative and infrastructure-less vehicular traffic management. In summary, we have:

1. **Routing Protocol using Mobility Pattern**

   This contribution concerns the proposal of a vehicle routing protocol. It is currently known that there are anonymized datasets concerning the mobility patterns of the drivers[6,7,8] in the urban center. Based on that, we proposed a protocol that uses historical knowledge of mobility patterns of the drivers to obtain a global view of the road network situations. Our approach has two different stages: *(i) - Offline*, in which the historical data processing of the global view of the road network is performed in order to generate the mobility patterns; and *(ii) - Online*, in which vehicles in the route to congested roads are re-routed. The proposed protocol acts as a traffic monitoring system, having an overview of road networks without needing to periodically exchange information status between the central server and the vehicles. Simulation results (presented in Section 3.4) have shown that such an approach could be a suitable alternative for traffic management. This protocol is fully explained in Chapter 3.

2. **Vehicle Ranking Mechanism**

   This contribution concerns the proposal of an innovative vehicle ranking mechanism called V_{rank}. It is known that Google's PageRank [108] algorithm ranks the importance of webpages based on the number of web-links directed towards it. The general idea of PageRank relies on a graph where nodes are webpages and edges depict the links between them. Thereby, PageRank uses the link structure as an indicator of an individual page's importance in the structure of the World Wide Web relative to other pages. In general, the higher the number of links, the greater the importance of the webpage. The idea of V_{rank} is to use the link structure of VANETs to compute the vehicle's score. To do this, we used the Egocentric Betweenness Metric [4, 7]. Betweenness is a measure of how often a node is located on the geodesic distance (shortest path) between other nodes in the network. It thus measures the importance to which the node can function as a point of control in the communication [107]. Intuitively, the betweenness metric measures the control a node has over communication in the network. High betweenness value, thus implying that a node can reach other nodes on relatively short path or that a node lies on a considerable fraction of shortest paths connecting pairs of other nodes. Simulation results (presented in

[6] http://kolntrace.project.citi-lab.fr/
[7] http://www.vehicularlab.uni.lu/lust-scenario/
[8] https://crawdad.org/epfl/mobility/20090224/

Section 5.4) have demonstrated that by using the V_{rank}, it is possible to make the system scalable. This mechanism is fully explained in Chapter 5.

3. **Collaborative Route Planning**

   This contribution consists of proposing, designing, and evaluating collaborative route planning to improve vehicular traffic management on urban road scenarios. Generally speaking, vehicles traveling in the congestion region collaborate by exchanging information about their alternate routes chosen that bypass the congestion. The idea here is that each vehicle plans its available alternative routes to the destination taking into account the alternative route information received from surrounding vehicles [8]. Through this collaboration, each vehicle can create an awareness to which roads vehicles are being moved to, thus planning the most suitable alternative route and avoiding potential future congestion. Simulation results (presented in Section 6.4.4) have shown that collaborative decision making is more efficient than selfish decision making in alternative routes planning. This mechanism is fully explained in Section 6.3.4.

4. **Infrastructure-less Vehicular Traffic Management System**

   This contribution consists of proposing, implementing, and evaluating a collaborative and infrastructure-less vehicular traffic management system in the urban scenario. It is worth mentioning that such a system takes into account the contributions presented in Items 2 and 3 previously presented to achieve its goal. Simulation results (presented in Section 6.4) have demonstrated that the proposed solution tends to be more scalable than infrastructure ones, and the collaborative routing strategy is more suitable in urban mobility management. This system is fully explained in Chapter 6.

1.4 Thesis Outline

The structure of this thesis is outlined in chapters as follows:

- **Chapter 2** presents an overview of the current taxonomy of ITS applications and also the concept and challenges of using VANETs;

- **Chapter 3** proposes and assesses a vehicular traffic routing that employs historical knowledge of mobility patterns of the drivers to obtain a global view of the road network, called APOLO. Such an approach does not require constantly exchanging information among the vehicles and the central server in order to obtain a global view of road traffic conditions;

- **Chapter 4** depicts a thorough study by implementing and evaluating how well egocentric betweenness performs compared to the sociocentric measure in VANETs. The main advantage of egocentric measures is to use only locally available knowledge of the topology to evaluate the importance of a node. In this study, using the egocentric betweenness measure in highly dynamic topologies has demonstrated a high degree of similarity compared to the sociocentric approach;

- **Chapter 5** proposes a system for information management and knowledge distribution called TRUSTed. The proposed system applies the egocentric betweenness measure, introduced in the previous chapter, to select the most relevant vehicle to carry out the tasks of information aggregation and knowledge generation;

- **Chapter 6** proposes and assesses a distributed system of urban mobility management based on a collaborative approach in vehicular social networks (VSNs), called SOPHIA. The VSN paradigm has emerged from integrating mobile communication devices and their social relationships in the vehicular environment. Therefore, social network analysis and social network concepts are two approaches explored in VSNs. Our proposed solution adopts both social network analysis and social network concept approaches for alternative route-planning;

- **Chapter 7** concludes this thesis with a summary, directions for future work to ensure continuous improvement in the current and related field of study and the publications produced from this thesis.

Chapter 2

Intelligent Transportation Systems

In the 20th century, the ITS concept was proposed by the United States (US); however, it has become a topic of research and development worldwide, particularly in the European Union (EU), Japan, and the US [17]. Although ITS may refer to all forms of transport, the EU has limited its applications in the field of road transport [53]. This chapter introduces an overview of ITS, as well as the background of VANETs on which this thesis is based.

2.1 Introduction

The increasing need for mobility in large urban centers has brought about important changes in transportation infrastructures. Moreover, it is well-known that such urban centers are increasingly overcrowded with vehicles, and the direct consequence of this is the population facing unpleasant situations in daily life such as growing vehicular traffic congestions, as well as unpredicted emergencies and accidents. The lack of mobility in urban areas has shown the need to develop more efficient and safer transportation systems. To this end, traffic management systems have applied information and communication technologies, emerging the so-called intelligent transportation systems [12, 65, 94]. ITS consist of different telecommunications and computer technologies designed and developed to improve the management, monitoring, control, and safety of vehicular traffic.

ITS usually consist of multi-subsystems that combine tasks of data gathering, storage, processing, and management tasks (Figure 2.1). Thus, real-time data sensing may be processed to compute the communication network state, to plan a route, to dynamically manage traffic flows in a particular area, and to report data from a logistics operator [45, 146]. In addition to that, such subsystems need to work synchronously to meet the global objective of the whole system [146]. In other words, ITS are made up of subsystems where each one has a well-defined task to provide useful information to the end-user. Summing up, when ITS solutions are designed, all the synergies among subsystems and the interests of all the stakeholders, such as end-users, companies, and governments, must be specified. Thus, the system provides a common goal, designed based on the user requirements and the scope for the planning of a smart transport system.

Public and private institutions play a vital role in promoting policies that help and support the development of systems that improve the efficiency of current ITS. A typical example of this

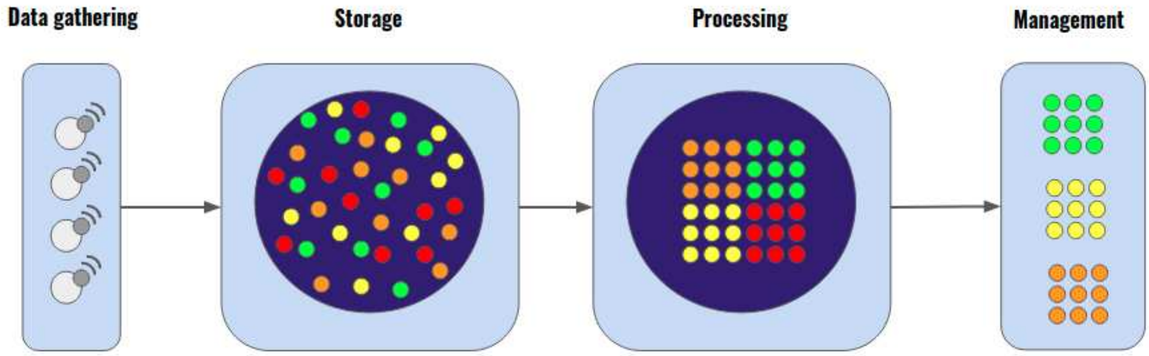


Figure 2.1: The key tasks of intelligent transportation systems.

is Horizon 2020's project[1] from the EU. This project includes a work-oriented towards *Smart, green and integrated transport*[2], which encourages projects and ideas related to 'Mobility for Growth' or 'Green Vehicles'.

In recent years, a large number of innovations, projects, and research have focused on issues involving intelligent transport systems that will be detailed in Section 2.2.

2.2 Taxonomy of ITS Applications

ITS applications are often classified into five main categories according to their functionalities: Environment, Assistance, Safety, User, and Traffic Management. Figure 2.2 provides a taxonomy of this classification. The following is a brief explanation of each of them.

Environment category - The environment category focuses on providing detailed information about the road environment situations, for example, weather prediction systems are based on surveillance, monitoring, weather forecasts, and roadway conditions to perform the proper management actions in order to improve the driving experience and alleviate the impacts of unfavorable conditions. Road weather systems can be used to help make decisions concerning strategies, route planning, and driver advisories. This type of system generally uses physical-sensing devices (weather stations, such as humidity sensors, and temperature sensors) usually deployed on roads to determine precipitation, air temperatures, smoke, fog, as well as other external factors which directly increase the risk situations for vehicle occupants or affect road maintenance decisions.

Assistance category - Assistance category aims to provide information, advice, and warnings that assist or intervene in vehicle control, besides avoiding dangerous driving situations. For example, parking spot locator systems indicate available parking places such as public roads, garages, or parking lots. In this type of system, the radio-frequency identification technology and GPS are commonly used to collect information from different parking spots, thus offering drivers ample opportunities to park their vehicles. Tourism and event systems are developed to attend the needs of travelers in unknown regions, indicating hotels, restaurants, concerts, or

[1] https://ec.europa.eu/programmes/horizon2020/en/h2020-section/smart-green-and-integrated-transport
[2] https://trimis.ec.europa.eu/programme/horizon-2020-smart-green-and-integrated-transport

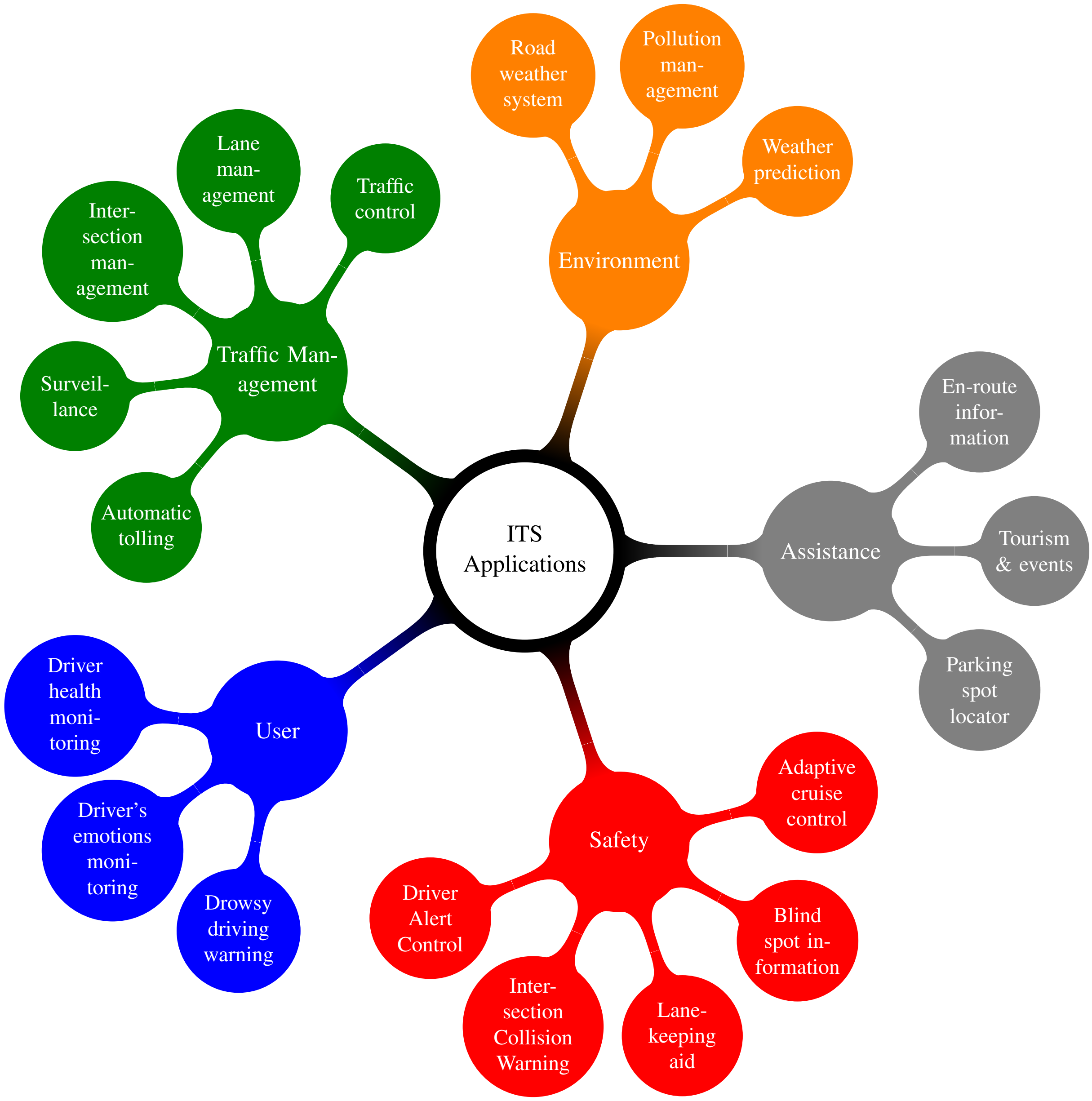


Figure 2.2: Taxonomy of ITS applications.

sports events according to traveler preference.

Safety category - This category focuses on improving the safety of drivers and passengers by reducing the number of accidents, injuries, and fatalities during the journey. A well-known application from the security category is the lane-keeping system which helps keep a vehicle within its lane. This system monitors road lane markings and recognizes any drifts outside of this lane using onboard vehicular cameras [43]. Another example of this category is the adaptive cruise control system which uses distance, speed, and radar sensors to manage the speed and keep a secure distance away from the vehicles in front [126]. Blindspot information is a system to alert the driver when a vehicle is detected to be approaching or entering the blind spot area [27, 81]. Intersection collision warning systems use speed and position information of vehicles to compute the likelihood of a collision. Every time the probability of collision is higher than some established security range, a warning signal is transmitted.

User category - The user category focuses on monitoring the drivers' behavior such as fatigue, alcohol levels, and emotional state disorders, which is essential for traffic safety and reducing accidents. Drowsy driving warning systems aim to prevent accidents by analyzing facial expressions, including eye closure duration, eyelid movement, and eye blink times [136, 142]. In addition to that, radar sensors are employed to monitor the car's movements and detect any abnormality. Driver's health monitoring systems are increasingly using low-cost, non-contact technologies to measure physiological information [26, 31]. Usually, drivers' physiological monitoring parameters are captured by on-board camera images [26]. One of the advantages of using camera images is that there is no electrical contact between the person and the equipment. Therefore, when the system identifies that there is something wrong with the driver's health, an emergency vehicle can be called automatically. A driver's emotion recognition system focuses on identifying signs of irritation or fatigue that impair driving performance. Such systems use electromyogram, respiration, and electrodermal activity signals combined with sophisticated algorithms such as support vector machines and adaptive neuro-fuzzy interference systems to classify and recognize these emotions [13].

Traffic Management category - The traffic management category has aimed to improve vehicular traffic flow efficiency. Surveillance systems can be classified into two categories: the first one, fixed surveillance systems that use cameras and sensors placed on the roads to monitor traffic conditions. The second one, vehicular onboard-surveillance systems use cameras and sensors embedded into support surveillance [33, 97]. Traditional traffic lights are increasingly being replaced by intersection management systems for intersection control. In these systems, vehicular and road infrastructure technologies and traffic control centers operate in an integrated fashion to coordinate traffic efficiently [32]. Lane management systems aim to manage the available road capacity in special circumstances such as incidents, high-risk weather, or emergency evacuations. This system utilizes cameras and different kinds of sensors (for example, infrared and radar) to identify occupancy, velocity, and the direction of vehicles [58]. Traffic management systems are becoming increasingly necessary in large urban centers, and the vehicular ad-hoc network is a promising paradigm to help such systems [21, 48, 100, 109, 137].

2.3 Vehicular Ad-hoc Networks

In recent decades, sensors have become increasingly ubiquitous in our daily environment due to their low production cost. Furthermore, we can observe sensors deployed in many areas such as agriculture [19, 105], forestry [46, 99], healthcare [11, 117], and vehicle [88, 98] monitoring. In the vehicular scenario, vehicle manufacturers are increasingly deploying sensors aiming to provide services to end-users and also increase their satisfaction levels. Figure 2.3 depicts an illustrative example of a set of sensors commonly found on current vehicles. Nowadays, the estimated number of sensors in a modern vehicle is nearby 100, and as vehicles become "smarter", this number might rise up to 200 sensors per vehicle [65].

In addition to advances in sensor technology, there have also been advances in information technology and communication. It is known that such advances allowed the emergence of a new network paradigm well-known as VANETs. VANETs are a particular case of Mobile Ad-hoc Networks (MANETs), whose nodes are made up of vehicles, and the orientation of public road

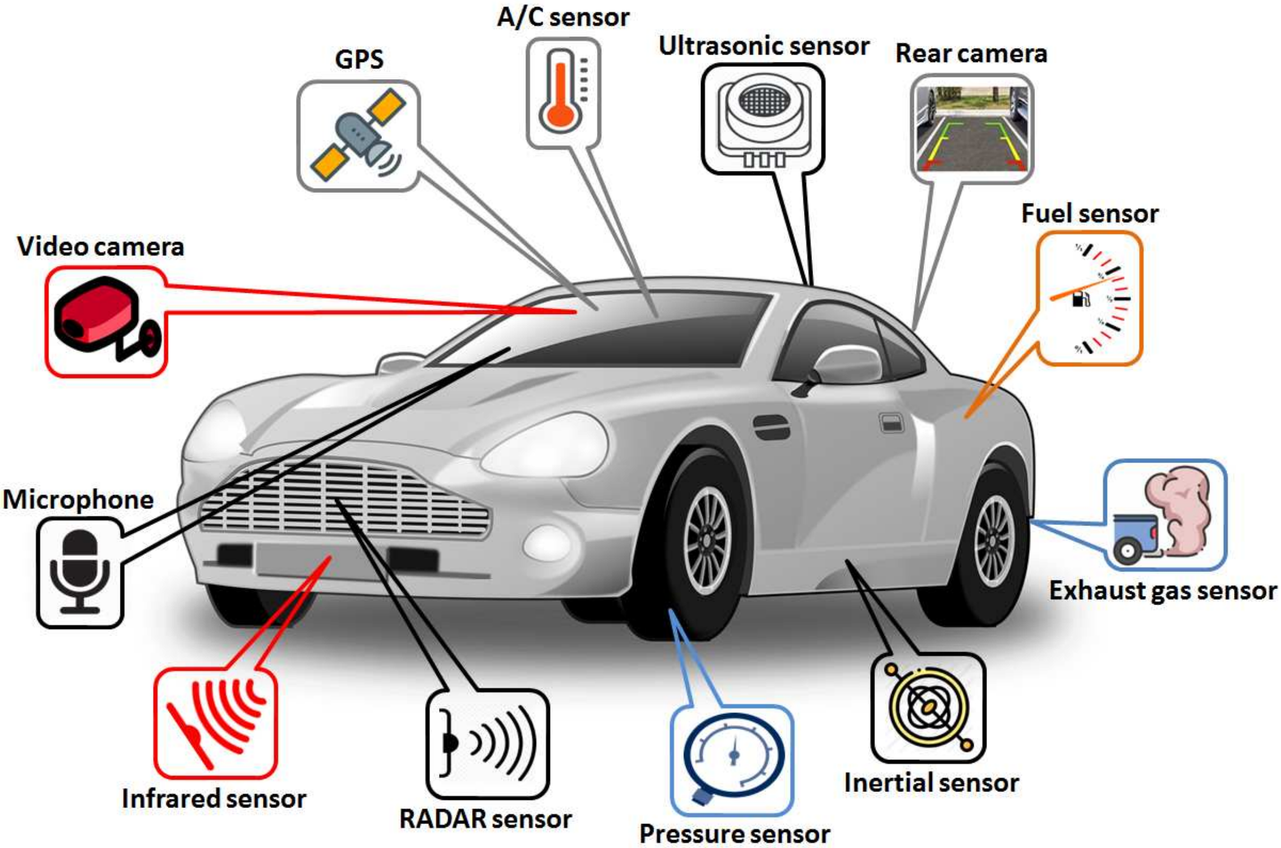


Figure 2.3: Different types of in-car sensors.

limit node movements. In this kind of network, vehicles have wireless communication, processing, and storage capabilities through onboard units, thus enabling to build communication networks on the go spontaneously. In the VANET, each vehicle works as a host and also as a router by forwarding packets to other vehicles inside the transmission range [93].

VANET support is of vital importance for near-future ITS applications [38, 52]. It is currently known that most vehicle manufacturers are supplying vehicles with onboard computational resources, wireless communication devices, and in-car sensors, in order to deploy large-scale vehicular networks. By using different sensors together (RADAR, infrared, and ultrasonic), cameras, computational resources, and wireless communication, vehicles can gather and process the data and return useful information or recommendations to help the driver to make a decision [88, 98]. In the remaining sections of this chapter, we discuss some aspects of VANETs that are necessary to understand the contributions made in this thesis.

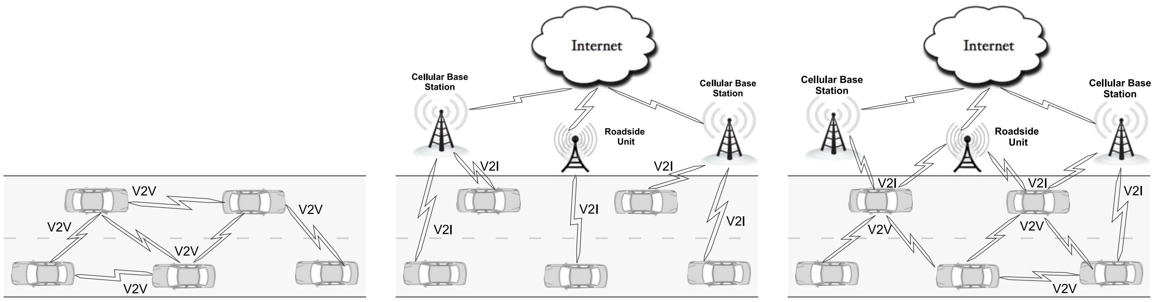
2.3.1 VANET Characteristics

Figure 2.4 portrays a classification of VANET communications. It is well-known that communication can take place between nearby vehicles and between vehicles and roadside units (RSUs), thus leading to the three communication possibilities, as explained below:

- *Vehicle-to-vehicle communication* - (V2V): The V2V provides direct communication between vehicles without relying on the support of static infrastructure. In this case, the vehicles themselves are responsible for data dissemination on the network by other vehicles through multiple hops. It is noteworthy that the V2V communication link directly depends on the density;

- *Vehicle-to-infrastructure communication* - (V2I): The V2I enables vehicles to establish a communication link with a communication infrastructure, known as RSUs. RSUs can serve as intermediate communication nodes or gateways, besides centralizing network traffic. The advantage of the V2I is to increase connectivity and the ability to communicate with other networks, such as the Internet. However, this benefit is only achievable by installing numerous RSUs at the roadside and/or highways, increasing the cost of implementation;

- *Hybrid communication*: It combines the benefits of V2V and V2I communications. In this case, the infrastructure is utilized to increase network connectivity, i.e., a vehicle can communicate with a fixed infrastructure in a single hop or multiple hops with other vehicles according to the node's location on the network.



(a) Vehicle-to-Vehicle (b) Vehicle-to-Infrastructure (c) Hybrid

Figure 2.4: VANET communications - Adapted from [40].

A VANET has some unique characteristics, listed below, which make it different from MANET and also has some similar characteristics, such as omnidirectional broadcast, short transmission range, and low bandwidth:

- *Highly dynamic topology*: The topology of vehicular networks keeps on changing because of the high-speed movement of the vehicle. Vehicles usually travel at a relative velocity in the order of 50 km/h in urban scenarios and over 100 km/h on highways [38]. In addition to that, vehicles can quickly join or leave networks, in very short time periods, because they may move in different directions;

- *Frequent disconnections*: The high dynamic mobility of vehicles leads to the reduction of communication link stability. Consequently, the communication link between vehicles can quickly disappear during data transmission;

- *No power limitation*: Unlike MANET nodes, nodes in VANET have minimal energy dependency. They have a reliable power supply (vehicle battery) and this allows the vehicle to have high computational power;

- *Constrained mobility*: It is known that VANETs display highly dynamic topology; however, vehicles are constrained by roads, streets, and highways layouts, traffic laws and regulations, and drivers' driving behaviors. Given the mobility restrictions, it is possible to predict the future position of the vehicle [38];

- *Variable network densities*: Network density can range from sparse, with few or no vehicles within the transmission range, to dense, with many vehicles, as vehicles move. For example, the density can be small as in rural areas or large as during rush hour in urban centers;

- *Variable signal propagation models*: VANET applications are usually designed to operate in one of these environments: urban, highway, and rural - or the combination of some of them. Typically, on a highway, the propagation model, as free-space, can be considered, but it is worth mentioning that this model can still experience interference by the reflection from objects located around the roads. Due to the presence of buildings, trees, and several other objects around in the urban scenario, the signal propagation in this environment experiences shadowing, multi-path, and fading effects. In a rural environment, the local topology should take into consideration (for example fields, dense forests, hills) in the signal propagation, because such a topology can interfere in the wireless communication.

All characteristics listed above pose huge challenges to the design and implementation of VANETs' applications. It should be mentioned that the spatial-temporal constraints, different types of vehicles, and drivers are factors that should be considered in the development of protocols and algorithms in this type of network. Furthermore, due to these intrinsic characteristics of VANETs, solutions developed for traditional ad-hoc networks, such as MANETs, typically experience severe performance degradation when applied to VANETs [92].

2.3.2 Protocol Stack

The protocol stack for VANETs has to deal with communication between vehicles and between vehicles and fixed roadside infrastructures. In the following sections, we present protocols for VANETs according to each layer of the network architecture.

Physical Layer

Due to the unique characteristics of VANETs such as high mobility of nodes, short connection time, and frequent network partitioning, the specification of the inter-vehicle communication (IVC) standard was required. To meet this goal, both the U.S. Federal Communication Commission (FCC) [36] and the European Telecommunications Standards Institute (ETSI) [55] reserve the 5.85 GHz frequency for the spectrum allocation, as shown in Figure 2.5. The spectrum is divided into seven channels of 10 MHz for the American spectrum, and five channels for the European spectrum. As depicted by the figure, both have four service channels (SCHs) for safety and non-safety data exchange and one control channel (CCH). The difference between them is that the American standard has expanded to include two more channels at both ends for special uses [78, 28]. In the European spectrum, on the other hand, the 20 MHz (ITS-G5B band) are allocated for the general-purpose of ITS and the 30 MHz (ITS-G5A band) for road safety services [55]. The main purpose is to enable vehicle-to-vehicle and vehicle-to-infrastructure communications, besides enabling public safety applications. Private applications are also allowed in order to lower costs and to promote DSRC development and adoption [28]. Furthermore, the DSRC supports a vehicle speed up to 200 km/h, the transmission range of 300 m (up to 1000 m), and the default data rate of 6 Mbps (up to 27 Mbps) [79].

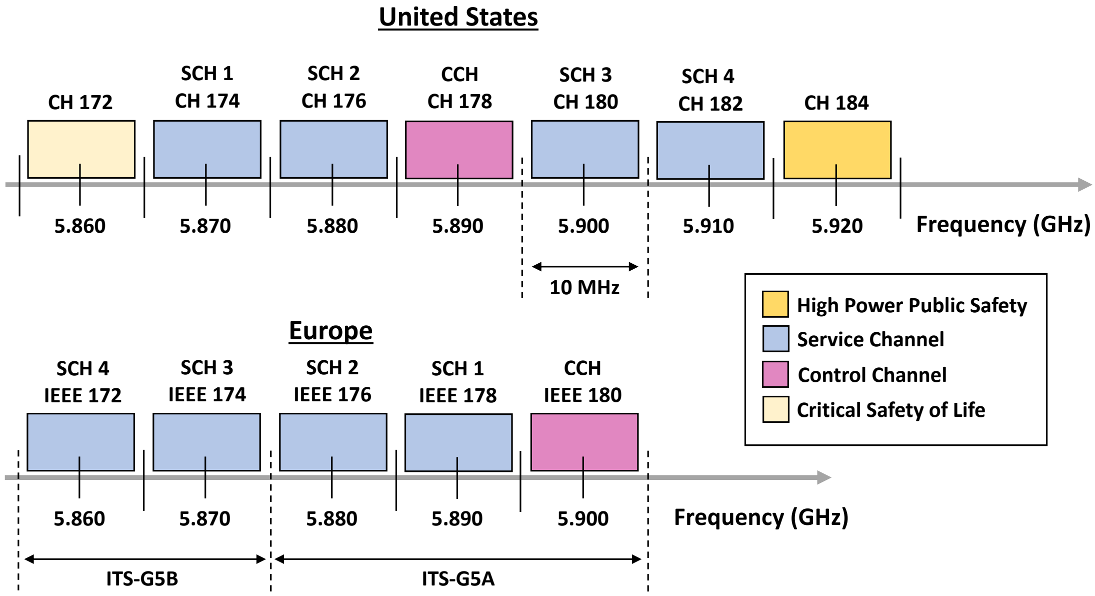


Figure 2.5: Frequency allocation of FCC (top) and ETSI (bottom).

DSRC radio technology is also known as IEEE 802.11p WAVE (Wireless Access in Vehicular Environments), a standard designed to support wireless access in VANETs. The IEEE 802.11p standard is meant to *(i)* describe functions and services that coordinate the operation in a rapidly varying environment and exchange the message without having to join a Basic Service Set (BSS), as in the traditional IEEE 802.11 use case; *(ii)* IEEE 802.11p also defines techniques and interface functions that are controlled by the IEEE 802.11 MAC layer. Therefore, it is limited by the scope of the IEEE 802.11 standard, which means that the physical and MAC layers work within a single logical channel [38].

MAC Layer

The intrinsic characteristics of VANETs, listed in Section 2.3.1, make their qualitative and quantitative analysis particularly critical, mainly when designing medium access control (MAC) layer protocols. In addition to that, the MAC protocol design should take into consideration different types of messages (event-driven messages, periodic messages, and informational messages) traveling on the network. Each type of message has different priorities and goals - for example, event-driven messages are alert messages broadcasted to other vehicles about unsafe situations that have been identified. Such messages have a very high priority [150]. This type of message is essential to the operation of VANET applications. The big challenge for applications using such a message is to make sure that all vehicles intended to benefit from this message receive it correctly and quickly [150]. Periodic messages are disseminated to notify nearby vehicles about the vehicle's current status (speed, position, and direction [122]). Usually, the data of this message is useful to all vehicles around the sender. Informational messages are non-safety messages, with a focus on infotainment applications in order to make driving more convenient and comfortable. Unlike Event-driven messages, this type does not need high priority but may require a high transmission rate.

Here, we have made a brief and concise classification of protocols dealing with MAC issues. MAC protocols can be divided into two top-level categories: *contention-free* and *contention-based*, as referred to in Figure 2.6. Contention-free MAC protocols are based on sharing the channel efficiently at high uniform load [14]. Time-Division Multiple Access (TDMA), Frequency-Division Multiple Access (FDMA), and Code-Division Multiple Access (CDMA) are some examples. Pure contention-free MAC protocols are more appropriate for static networks and/or networks with centralized control [131].

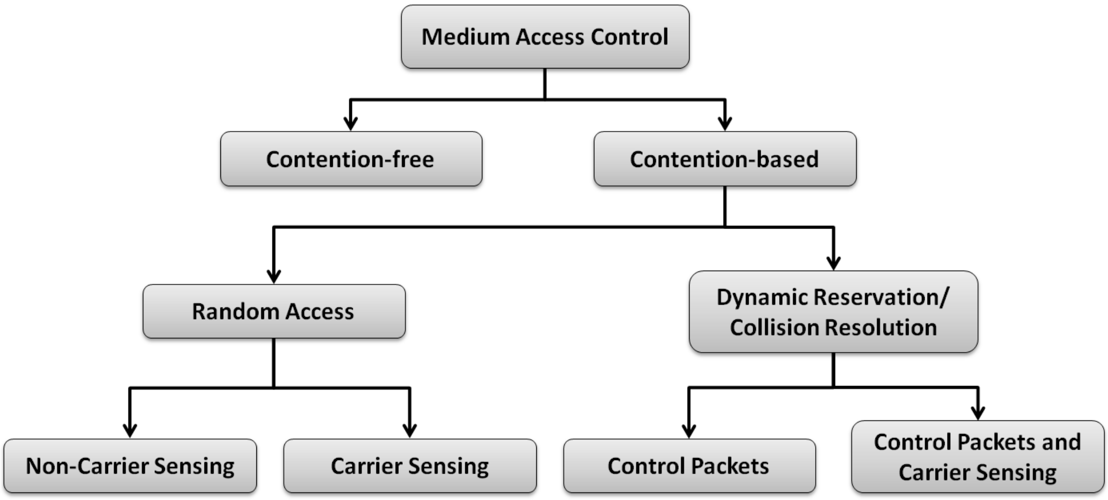


Figure 2.6: Classification of ad-hoc MAC protocols [131].

Contention-based MAC protocols, on the other hand, are based on competition for shared wireless channel access among attending nodes. Competition-based protocols are classified into *random access* and *dynamic reservation/collision resolution* protocols. In random access protocols, such as ALOHA, a node may access the channel whenever it is available. A modification of ALOHA, namely Slotted ALOHA, includes synchronized transmission time-slots alike to TDMA protocol. In this case, nodes can transmit only at the beginning of a time-slot leading to doubled synchronization [141]. The Carrier Sense Multiple Access (CSMA) is another random access protocol that decreases the possibility of packet collisions and improves the throughput due to carrier sensing mechanisms. The main advantage of random access protocols is that they are not susceptible to mobility and topology changes. Therefore, vehicle movements do not impose any reconfiguration overhead due to the network topology changes. In order to deal with the hidden and exposed terminal station problems, researchers have designed several protocols utilizing dynamic reservation and/or collision resolution such as Multiple Access Collision Avoidance (MACA) and MACA for Wireless LANs (MACAW). Both protocols apply Request-To-Send/Clear-To-Send (RTS/CTS) control packets to prevent collisions. Besides them, there are others that combine both carrier sensing methods and control packets such as Floor Acquisition Multiple Access (FAMA), Sensor-MAC (SMAC), and IEEE 802.11 CSMA/CA [131].

Network Layer

In VANETs, routing protocol strategies must be designed and implemented to provide reliable communication and minimum interruption probability. Vehicular networks can support different communication approaches, such as:

- *Unicast communication*: The main purpose is to perform data transportation through the ad-hoc network from a source node to a certain destination (the other vehicle or RSU). The communication may consist of just a single-hop to route the message or over multiple hops toward the destination node. For multihop routing, a number of routing protocols for ad-hoc networks can be considered proper. The destination node may be at either a known location or an estimated location inside a specified range. There are two ways to implement unicast routing such as uni- or bidirectional. The latter is the case for applications that require connection-oriented communication as opposed to several warning applications, for which widespread unidirectional distribution is essential [123]. The goal is to use the vehicular network for transporting messages, not the distribution of messages [123];

- *Geocast communication*: The main purpose is the immediate distribution of information in a geographic area, for example, to alert approaching vehicles about an unexpected situation or unusual road condition that requires attention by drivers. In the geocast mode, the sender of the message defines a target region for the message to be sent and attaches such a region to the message. After that, the message is transmitted to all immediate neighbors within the transmission range. Each receiver located inside the specified destination region sends the message in a broadcast fashion. It is worth mentioning that in a situation of high vehicle density, the forwarding protocol may be optimized to reduce redundancy and improve scalability;

- *Broadcast communication*: Just like the geocast communication, broadcast communication has as the main purpose of the omnidirectional distribution of information. That is, the neighboring nodes that received the message simply forward it to all other neighbors in order to reach the maximum number of nodes. Here it is also necessary to implement some broadcast suppression mechanism to avoid communication overhead. The broadcast communication strategy is also applied at the discovery phase of some unicast routing protocols in order to determine an efficient route from the source vehicle to the target vehicle [59, 84].

Transport Layer

The traditional transport layer is responsible for delivering data to the application process between host computers. The Transport Control Protocol (TCP) is a well-known transport layer protocol that provides reliable end-to-end communication among application processes. To this end, it incorporates different mechanisms such as flow rate control, error recovery, and congestion avoidance. In traditional wired networks, the packet losses or transmission errors are considered to be a consequence of network congestion, since the problems due to route disconnection are minimal [76]. When network channel congestion is detected, the TCP sender often reduces the sending rate, and also adjusts the congestion window size to decrease the system load. Another well-known transport layer protocol is the User Datagram Protocol (UDP). Unlike TPC, UDP provides no guarantee of package delivery between the application process.

It is well-known that TCP and UDP, end-to-end control protocols, were initially designed for the wired network and they do not perform well on wireless networks [135]. In a VANET environment, the frequent communication disruptions, caused by vehicle movement or natural obstacles (trees, buildings, and others), are one of the main problems for the transport protocol. Furthermore, in a VANET, packet loss and end-to-end delay can be caused by the high channel contention, channel interference, or frequent connection breakage. Thus, it is essential for the transport layer to have awareness about the channel quality condition. Through awareness, VANET applications can work adaptively according to the situation.

A VANET transport layer has to deal with multi-hop and broadcast communication, as well as considering routing protocols. The routing protocol design should be taken into consideration to minimize the end-to-end delay during broadcasting messages. During the design, one should also implement a control mechanism to avoid broadcasting storm problems because such a problem occurs whenever the wireless channel is accessed simultaneously by all vehicles inside the transmission range. In addition to avoiding the broadcast storm problem, network underutilization of bandwidth and unnecessary retransmission can also be avoided by the mechanism. In order to accomplish multi-hop data dissemination and to restrict excessive retransmission of messages, VANET applications need to pick up an optimal next hop vehicle as the forwarder to continue data dissemination. Several strategies have been applied for picking up a proper forwarder. Most of the proposed multi-hop broadcast protocols pick up the farthest vehicle in the transmission range as the forwarder [6, 10, 106]. Depending on the wireless channel conditions, the farthest vehicle will not always be the best one to forward the message. Thus, the best link quality is another strategy applied, i.e., the vehicle with the best channel condition will be picked up as the next forwarder [114, 119, 144]. Probability-based forwarding is another well-known strategy. In this strategy, the vehicles will forward the message with a certain probability attributed to them, thus the number of rebroadcasted messages will be reduced as only a few vehicles will participate in the forwarding process. Usually, these protocols dynamically assign value according to vehicle location and density of the network [89, 121, 147].

2.4 Final Remarks

The vehicular network is an essential paradigm for near future ITS applications, smart vehicles, and smart infrastructure. VANETs comprise vehicles equipped with the capability to establish wireless communications and self-organize into a collaborative network. Through this kind of network, countless applications can be proposed and implemented, making travel safer, more efficient, and more pleasant to end-users. In fact, VANETs are likely to become the most important achievement of MANETs. This chapter has brought discussions on the main characteristics of vehicular ad-hoc networks, architecture details, constraints of layers, and protocols, also including a discussion about intelligent transportation systems.

Chapter 3

A Mobility Pattern Analysis Approach to Improve Urban Mobility

3.1 Introduction

Every year, the number of vehicles in urban areas increases exponentially, which is not followed by road infrastructure expansion. This scenario leads to road traffic congestion, which is a significant problem in modern societies, resulting in millions of gallons of fuel consumed and time wasted in traffic. Consequently, the performance of many sectors in urban services (such as health, economy, environment, and daily routine activities) is compromised [21]. Additionally, this existing scenario generates high financial losses. For example, A&M Transportation Institute and INRIX calculates that in 2018, the US lost [74]: *(i)* $87 billion dollars due to traffic congestion; *(ii)* 6.9 billion hours of delayed person-hours; and, *(iii)* 3.1 billion gallons of wasted fuel.

Therefore, over the last years, many researchers from both industry and academia are concentrating their efforts to deploy ATMS into urban centers. Such systems aim to explore different technologies, such as sensing and wireless communication, in order to improve urban traffic management [47]. In recent years, ATMS solutions based on VANETs were proposed to generate a global view of the road network, allowing the detection of all possible road traffic congestions [21, 48, 100, 109, 137]. These solutions apply real-time processing of information about the route to be traversed by all vehicles. This kind of approach has two problems: *(i)* the data computing to assign new alternative route for each vehicle is very intensive, i.e., if the processing time is too long, then vehicles can be already on congested roads; *(ii)* the intensive communication among the vehicles and between vehicle and central server results in a network overload.

In order to overcome these problems, we proposed **APOLO** (context-**A**ware and **P**e**O**ple-centric vehicu**L**ar traffic rer**O**uting), a people-centric (driver's information) approach based on VANET technologies to improve urban mobility. Our approach has two distinct stages: *(i)* - *Offline*, in which the historical data processing of global view of road network is performed in order to generate the mobility patterns; and *(ii)* - *Online*, in which vehicles in the route to congested roads are re-routed.

APOLO acts as a centralized traffic monitoring system, having an overview of road networks without the need for periodical information about the vehicles to perform real-time processing that information. Furthermore, APOLO pro-actively classifies, in advance, traffic levels on the road network based on historical knowledge of mobility patterns of drivers. Since the human movement has a high degree of spatio-temporal regularity [64], APOLO uses space and temporal information about mobility patterns, as well as information about the road map, in order to identify traffic congestion, allowing the rerouting planning of vehicles. Employing these two parameters, we can bring awareness and enhance the intelligence of systems by analyzing the spatio-temporal data. The purpose of using historical mobility is twofold: *(i)* to obtain a global view of the road network; and *(ii)* to avoid the constant data exchange between the vehicle and the central server.

The chapter organization is the following. The next section discusses realistic mobility traces available and analysis of the driver's mobility patterns according to the chosen dataset. Section 3.3 presents the proposed solution for vehicular traffic rerouting based on mobility patterns of drivers. Performance evaluation and results are discussed in Section 3.4 and Section 3.5 concludes the chapter.

3.2 Realistic Mobility Traces

Although the movement behavior of humans has a significant degree of variation, human mobility can generate structural movement patterns [35, 64, 149]. This fact occurs because the displacement of people is constrained by a distance that they can travel during a period of time. Thus, mobility patterns can be shaped by social relationships once we may be more likely to visit favorite places, friend's houses, and workplaces.

Therefore, a better understanding of human mobility can bring benefits to many urban services, mainly the ones related to the user's location. For example, location-based recommendation, content-based delivery networks, and traffic management [57]. Next, we describe situations where this kind of information can improve urban services.

3.2.1 Trace of Luxembourg

A trace usually describes the movement of objects by a temporal sequence of spatial points with their timestamps. It has information about people and dynamic cities, such as vehicular mobility, human activity, and social events.

Vehicular traces are a kind of dataset that contains information about the movement of vehicles within a specific area. In this way, these traces can present the behavior of drivers in a particular scenario. The vehicular mobility traces are created through the merge of the map and vehicular traffic information with a vehicular mobility simulator. Several vehicular mobility traces of real cities can be found in the literature, such as San Francisco[1], Shanghai[2], Cologne[3], and Luxembourg[4]. However, only the Luxembourg trace includes a realistic trace with public

[1]http://crawdad.org/epfl/mobility/20090224/

[2]http://wirelesslab.sjtu.edu.cn/taxi-trace-data.html

[3]http://kolntrace.project.citi-lab.fr/

[4]http://www.vehicularlab.uni.lu/lust-scenario/

and personal vehicles within a period of at least 24 hours. Therefore, the Luxembourg trace was applied in this thesis.

3.2.2 Analysis of the Driver's Mobility Patterns

Mobility patterns can be modeled through spatial and temporal variables [47, 64]. Firstly, it was necessary to analyze the dataset on the temporal context, as shown in Figure 3.1. From the Luxembourg trace, it is possible to observe two rush-hour peaks (blue line), one in the morning (08:00), and another one in the evening (18:30), as well as the off-peak period around lunchtime. Additionally, it is also possible to observe that during rush hours, vehicles stuck on traffic congestions (red line) increase significantly. In each rush-hour peaks, there is around 5000 cars in the scenario, where among them, 1500 cars are stuck in the road traffic congestion.

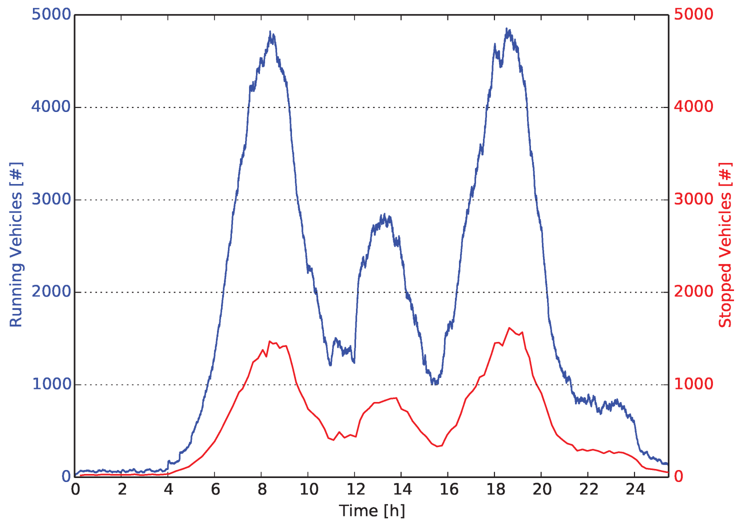
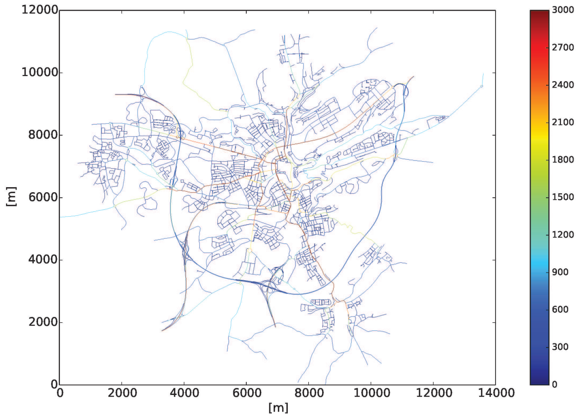


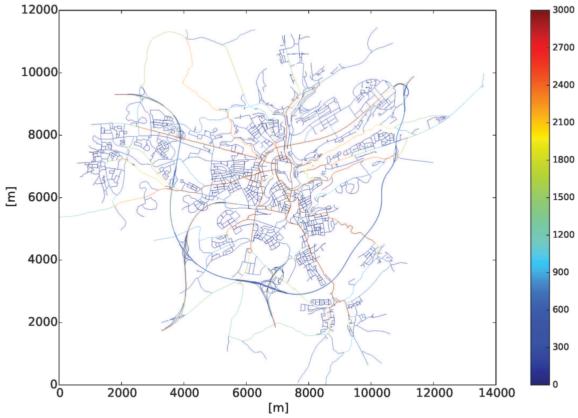
Figure 3.1: Traffic information from Luxembourg trace.

In the next step, the data set in the spatial context was analyzed. Figures 3.2(a), 3.2(b) and 3.2(c) illustrate heat maps of the traffic simulation at 08:00, 13:45, and 18:30, respectively. In these figures, map colors range from blue to red, where the red color represents roads with high vehicle density and the blue color depicts the low-density case. According to the information presented, drivers prefer to travel through avenues of the city, instead of the side streets. To confirm this conjecture, we measured the cumulative distribution function (shown in Figure 3.3) of the vehicle density. We can see that the vehicle density of 90% of road segments is around 10 vehicles/km and this confirms the analysis of Figures 3.2(a), 3.2(b), and 3.2(c).

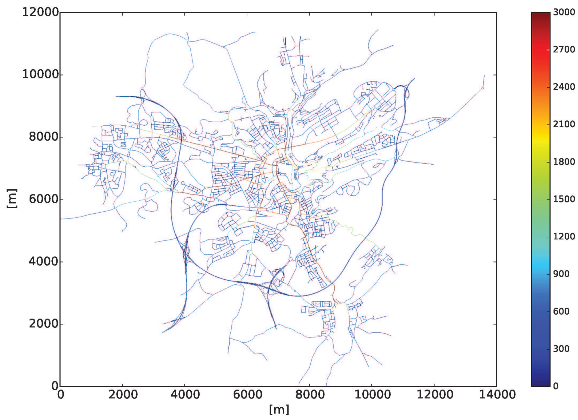
Future road traffic congestion happens when several drivers take exactly the same road inside the same future time window. We assume that our system has historical information on mobility patterns of drivers, road network topology, road network capacity, legal speed limits, and average speed. In this stage, we generated a set of weighted graphs, where the weight represents the level of network density (more details will be presented in Subsections 3.3.1 and 3.3.2). Each weighted graph of the set represents a determined time, and all these data processing tasks were made offline.



(a) Morning rush hours peak.

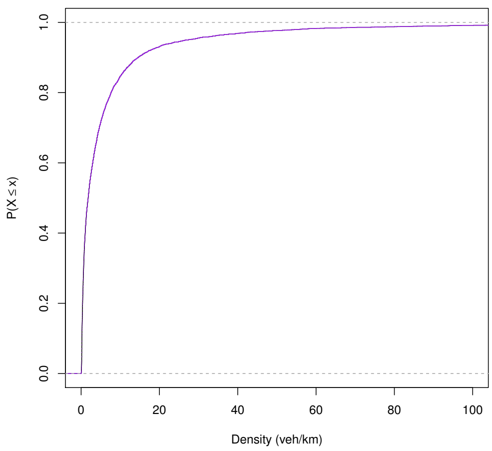


(b) Lunch time.

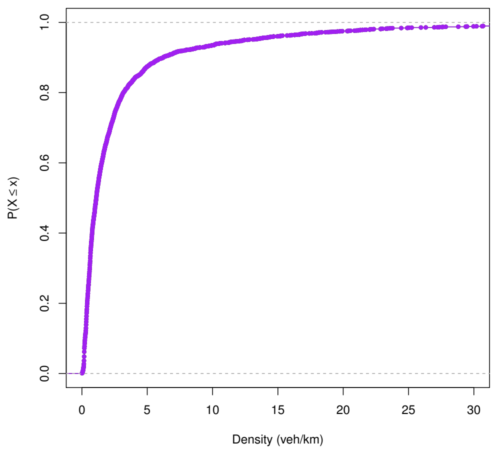


(c) Evening rush hours peak.

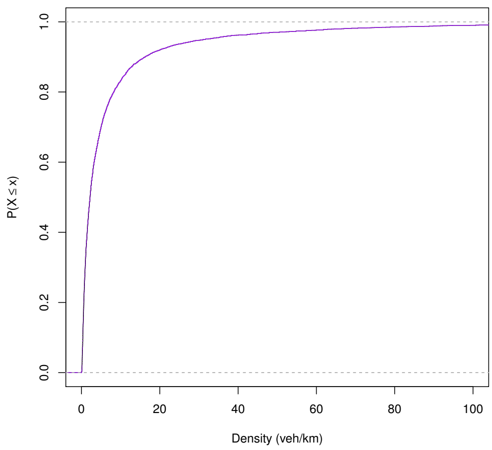
Figure 3.2: Heat maps of traffic simulation.



(a) Morning rush hours peak.



(b) Lunch time.



(c) Evening rush hours peak.

Figure 3.3: Cumulative distribution function.

3.3 APOLO: Context-Aware and People-Centric Vehicular Traffic Rerouting

This thesis presents the context-**A**ware and **PeO**ple-centric vehicu**L**ar traffic rer**O**uting (**APOLO**) approach to avoid road congestion caused by an expected event (traffic jam). APOLO makes use of the extensive knowledge of the vehicular traffic behavior of the city to achieve the desired goal.

3.3.1 Road Network Representation and Estimation

APOLO builds a set of weighted graphs using the road map information and mobility patterns of drivers (based on spatial and temporal analysis). Before explaining how the weight was calculated, a definition of the road network structure is necessary.

Definition 1 *We consider a road network as a graph $G = (V, E)$, where intersections or dead ends correspond to set of vertices $V = \{v_1, v_2, ..., v_r\}$, while road segments correspond to set of edges $E = \{e_1, e_2, ..., e_s\}$, and an edge k is represented as $e_k = (v_i, v_j) \in E$ and $i \neq j$. Let p_i be the route of a vehicle i from two points (origin and destination), i.e., set of ordered edges. Let $N = \{n_1, n_2, ..., n_t\}$ be a set of nodes (vehicles) and $P = \{p_1, p_2, ..., p_t\}$ a set of path (routes) for each n_i that can be defined in G. Then, the route of a particular vehicle k can be defined as follows $p_k = \{e'_1, e'_2, ..., e'_u\}$, where e'_i represents the ith edge and u represents the total number of path to be covered.*

Furthermore, the weight of each edge (e_i) denotes the road traffic density and it is represented by $W = \{w_1, w_2, ..., w_i\}$. The weight equation was modeled to be inversely proportional to the vehicular traffic condition, i.e., congested roads have greater weight than free-flow roads, as shown in Equation 3.1. In this case, edges with low utilization rate are associated with lower weight, where $w_i \in (0, 1]$.

$$w_i = 1 - \frac{v_i^{avg}}{v_i^{lim} \times d_i} \mid d_i > 0 \tag{3.1}$$

where v_i^{avg}, v_i^{lim}, and d_i represent the average speed, maximum road speed, and density, respectively, of e_i.

3.3.2 Traffic Condition Classification

In this thesis, K-Nearest Neighbor (KNN) [85] algorithm is used to perform traffic condition classification, since it often achieves near-optimal results with low complexity in many domains [42].

Next, the KNN algorithm is explained, where the following notation is applied. Assuming the need to classify M entities into N classes, let $C = \{c_1, c_2, ..., c_N\}$ be a set of classes, while training dataset corresponds to $T = \{(x_1, c_1), (x_2, c_2), ..., (x_s, c_t)\}$ of M entities x_s ($s = 1, 2, ..., M$) and their corresponding class label c_t ($t = 1, 2, ..., N$) in C. According to KNN algorithm, an unclassified example, x_i, is attributed to a class that represents the class majority of its k-nearest neighbor in T.

Overall, the KNN algorithm needs a sample database to be trained. To this end, we built a synthetic dataset according to the Highway Capacity Manual (HCM) [95]. This manual includes guidelines, concepts, and methods for measuring the quality of service, based on speed, vehicle length, road capacity, and density of vehicles. Such a dataset was developed based on the Level-Of-Services (LOS) from the manual. HCM employs six different LOS (A to F) to define the traffic conditions on each road segment. A denotes the best quality of service (free-flow conditions - 0) and F denotes the worst (severe traffic congestion - 1).

From the training dataset, it is possible to identify all levels of service on the road, where each one of them is based on traffic density proposed by HCM. Thus, traffic condition classification was used to generate the set of weighted graphs. Furthermore, to avoid false positives, the road density was combined with the average road speed to define the traffic condition. It is worth mentioning that the density is given by the percentage of vehicles on the road by the maximum capacity. Therefore, the traffic congestion classification is defined such as free-flow = 0, slight congestion = 1/3, moderate congestion = 2/3 and intense congestion = 1 [21], as shown in Table 3.1. This classification is constantly made at a predefined interval t, where t is defined by the application.

Table 3.1: Traffic Condition Classification.

		Density			
		Low	Medium	High	Very High
Speed	Fast	Free-flow	Free-flow	Free-flow	Light
	Medium	Free-flow	Free-flow	Slight	Moderate
	Slow	Free-flow	Free-flow	Moderate	Moderate
	Very Slow	Slight	Moderate	Moderate	Intense

3.3.3 Congestion Identification and Rerouting Strategy

Periodically, APOLO checks, in advance, the level of network density in the set of weighted graphs to detect signs of road traffic congestion (Moderate and Intense). Thus, APOLO identifies a sign of road congestion, it plans the rerouting just the vehicles that will move toward the congested road, and their final destination is not on it.

Our rerouting strategy uses a globally optimal approach for all vehicles in the road network. Unlike traditional strategies, where optimal routes are selected individually for each vehicle (it may cause switch traffic congestion to another spot), APOLO applies a global strategy. The global strategy can maintain high traffic flow, and for that, some vehicles may have an additional travel distance in their route.

The rerouting strategy was implemented as a greedy search algorithm based on the weight of each edge, i.e., the next road is selected according to the lowest weight. Moreover, after selecting the next road, APOLO updates the edge weights, based on new vehicle routes, and the mechanism continues until establishing a complete route. Every time a new route is built to a particular vehicle, APOLO sends the updated route information directly to the defined recipient.

Algorithm 1 presents the main process of our approach. It has as input the set of vehicles (represented as N), road network (represented as R), and set of historical data of mobility patterns of drivers (represented as I). The $output$ represents a new alternative route of each vehicle that will move toward the congested road. Algorithm 1 has two phases: *(i)* - *Offline* - historical data processing to generate a set of weighted graphs; and *(ii)* - *Online* - the selection of vehicles to be rerouted, when traffic congestion is identified.

Algorithm 1: Congestion avoidance and control

```
   inputs : N set of vehicles; R road network; I set of historical data of mobility patterns of
            drivers
   output : New alternative path
   Offline: Generation of a set of weighted graphs:
            setGraph = setWeightedGraph(R, I)
   Online : Selection of vehicles to be rerouted, when traffic congestion is identified
1  foreach period of time do
2      congestedRoads = congestionIdentification(setGraph, N);
3      if #congestedRoads > 0 then
4          foreach road ∈ congestedRoads do
5              vehicles = selectedvehicles(roads)
6          foreach veh ∈ vehicles do
7              if vehDest ≠ congestedRoad then
8                  newRoute = getNewRoute(veh, setGraph);
9                  setGraph = updateSetGraph(newRoute);
10                 sendMessage(veh, newRoute)
```

APOLO periodically checks possible road traffic congestion (Lines 1 and 2 in Algorithm 1). Whenever a possible congestion is identified (Line 3), APOLO reroutes just the vehicles that will move toward it and their final destination is not the congested road (Lines 4 to 8). After updating the graphs (Line 9) a message is sent to the vehicle to deploy the new path (Line 10).

3.4 Performance Evaluation and Results

It is worth mentioning that no selfish nodes were considered (i.e., all vehicles travel by the recommended path from APOLO). Additionally, CO_2 emission and fuel consumption were calculated from the model implemented in SUMO (HBEFA-v3.1-based[5] - Handbook Emission Factors for Road Transport) functionality to evaluate emissions. Finally, the results are presented with a confidence interval of 95 %.

During the performance evaluation, the value of the K (number of neighbors of the K-NN classifier algorithm) was experimentally chosen to be 5. In order to evaluate the performance of the APOLO, six metrics were used and are described in detail below:

- *Travel distance*: Average distance traveled by all vehicles;

[5]https://sumo.dlr.de/docs/Models/Emissions.html

- *Travel time*: Average travel time of all vehicles;
- *Fuel consumption*: Average fuel consumption of all vehicles;
- *Idle time*: Average time that all vehicles got stuck in a traffic jam;
- *CO_2 emission*: average CO_2 emission of all vehicles;
- *Average speed*: Average speed of all vehicles.

We compare APOLO with the well-known routing algorithms called Dynamic Shortest Path (DSP) and Random k-Shortest Paths (RkSP) [109], and both have a global view of the road network. Both algorithms use a classical rerouting strategy, based on the shortest time path, to determine a new path. The main difference between the two algorithms is that RkSP makes a random selection, among k shortest paths, between the current position and the destination. The parameter k and routing interval of the RkSP algorithm, were configured according to reference [109] (routing interval = 150 s and k = 3). Additionally, we compare our proposal with the original vehicular mobility trace of Luxembourg (OVMT), in other words, in the absence of any rerouting strategy.

3.4.1 Performance Analysis and Discussion

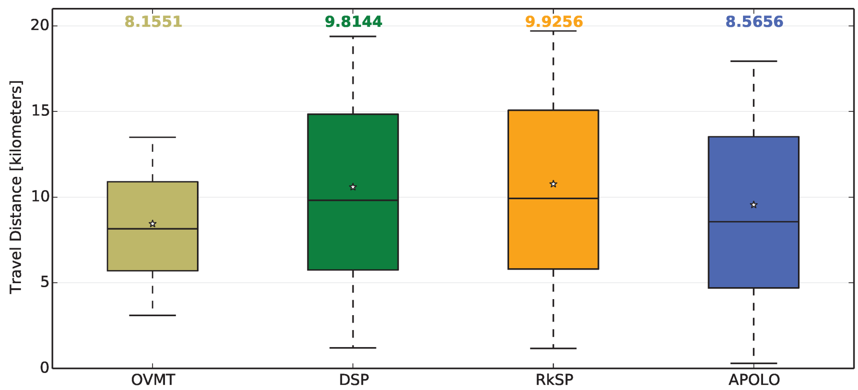


Figure 3.4: Travel distance.

Figure 3.5 shows that our approach has average travel time significantly reduced of approximately 20 %, 13 %, and 17 % compared to OVMT, DSP, and RkSP, respectively. Travel time is reduced due to the knowledge of mobility patterns of drivers and thus identifying future congestion formation. Furthermore, improving the vehicular traffic flow on the entire Luxembourg scenario. The strategy applied in DSP and RkSP, which rerouting periodically all vehicles moving toward road traffic congestion, reduces the average travel time of 9 % and 5 % compared to OVMT, respectively. Moreover, this constant rerouting process increased the average travel distance for both approaches, see Figure 3.4.

Similarly, as DSP and RkSP, APOLO presents a small increase in average travel distance of approximately 5 % compared to OVMT, as shown in Figure 3.4. The reason for this is that APOLO employs an alternative route for all vehicles, whose trajectory will travel along the congested roads. Therefore, some vehicles can have to travel a greater distance to achieve their

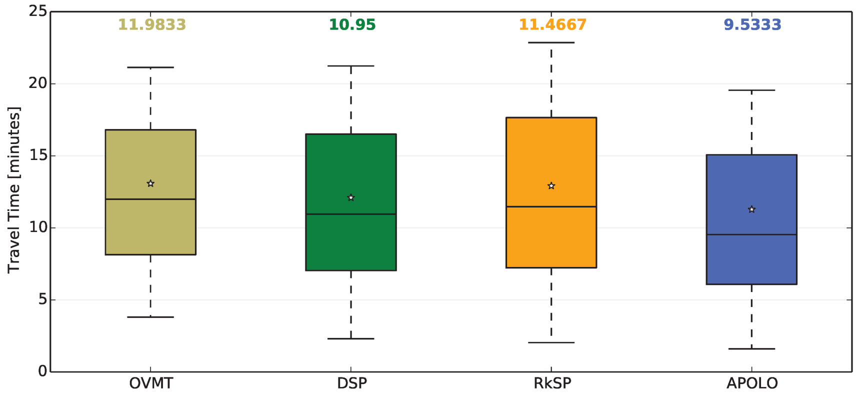


Figure 3.5: Travel time.

destination. On the other hand, APOLO has a shortened average travel distance by approximately 13 % and 14 % compared to DSP and RkSP, respectively. We can observe that our approach has a significant travel time reduction with a small increase in travel distance. RkSP randomly chooses a route from a set of k shortest path routes, in order to avoid switching traffic congestion to another spot. However, this strategy increases travel distance; i.e., the random selection policy of k possible trajectories tends to choose a longer path, as can be observed in Figure 3.4.

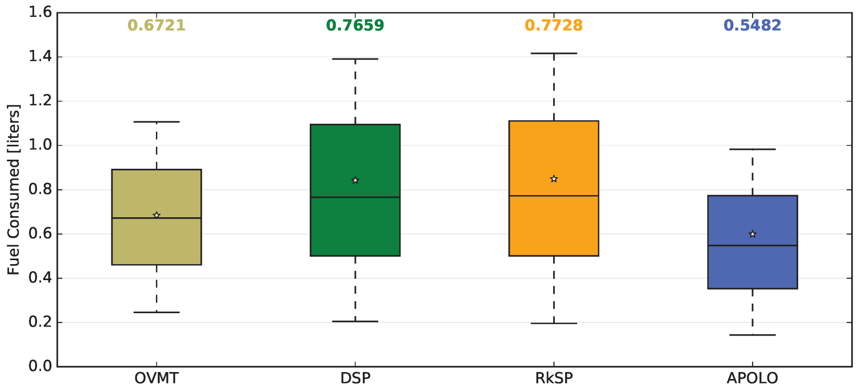


Figure 3.6: Fuel consumed.

Figure 3.6 shows fuel consumption results, as can be observed APOLO shows an average fuel consumption of 0.5482 liters, while OVMT shows an average 0.6721 liters, DSP consumes 0.7659 liters, and RkSP consumes 0.7728 liters. In other words, APOLO has notable saving in fuel consumption of approximately 19 % compared to OVMT and a saving of approximately 29 % and 30 % compared to DSP and RkSP, respectively. In APOLO's approach, the vehicles travel a greater distance, on average, in comparison to OVMT (see Figure 3.4), however, it has the shortest travel time (see Figure 3.5). That is, APOLO performs the rerouting of vehicles so that the new path has the lowest possible traffic of vehicles. As a consequence of this approach is the shortest travel time, because the vehicles do not get stuck in traffic congestion. Furthermore, the new routes help reduce the number of accelerations and decelerations caused by traffic congestions, thus saving fuel consumption. The behavior in Figure 3.6 is observed in Figure 3.8 as well, because both, fuel consumption and CO_2 emissions, are directly related.

Figure 3.7 shows the average idle time. Among all, APOLO has the lowest idle time of

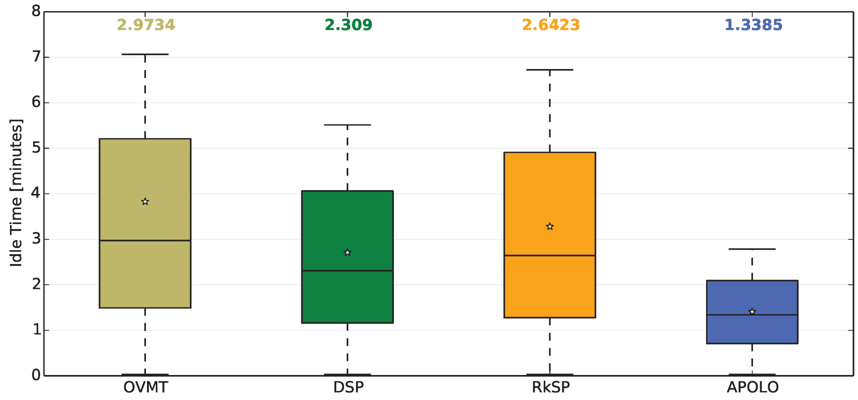


Figure 3.7: Idle time.

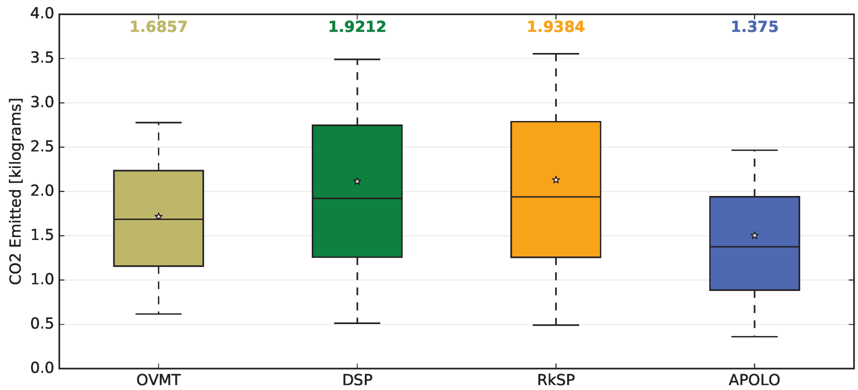


Figure 3.8: CO_2 emitted.

approximately 55 %, 43 %, and 50 % lower than OVMT, DSP, RkSP respectively. The APOLO has the lowest idle time because it uses information, such as historical knowledge of mobility patterns of drivers and global view of traffic conditions to select the best path of each vehicle. Despite both approaches, DSP and RkSP, have a global view, and their rerouting decisions are made in real-time, sometimes this decision may be too late. That way, the vehicles can go into a congested road and have no option to exit, thereby increasing the idle time. Finally, our approach has an increase in average speed to approximately 39 %, 3 %, and 6 % compared to OVMT, DSP, and RkSP respectively (Figure 3.9). Although the average speeds between DSP and APOLO are close to 54 km/h, the APOLO stands out for having the lowest idle time (Figure 3.7).

3.5 Final Remarks

In this chapter, we proposed a people-centric approach for vehicular traffic management in urban centers, named APOLO. The main idea of APOLO is to periodically analyze the spatial and temporal parameters of mobility patterns of drivers to manage vehicular traffic flow in urban centers. This approach makes APOLO different from existing approaches that explore only the characteristics of public roads (speed limit and length of the roadway) and vehicle density at each moment. Based on the observation of simulation results, our approach has shortened travel time significantly, e.g., approximately 20 % compared to OVMT, with a small increase

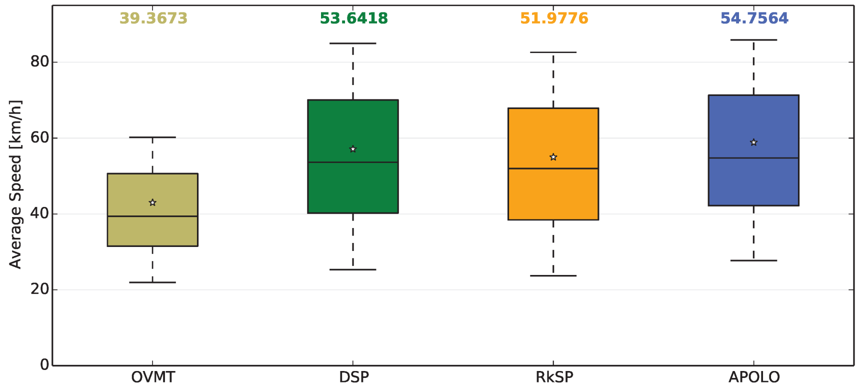


Figure 3.9: Average speed.

in average travel distance. Despite the increase in travel distance, average fuel consumption was lower than other approaches, by minimizing the unnecessary acceleration and deceleration caused by road traffic congestion. Besides, the APOLO can reduce around 55 % idle time and increase of 39 % in average speed compared to OVMT. Through the numerical results, we observe that our approach may represent an exciting alternative to improve ATMS services. These results show that people-centric services applied to traffic management can help drivers shorten their commuting time.

Chapter 4

Egocentric and Sociocentric Betweenness Measure in VANETs: A Comparative Study

4.1 Introduction

Centrality is a concept widely employed in social network analysis (SNA) to classify nodes as central or, more important, in the network [23, 139]. Several approaches have been developed to compute node centrality [23]; however, the three most commonly used approaches in SNA are degree centrality, closeness centrality and betweenness centrality [61, 102]. Although there are different centrality metrics in the literature, most of them fit into two categories such as radial and medial measures [23]. Radial measures assess information flow that originates from or ends at, a given node. It includes degree and closeness centrality. On the other hand, medial measures assess the geodesic distance that crosses a given node [23], which includes all variations of the betweenness centrality.

The calculation of centrality measures requires global knowledge of the network topology [96], but very often, this knowledge is not available. Besides, it is usually difficult to obtain this information in large-scale or highly dynamic networks. Taking this into consideration, the concept of the ego-network has attracted great attention in the scientific community. This stems from the fact that its topological analysis can be carried out locally by individual nodes without the need for global knowledge of the network [56, 86, 96]. Another advantage of the ego-network is the simple structure to collect data compared to collecting data from the entire network. By definition, the ego-network is a subnetwork centred on a single node, called the ego, whereas one-hop nodes are called alters [86, 96]. In an ego-network, only the nodes that are directly connected to the ego belong to the subnetwork [86, 96].

It is known that the message delivery in VANETs is a difficult task due to the highly dynamic topology [9, 10]. Therefore, a key challenge, in this type of network, is to find a path among the nodes that can provide good information flow. A good alternative is to apply centrality measures. However, some centrality measures may not be appropriate enough in the information flow in the network. For example, degree measure is not suitable for that [4, 7, 41]. On the other hand, the betweenness centrality is more suitable to deal with flow information through the network [39, 41, 133]. Based on the idea of ego-networks and the betweenness centrality measure, another perspective in SNA has emerged, named egocentric betweenness [96]. The egocentric betweenness measure has been adapted for several types of networks such as wireless sensor networks [39], delay-tolerant networks [41] and wireless mesh networks [133]. However, this measure has not been systematically investigated in vehicular ad hoc networks (VANETs), which have unique characteristics such as high mobility of nodes, short connection time, and frequent network partitioning.

The development of services over VANETs has attracted researchers from both academia and industry due to the wide diversity of applications. They can range from vehicle traffic monitoring, system-aided navigation, and cooperative collision warning, to infotainment [15, 67, 68, 128]. Many of these services need to be aware of the local situation [67, 68]. To reach this awareness, one can take advantage of either cooperative awareness message (CAM) [54] (European standard) or the basic safety message (BSM) [75] (American standard). In both standards, the messages contain information regarding vehicle status such as position, speed, direction, location coordinates, and other vehicle information [122]. The process of acquiring local awareness is usually performed by broadcasting one-hop messages. As a result, each vehicle will be aware of its neighbor vehicles within its transmission range. The periodic exchange of one-hop messages is known as beaconing [122].

Due to the instabilities in the communication links induced by the highly dynamic topology, calculating the betweenness centrality scores in a VANET is a challenging task. On the other hand, once having identified the highest-betweenness centrality node in the network, it can be used as a facilitator node to spread the information flow [41]. This measure has been frequently applied in the design of efficient data forwarding algorithms, for instance, in wireless sensor networks [39].

A distributed approach to calculate the egocentric betweenness score was implemented and evaluated with the sociocentric metric in order to prove the feasibility of the egocentric betweenness measure in VANETs. To this end, we use a beaconing mechanism to broadcast one-hop messages about its local information. Once local information is received, each vehicle can compute its egocentric betweenness score. The main goal is to present the similarity of betweenness centrality considering two approaches: local knowledge-based (egocentric) and global knowledge-based (sociocentric).

The remainder of this chapter features the egocentric betweenness measure applied in different areas (Section 4.2). This is followed by the calculation of centrality in sociocentric and egocentric networks in Section 4.3. Section 4.4 presents how the egocentric betweenness measure was computed in VANETs. Simulation experiments and results are presented in Section 4.5. In the end, Section 4.6 gives the final remarks.

4.2 Egocentric Betweenness Measure in Different Areas

In this section, we survey the works that use the egocentric betweenness measure in different areas, such as wireless sensor networks [39], mobile ad hoc networks [41] and wireless mesh networks [133]. Each distinct area has had to deal with several critical issues related to their own characteristics.

Cuzzocrea et al. [39] investigated the problem of the quality of service (QoS)-based topology control over wireless sensor networks. To this end, a weighted, bidirectional topology-control algorithm named edge-betweenness centrality (EBC) was proposed. EBC selects the suitable set of neighbours in which input QoS requirements may be satisfied. The idea here is to select from the target network appropriate logical neighbours of the former nodes, i.e., a subset of neighbours that can be employed to perform application-specific procedures (for instance, message delivery) without the need to include all nodes of the network. The authors have demonstrated that this approach allows achieving a high QoS in wireless sensor networks using evaluating the relationships between entities of the network (i.e., edges). This provides the capability of controlling the information flow, the message delivery, the latency, and the energy dissipation among nodes.

The authors of SimBet routing [41] proposed an algorithm for forwarding data packets in disconnected delay-tolerant MANETs based on social network analysis techniques. For this purpose, they designed and implemented the routing protocol, which used two components: *(i)* betweenness utility, which exploits the exchange of pre-estimated egocentric betweenness centrality scores; and *(ii)* similarity utility, which selects the node that provides the maximum utility for carrying the message. Based on these components, SimBet chooses which node provides the maximum utility for carrying the message. Simulation results have shown that it achieves good performance comparable to epidemic routing, with low network overhead. Additionally, the authors have illustrated that the employment of the egocentric betweenness metric may prove useful in any distributed systems, where global topology knowledge is inaccessible and, especially, where the underlying networks present small-world characteristics.

Vazquez-Rodas et al. [133] proposed a protocol for topology control in wireless mesh networks to improve energy efficiency and the battery lifetime. The proposed mechanisms choose which devices must act as routers, forwarding the data packets received from other hand-held devices to it. In order to select the devices, centrality metrics are applied, from social network analysis, to build a topology control mechanism based on a connected dominating set. The mechanism's implementation and evaluation have been carried out in two modes, i.e., centralized and distributed. In the centralized mode, the three most common centrality measures (degree, closeness, and betweenness) were employed. In the distributed mode, the egocentric betweenness measure was applied. Through the experiment results, it was verified that the use of the centrality measures contributes to better network performance.

4.3 Sociocentric and Egocentric Centrality Measures

In SNA, the centrality measures indicate the importance of a node within a graph. This is performed by taking into account all connections from the node (or the ones that pass through it) to other nodes [41, 61]. The importance of a node can be computed by means of centrality measures such as degree, closeness, betweenness, among many others. SNA can be divided into two network analysis approaches: ego-network analysis (egocentric) and global network analysis (sociocentric). The former studies the relationships existing from the perspective of a participant. The latter tries to observe all relationships between the participants within the network. In this section, we will study the difference between sociocentric and egocentric centrality measures for network analysis. In Subection 4.3.1, the most commonly-used centrality measures in sociocentric analysis will be described, while in Subection 4.3.2, the centrality measure used in the egocentric analysis will be detailed. Finally, Section 4.3.3 gives the complexity analysis of both measures.

4.3.1 Sociocentric Centrality Measures

Centrality measures are the most useful mathematical models developed for SNA [82]. These measures aim to understand the structural properties of social relationships. For instance, a participant with a high centrality score usually has a higher degree of influence than other participants within the network. According to the SNA, the network structure consists of an undirected graph, and its definition is presented below.

Definition 2 *Let $G = (V, E)$ where V corresponds to a set of nodes (v), also called vertices or actors and E corresponds to a set of edges (e, where $e \in E \subseteq V \times V$ is identified by a pair of nodes), also called ties. We represent the neighbourhoods of the node v' as the set of nodes $v \in V$ reachable in r hops (N_r^v). Thereby, $N_r^v = \{v' \in V | v' \neq v \wedge d(v, v') \leq r\}$, where d represents the geodesic distance between nodes. Furthermore, a graph can be defined as a two-dimensional adjacency matrix A, where each element a_{ij} takes a value of one if an edge connects the node i to the node j ($i \neq j$) and zero otherwise.*

Freeman's degree, closeness and betweenness measures are the most commonly-used centrality metrics in sociocentric analysis [23, 41, 61]. They are briefly described below.

Degree centrality is the simplest and the most well-known measure. It assesses the number of direct ties that involve a given node, i.e., it is the number of adjacent edges [61]. A node with a high degree of centrality can be seen as popular because it has a large number of ties to others [20]. According to the work of Wasserman and Faust [139], the degree can also be considered as a measure of local centrality. Therefore, degree centrality of a given node, p_i, can be mathematically represented as:

$$C_D(p_i) = \sum_{j=1}^{N} e(p_i, p_j) \tag{4.1}$$

where $e(p_i, p_j)$ = 1 means a direct link exists between p_i and p_j, otherwise $e(p_i, p_j)$ = 0.

Closeness centrality is defined by the geodesic distance d of a subset of nodes that are mutually connected in the network [61], i.e., it measures how close a node is in relation to all other nodes in the network. This measure can be represented as an indicator of how long information will take to be propagated from a given node to other nodes within the network [102]. Therefore, closeness centrality for a given node, p_i, can be mathematically represented as:

$$C_C(p_i) = \frac{(N-1)}{\sum_{j=1}^{N} d(p_i, p_j)} \tag{4.2}$$

where N is the number of nodes in the network and $i \neq j$.

Betweenness centrality is usually calculated as a fraction of the geodesic distance between all node pairs that pass by a determined node [102], i.e., it is based on the idea that a node is central if it is located on the shortest path between other pairs of node sets within the network. This measure is often applied as a metric of the influence of a node on the spread of information compared to other nodes of the network [20]. Therefore, betweenness centrality for a given node, p_i, can be mathematically represented as:

$$C_B(p_i) = \sum_{j=1}^{N} \sum_{k=1}^{j-1} \frac{g_{jk}(p_i)}{g_{jk}} \tag{4.3}$$

where $g_{jk}(p_i)$ represents the number of geodesic paths that pass through node p_i and g_{jk} represents the total geodesic path between p_j and p_k.

Freeman's centrality measures usually require global knowledge of all network nodes and their interconnections [41, 86, 96]. The problem here is that this knowledge is not always accessible. Furthermore, the applicability of these measures is often difficult in large-scale networks (World Wide Web) and highly dynamic networks (VANETs). This is true because in the first one, it requires a high computational power to compute all the measures, while in the second one, the interconnection topologies change rapidly over time. For this reason, the concept of ego-networks has been introduced [86, 96]. The ego-network analysis can be carried out using only local knowledge, without the need for complete knowledge of the network topology.

4.3.2 Egocentric Centrality Measures

First of all, the definition of ego-networks is needed in order to understand the concept of egocentric centrality measures. By definition, an ego-network is a local subgraph consisting of a single node (ego) in addition to nodes that are connected to it (alters) and all the interconnection links among alters [56, 96]. Figure 4.1 highlights a local subgraph where n represents ego and the one-hop neighbours (1, 2, 3, 4 and 5) denote the alters.

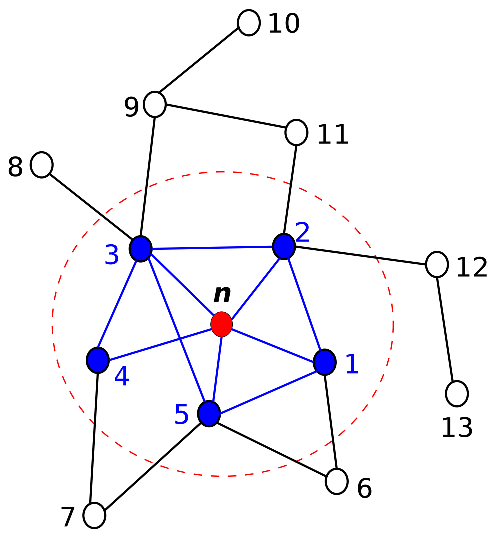


Figure 4.1: An illustration of the ego-network (local subgraph), where n represents the ego and the nodes (1, 2, 3, 4 and 5) denote the alters.

Inside the ego-network, the degree centrality of the nodes can be easily computed, as it is the number of direct connections of one node to its immediate neighbourhood. Because of that, it is possible to conclude that the degree centrality is similar to both egocentric and sociocentric network topologies. Incidentally, this same conclusion was reported by Wasserman and Faust [139]. On the other hand, the closeness centrality measure concerns the geodesic distances from a given node to all other nodes within the network. It is possible to notice that this measure requires the participation of all nodes involved in the network. Thereby, this measure cannot be directly applied in ego-networks, since all geodesic distances from the ego to other nodes are one-hop neighbours by definition, and this holds true because geodesic paths are no greater than two. Among the three measures presented in Subection 4.3.1, the betweenness centrality measure is the most studied in several fields [25, 96]. However, the literature lacks an investigation of this measure on VANETs.

The betweenness centrality in ego-networks will be analysed in the remainder of this section. From now on, we are going to call it the egocentric betweenness measure (EBM). The definition and how it is computed are presented below.

Definition 3 *Once again, let an undirected graph $G = (V, E)$ where V corresponds to a set of nodes (v) and E corresponds to a set of edges (e, where $e \in E \subseteq V \times V$ is identified by a pair of nodes). The neighbourhoods of the node v' are expressed as set of nodes $v \in V$ reachable in r hops. Let N_n^r be the set of nodes that is r hops away from n (ego), i.e., $N_n^r = \{v' \in V | v' \neq n \wedge 1 \leq d(n, v') \leq r\}$, where $d(n, v')$ denotes one hop between n and v'. Thereby, the first-order of node n consists of an undirected graph $G = (V_n^1, E_n^1)$, where the set of nodes corresponds to $V_n^1 = \{N_n^1 \cup \{n\}\}$ and the set of edges corresponds to $E_n^1 = \{(i, j) \in E_n^1 | i, j \in V_n^1\}$.*

The EBM of a certain node, n, can be calculated by the sum of reciprocal values of the $A_n^2[1 - A_n]_{i,j}$, as defined in Equation (4.4) [56].

$$EBM_{(n)} = \sum_{A_n(i,j) \neq 0, i<j} \frac{1}{A_n^2[1 - A_n]_{i,j}} \tag{4.4}$$

where A_n depicts the adjacency matrix of the node n, 1 is a matrix of all ones and the matrix A_n^2 provides the number of geodesic distances of a length of two between node pairs i and j.

Mathematically, an adjacency matrix ($A_{k \times k}$) can represent node-to-node inter-communication links, where k is the number one-hop neighbours. Thereby, each element of the adjacency matrix, $a_{i,j}$, is given by:

$$a_{ij} = \begin{cases} 1 & \text{if a direct link exists between } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$

To demonstrate the calculation of the egocentric betweenness measure using the adjacent matrix, we employed a classical graph example [96], see Figure 4.2.

Just to give one example, the egocentric betweenness score from the perspective of node $W4$ of Figure 4.2 is computed. The following adjacency matrix describes a view of all connection links between $W4$ (ego) and its alters, as well as the connection links between the alter pairs.

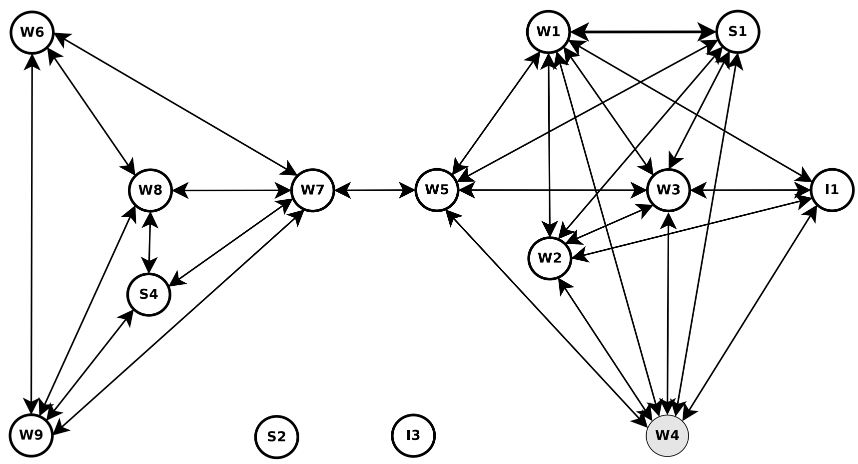


Figure 4.2: A classical graph example [96].

$$W4 = \begin{array}{c} \\ \mathrm{W4} \\ \mathrm{I1} \\ \mathrm{S1} \\ \mathrm{W3} \\ \mathrm{W1} \\ \mathrm{W2} \\ \mathrm{W5} \end{array} \begin{array}{c} \begin{array}{ccccccc} \mathrm{W4} & \mathrm{I1} & \mathrm{S1} & \mathrm{W3} & \mathrm{W1} & \mathrm{W2} & \mathrm{W5} \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Since the adjacency matrix $W4$ is symmetric and according to Equation (4.4), only the non-zero values above the primary diagonal need to be analysed ($i < j$). In this case, the remaining entries of $W4^2[1 - W4]$ are 4, 3 and 4, as shown in the matrix below.

$$W4^2[1 - W4] = \begin{array}{c} \\ \mathrm{W4} \\ \mathrm{I1} \\ \mathrm{S1} \\ \mathrm{W3} \\ \mathrm{W1} \\ \mathrm{W2} \\ \mathrm{W5} \end{array} \begin{array}{c} \begin{array}{ccccccc} \mathrm{W4} & \mathrm{I1} & \mathrm{S1} & \mathrm{W3} & \mathrm{W1} & \mathrm{W2} & \mathrm{W5} \end{array} \\ \begin{bmatrix} * & * & * & * & * & * & * \\ * & * & 4 & * & * & * & 3 \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & 4 \\ * & * & * & * & * & * & * \end{bmatrix} \end{array}$$

Therefore, the egocentric betweenness score of the ego node $W4$ is 0.83 ($1/4 + 1/3 + 1/4$). In this way, by using only the local knowledge available, each node can compute its egocentric betweenness score. Table 4.1 shows the scores of all nodes from the example of Figure 4.2, based on both betweenness centrality measures. Since egocentric betweenness is computed over the geodesic paths of the maximal length of two, the scores found in the egocentric betweenness measure are usually smaller than their sociocentric equivalents. However, an observation that is important to highlight is the similarity ranking of nodes.

The illustrative example given here was based on static networks; however, one of our major challenges is to perform the same calculation in highly dynamic network scenarios such as VANETs. In these networks, the egocentric betweenness score should be updated whenever a

Table 4.1: Egocentric and sociocentric betweenness scores of Figure 4.2.

		Betweenness Centrality	
		Sociocentric	**Egocentric**
	W1	3.75	0.83
	W2	0.25	0.25
	W3	3.75	0.83
	W4	3.75	0.83
	W5	30.00	4.00
	W6	0.00	0.00
Nodes	W7	28.33	4.33
	W8	0.33	0.33
	W9	0.33	0.33
	S1	1.50	0.25
	S2	0.00	0.00
	S4	0.00	0.00
	I1	0.00	0.00
	I3	0.00	0.00

new communication link is established or when a communication link ceases to exist.

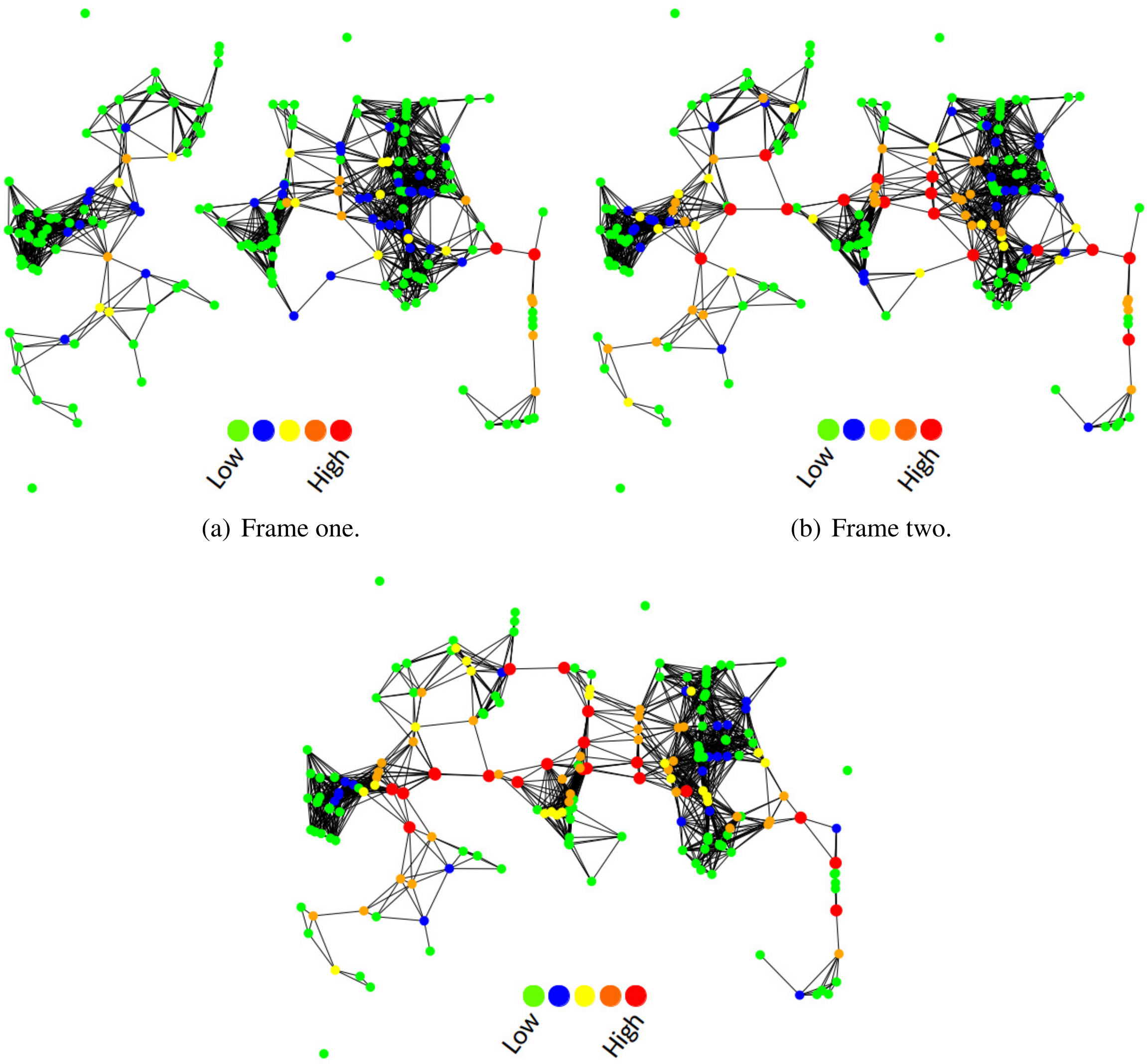
To exemplify how each node behaves and how the network structure can change in a highly dynamic network scenario, in relation to the betweenness centrality score, a set of footprints that describe a frame sequence (Figure 4.3) was illustrated; see Figures 4.3(a), 4.3(b), and 4.3(c). It shows the behaviour of the network topology (or temporal graphs) through a heat map set of our experiment scenario that will be presented later. The density is 150 vehicles/km^2, and the transmission range is 287 m. Each node (or vehicle) is represented by a circle, and every communication link is represented by a bar. Moreover, each node can have five different colours according to the betweenness centrality score, ranging from low to high, as shown in Figure 4.3.

4.3.3 Complexity Analysis of the Sociocentric and Egocentric Measures

In this section, the complexity of the sociocentric and egocentric betweenness metrics is analysed. The main goal is to assess message overhead and time complexity.

For the sociocentric betweenness measure, the nodes need to collect the global network topology information before performing the calculation. A straightforward way is as follows: *(i)* compute the length and the number of geodesic distances between all node pairs; *(ii)* for each node, calculate every pair-dependency, and sum them up. Consequently, this naive algorithm will consume $\Theta(N^3)$ time, where N is the number of nodes of the network. The well-known Brandes' algorithm can be efficiently calculated in $O(NM)$ time [24], where N and M represent the number of nodes and edges of the network, respectively. The message overhead over the entire network generally needs $O(N)$ message copies and $O(D)$ time steps for each node's message, where D represents the network diameter [24].

For the egocentric betweenness measure, the nodes require only local network topology information to carry out the calculation. The EBM calculation demands a computation complexity equal to $O(k^3)$ for a square matrix of $k \times k$ dimensions, where k is the number of alters.



(a) Frame one.

(b) Frame two.

(c) Frame three.

Figure 4.3: frames sequence.

The message overhead over the entire ego-network topology is $O(k)$, since each node needs to send the identification of its neighbouring nodes. Table 4.2 depicts the complexity analysis of the sociocentric and egocentric measures.

Table 4.2: Complexity comparison between sociocentric and egocentric measures.

Measure	**Time Complexity**	**Message Overhead**
$C_B(p_i)$	$O(NM)$	$O(DN)$
$EBM_{(n)}$	$O(k^3)$	$O(k)$

Since it is known that k is typically much smaller than N ($k \ll N$), therefore the local measure approach can bring computational benefits for calculation.

4.4 Egocentric Betweenness Measure in VANETs

Due to the high mobility of the vehicles in VANETs, getting all network topology knowledge is not an easy task. The egocentric betweenness measure is computed using only the available local knowledge; in that case, the adjacency matrix of one-hop neighbours. Each vehicle gets the local knowledge of the network topology by means of periodic beacon packets broadcast by its neighbours. The beacon transmission frequency employed was 1 Hz. Since the vehicle's beacon packets are only useful to adjacent neighbours, the beacons received are not forwarded. Therefore, the information exchanged among vehicles is lists of neighbours, as illustrated in Figure 4.4. In this example with four vehicles, the grey vehicle (labelled as 1), receives the lists of neighbours of all vehicles that are currently within its transmission range (vehicles labelled as 2, 3 and 4). Once having received the lists, the vehicle constructs the adjacency matrix representation and calculates the egocentric betweenness score, according to Subection 4.3.2. Each vehicle updates the egocentric betweenness score, whenever a new list is received.

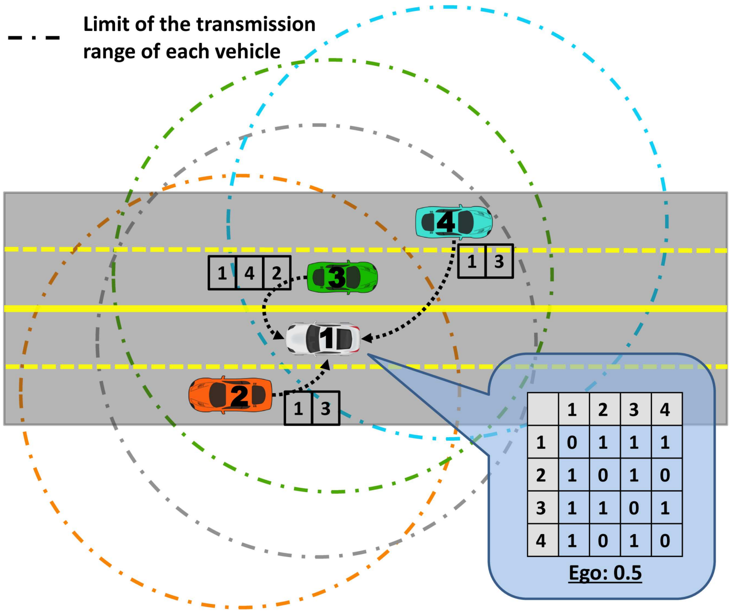


Figure 4.4: An illustrative example of the beacon packets' exchange among the vehicles to calculate the egocentric betweenness score. In this case, the grey vehicle, labelled as 1, is doing the calculation.

The main steps of our proposed approach are presented in Algorithm 2. The algorithm requires the list of neighbours of all vehicles that are currently within the transmission range (represented by L), as input information. The output information is the current list of neighbours and the egocentric betweenness score. Upon receiving a new list of neighbours, the adjacency matrix is updated to represent a new ego-network topology (Lines 2 and 3). After the adjacency matrix is updated, the algorithm computes the egocentric betweenness score (Lines 4, 5 and 6). Thereafter, the list of neighbours is also updated (Line 7). Lastly, a beacon packet containing a current list of neighbours is broadcast (Line 8).

Algorithm 2: Calculation of the egocentric betweenness scores.

inputs : $L = \{l_1, l_2, ..., l_n\}$ list of neighbours of all vehicles that are currently within the transmission range

output: Egocentric betweenness score and list of neighbours

foreach $l_i, i \in [1, n]$ **do**
  if $isNew(l_i)$ **then**
    $A = updateAdjacencyMatrix(l_i)$;
if $wasUpdate(A)$ **then**
  $E = A^2[1 - A]$;
  $egoValue = computeEgoBetweenness(E)$;
  $myListNeighbors = updateMyListNeighbors()$;
$sendBeacon(myListNeighbors)$;

4.5 Simulation Experiments and Results

This work uses a distributed approach to perform the calculation of egocentric betweenness scores in vehicular networks. It consists of four stages, as depicted in Figure 4.5. For the sake of clarity, the figure is divided into four different layers (in a bottom-up fashion). The bottom layer represents the chosen map segment for the evaluation. The layer above it describes the road topology structure of that segment. The third layer shows the vehicle routes and the inter-vehicle communication produced in the simulation. Finally, the top layer depicts the egocentric betweenness calculation results. The next two sections describe the experimental settings (Subection 4.5.1) employed in our simulations and the analysis of the simulation results (Subection 4.5.2), respectively.

4.5.1 Simulation Setup

The experiments were carried out with the aid of three different simulators. It is worth mentioning that all the experiments performed in this thesis were used in the same version of the simulators, as presented below:

- SUMO 0.29.0[1]: An open source road traffic simulation package designed to handle large road networks. SUMO is licensed under GPL;

- OMNeT++ 5.0[2]: A C++ based discrete event simulator for modeling communication networks, multiprocessors and other distributed or parallel systems. OMNeT++ is public-source, and can be used underthe Academic Public License;

- Veins 4.5[3]: An open-source framework for running vehicular network simulations. Such framework integrates OMNeT++ and SUMO and it offers a suite of models for inter-vehicular communications simulation. The physical (PHY) and medium access con-

[1]http://sumo.sourceforge.net/
[2]https://omnetpp.org/
[3]https://veins.car2x.org

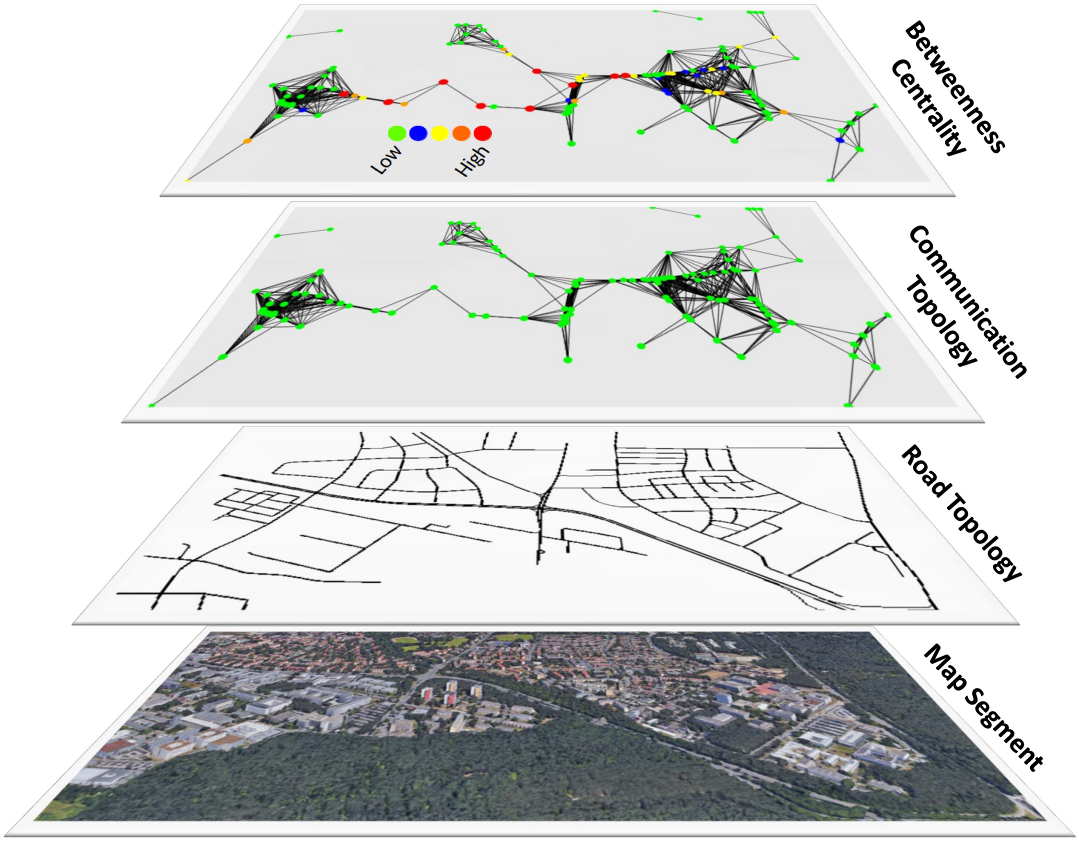


Figure 4.5: The simulation setup layers.

trol (MAC) layers were implemented based on the WAVE (Wireless Access in Vehicular Environment) standard, also known as IEEE 802.11p.

As for simulation parameters, each vehicle had a transmission rate of 6 Mbps, a transmission power of 0.98 mW, a receiver sensitivity of -82 dBm and a transmission range of 287 m. Channel 178 (control channel–CCH) was used to exchange beacon packets, thereby excluding the effects caused by channel switching between the CCH and the SCH (service channel).

In order to evaluate the applicability of the egocentric betweenness approach in vehicular networks, a real map clipping of the Erlangen area (Germany), obtained from OpenStreetMap[4], was used (Figure 4.6). Meanwhile, a set of feasible vehicle routes was synthetically generated with the aid of SUMO. Vehicle mobility used the Krauss car following model [83]. Five different sets of vehicle traffic densities were generated to validate our approach (40, 60, 80, 100 and 150 vehicles/km^2).

Finally, all experimental results of this work were executed thirty-three times on different vehicle traffic densities with a confidence interval of 95 %. Table 4.3 summarizes the simulation parameter settings.

In order to evaluate the performance of the proposed approach, eight metrics were used and are described in detail below.

- *Overhead*: shows the number of beacon packets transmitted in the network by all vehicles during the simulation run;

[4]www.openstreetmap.org

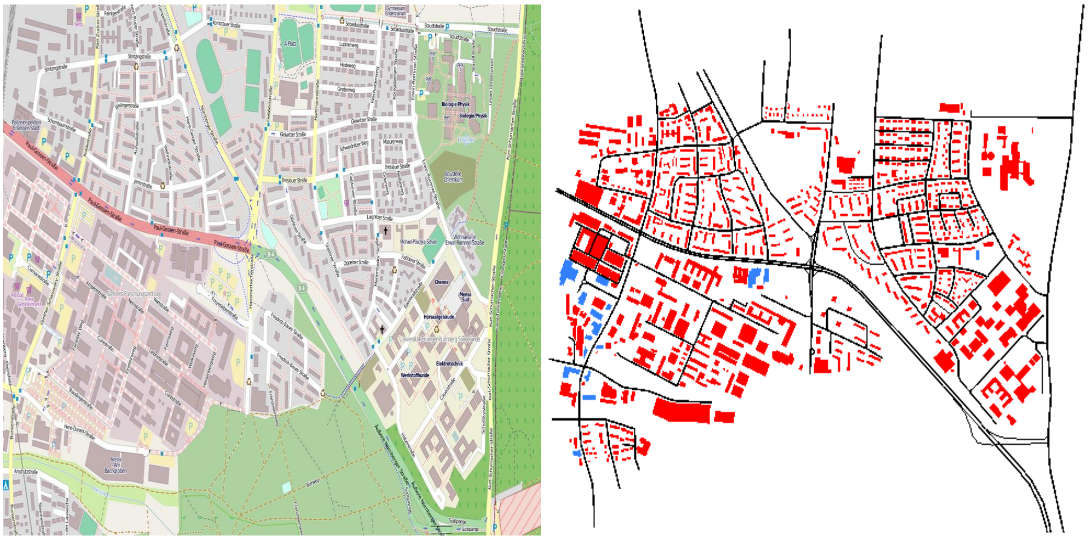


Figure 4.6: Map clipping from Erlangen, Germany. The figure on the left was imported from OSM and on the right represents the road topology used in our simulations.

Table 4.3: Simulation parameters.

Parameter	**Value**
Density of vehicles	40–150 vehicles/km^2
MAC layer	802.11p
Channel	178 (5.89 GHz)
Bandwidth	10 MHz
Transmission power	0.98 mW
Bitrate	6 Mbps
Sensitivity	-82 dBm
Transmission range	287 m
Beacon transmission frequency	1 Hz
Simulation time	350 s
Confidence interval	95 %

- *Beacon transmitted per vehicle*: gives the number of beacon packets transmitted per each vehicle during the simulation run;
- *Beacon received*: displays the number of beacon packets received per vehicle during the simulation run;
- *Total of lost packets*: is the sum of both RxTx (receive/transmit) and SNIR (signal to noise plus interference ratio) lost packets; the first one occurs due to the busy communication channel, whereas the second one occurs due to bit errors in received packets;
- *Channel busy ratio*: indicates the fraction of the time in which the channel is identified as busy;
- *Regression analysis*: is a set of statistical processes to estimate the linear relationships between two datasets;
- *Pearson correlation coefficient*: expresses the strength of a linear association between two datasets;
- *Window time*: points out the smallest window time under which there are no changes in the egocentric betweenness.

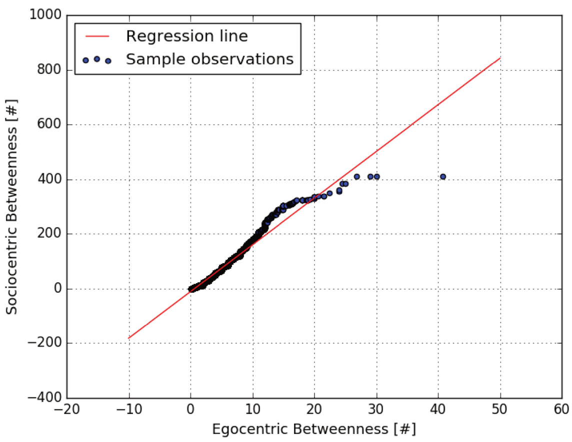
In order to provide a better understanding of our approach, results are compared to the ones obtained from the sociocentric betweenness approach. For this purpose, a dynamic graph was generated, with the aid of the Dynamic Graph Library [50], to perform the sociocentric betweenness calculation [24]. This library requires floating car data (FCD) as the input parameter. FCD is a method applied to gather traffic knowledge. In the sociocentric approach, all the vehicle network topology knowledge was used as input.

4.5.2 Simulation Results

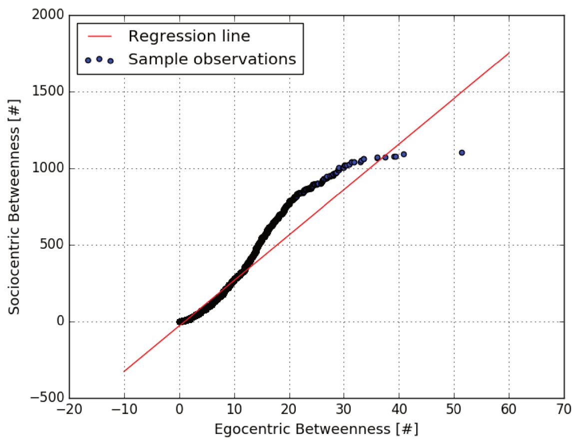
The first set of experiments investigated the correlation between egocentric and sociocentric betweenness scores in a VANET scenario. In other words, how accurate the results were when using only the local knowledge of the network topology to compute the betweenness score, instead of using global knowledge of the topology. The results of this approach are shown in the scatter diagram set in Figure 4.7, which compares the two approaches for each vehicle traffic density.

A scatter plot revealed the relationships between two variables (in our case, such variables were the sociocentric and the egocentric scores). The relationship between two variables is known as correlation. The higher the correlation between the two variables, the closer the sample observations will be to a straight line. If the sample observations go along a straight line (or regression line) from the origin to high x- and y-values, then the variables are assumed to have a positive correlation. Thus, it is possible to observe in Figure 4.7 that the egocentric and the sociocentric betweenness scores have a positive correlation.

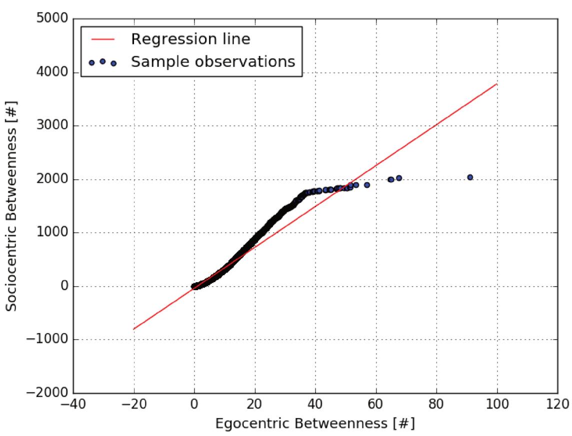
Figures 4.7(a), 4.7(b), 4.7(c), 4.7(d), and 4.7(e) show the scatterplots for densities of 40, 60, 80, 100 and 150 vehicles/km^2, respectively. As can be seen in these figures, these two measures do not provide the same betweenness scores, as expected. The egocentric betweenness



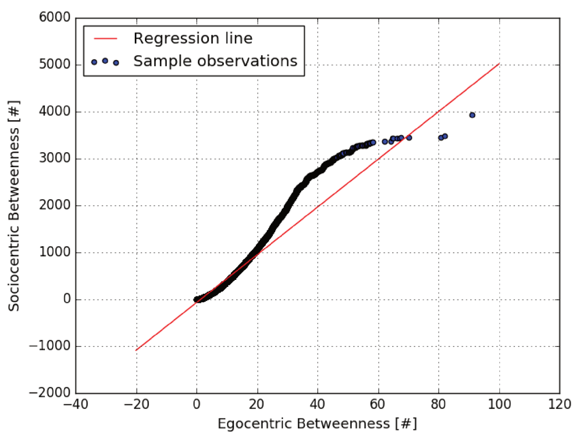
(a) 40 vehicles/km^2.



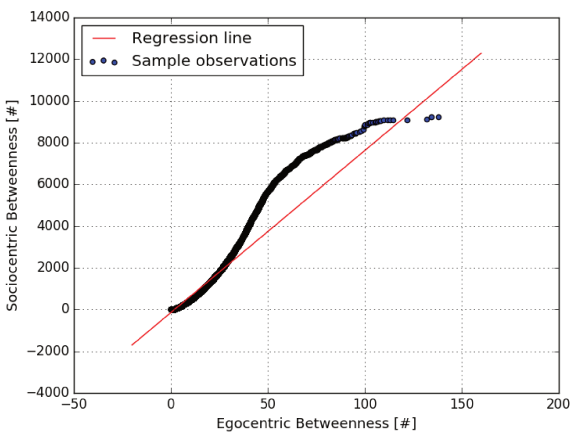
(b) 60 vehicles/km^2.



(c) 80 vehicles/km^2.



(d) 100 vehicles/km^2.



(e) 150 vehicles/km^2.

Figure 4.7: Scatterplot of sociocentric vs. egocentric betweenness for each vehicle traffic density.

scores (x-axis) were smaller than the sociocentric betweenness scores (y-axis). This can be explained by the fact that in the ego-network topology, the maximal geodesic distance between nodes was two, and this limitation did not apply to the sociocentric betweenness. On the other hand, through the analysis of the figures, the egocentric and the sociocentric betweenness scores have demonstrated a high degree of similarity regarding the ranking of nodes. This similarity can be confirmed in Table 4.4. The table depicts the Pearson correlation coefficient (PCC) between the egocentric and the sociocentric betweenness approaches. The presented values ranged from 0.953–0.983 (where 1.0 represents a perfect linear relationship between the two datasets analysed), in all traffic densities.

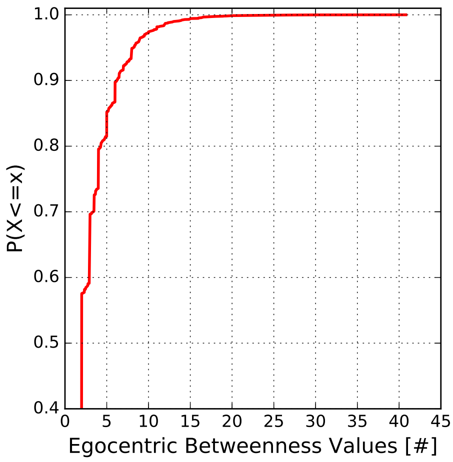
Table 4.4: Pearson correlation coefficient (PCC) of egocentric and sociocentric betweenness.

Density (vehicles/km^2)	**PCC**
40	0.983
60	0.962
80	0.971
100	0.964
150	0.953

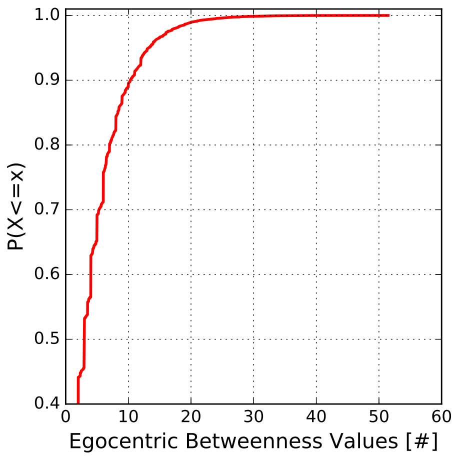
Lastly, it is possible to notice that some scores lie relatively away from the regression line (red line). Even so, there is a clear positive relationship between the two betweenness measures in VANETs.

Figures 4.8 and 4.9 depict the cumulative distribution function (CDF), in each vehicle traffic density, of the egocentric betweenness scores and the number of one-hop neighbours, respectively. The CDF measure is an interesting way of observing the behaviour of analysed variables. As can be observed in Figure 4.8, the egocentric betweenness scores fluctuate in the same range as in Figure 4.7, according to the vehicle traffic density. Another important information is to analyse the distribution of these scores. It is possible to observe that 90 % of the samples, for densities of 40, 60, 80, 100 and 150 vehicles/km^2, were lower than 7, 11, 16, 18 and 30, respectively. In other words, these scores were close to the regression line (red line of Figure 4.7), i.e., 90 % of the samples of the two variables had a high correlation. The same distribution analysis was performed for the number of one-hop neighbours, as shown in Figure 4.9. In this example, it is possible to notice that 90 % of the samples, for densities of 40, 60, 80, 100 and 150 vehicles/km^2, were lower than 7, 9, 12, 14 and 21 neighbours, respectively.

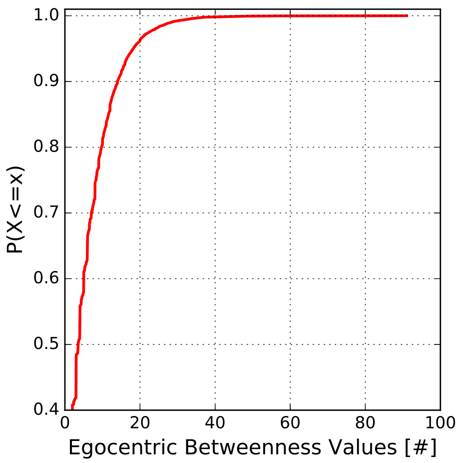
The relationship between the egocentric betweenness scores and the number of one-hop neighbours is depicted in Figure 4.10. This figure shows the average egocentric betweenness score (red line) and the average number of one-hop neighbours (blue line) for all vehicle traffic densities. Therefore, it summarizes all the information presented in the two sets of Figures 4.8 and 4.9. The observed behaviour of both measures is in agreement: as the traffic density increased, the number of vehicles in the vicinity and the egocentric betweenness scores also increased. For instance, in a low traffic density (40 vehicles/km^2), the egocentric betweenness score was around 2.5, and the number of one-hop neighbours was around 3.9, on average. On the other hand, in a high traffic density (150 vehicles/km^2), the egocentric betweenness score and the number of one-hop neighbours were around 12.2 and 9.8 on average, respectively.



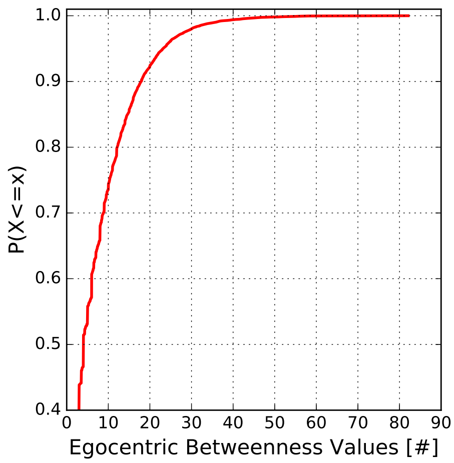
(a) 40 vehicles/km^2.



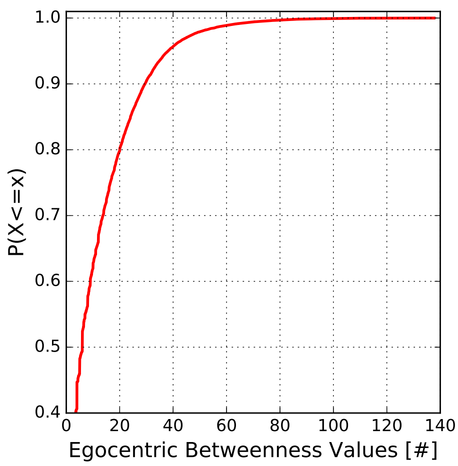
(b) 60 vehicles/km^2.



(c) 80 vehicles/km^2.

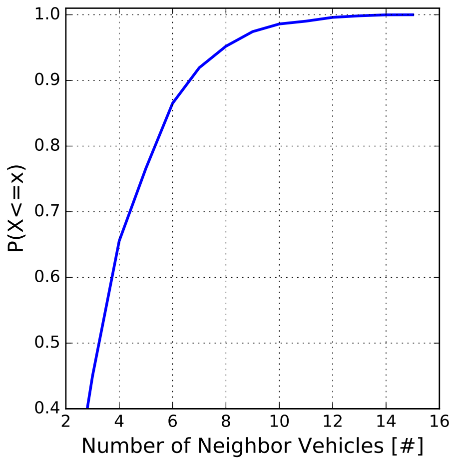


(d) 100 vehicles/km^2.

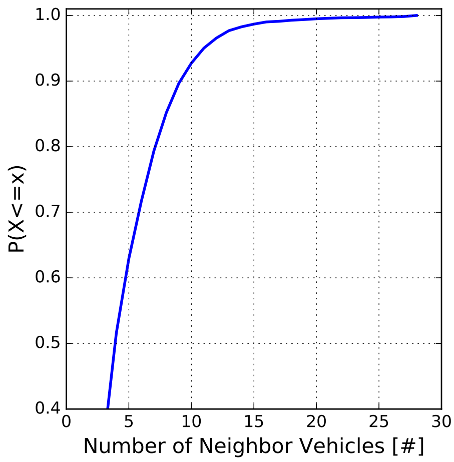


(e) 150 vehicles/km^2.

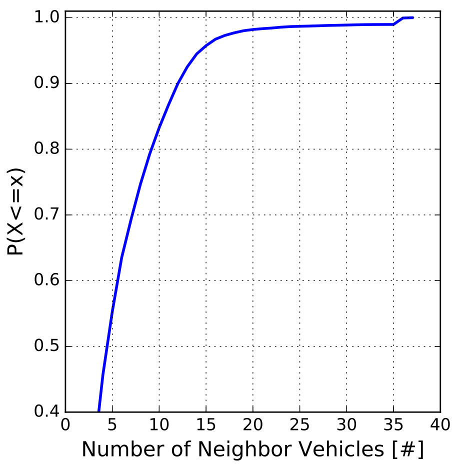
Figure 4.8: CDF of the egocentric betweenness scores in relation to the vehicle traffic densities.



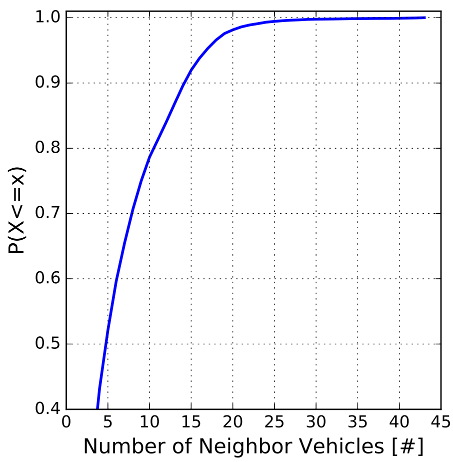
(a) 40 vehicles/km^2.



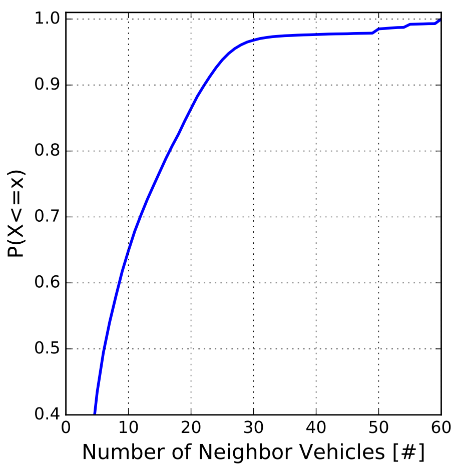
(b) 60 vehicles/km^2.



(c) 80 vehicles/km^2.



(d) 100 vehicles/km^2.



(e) 150 vehicles/km^2.

Figure 4.9: CDF of the number of one-hop neighbours in relation to the vehicle traffic densities.

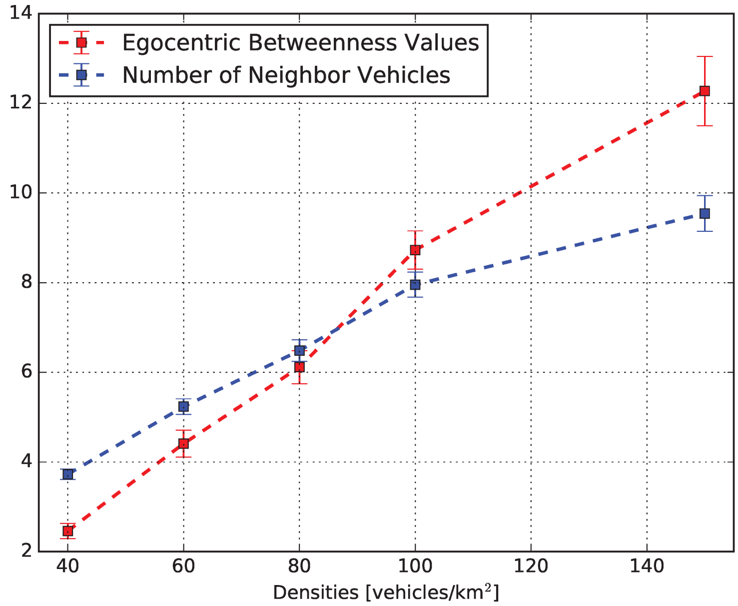


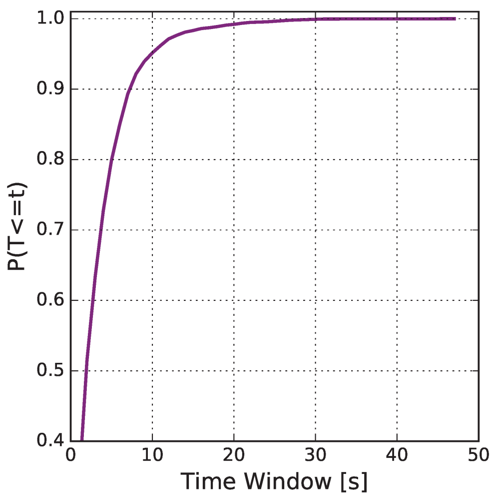
Figure 4.10: The relationship between the egocentric betweenness score and the number of one-hop neighbours.

Another important analysis that can be performed in the egocentric betweenness measure is the calculation of the smallest time window duration in which there were no changes to the egocentric betweenness scores in relation to the vehicle traffic densities. The CDF of the time window duration in each traffic density is shown in the Figure 4.11 set. In this case, it is possible to notice that 90 % of the samples, for densities of 40, 60, 80, 100 and 150 vehicles/km^2, have time window durations that were lower than 9, 8, 7, 6 and 5 s, respectively.

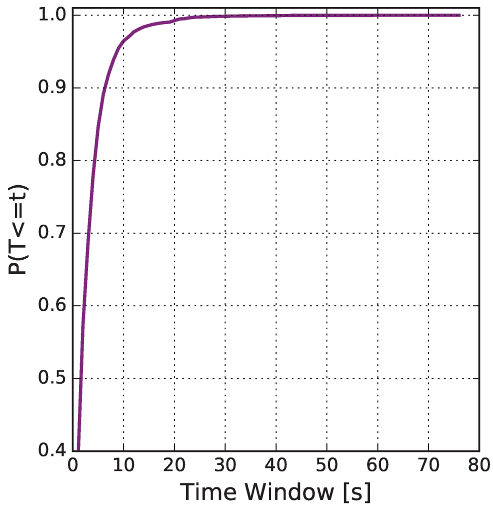
Figure 4.12 shows the average time window duration in each traffic density. This metric is important in vehicular networks because many applications rely on a stable period of connectivity between nodes [34, 125, 140]. The figure shows that as the traffic increased, the average time window duration decreased, until reaching a stable plateau. For example, when the density was 40 vehicles/km^2, the average time window was around 3.55 s. When the density increased, the average time window rapidly decreased until reaching the plateau at 2.95 s, for the cases of 100 vehicles/km^2 and 150 vehicles/km^2. For many distributed applications, the real-time content distribution within the area of interest was less than 2 s [34, 125]. Therefore, the average time window reached into all densities of the simulations was sufficient to meet the requirements of such applications. The behaviour depicted in the picture confirmed our expectation: as traffic increased, the trend was that the list of one-hop neighbours fluctuated rapidly over time. One point worth highlighting is that the time can vary according to the scenario used, as well as the mobility model and the vehicle traffic densities applied.

The second set of experiments consisted of performing the analysis of the network traffic. This analysis is needed to demonstrate the scalability of our proposed approach, since the periodic exchange of beacon packets, to stay aware of the one-hop neighbour topology, was carried out by means of vehicle-to-vehicle communications. The experiment results of the metrics such as overhead, beacon transmitted per vehicle, beacon received and total lost packets are depicted in Figure 4.13. The detailed results of each one of these metrics are given below.

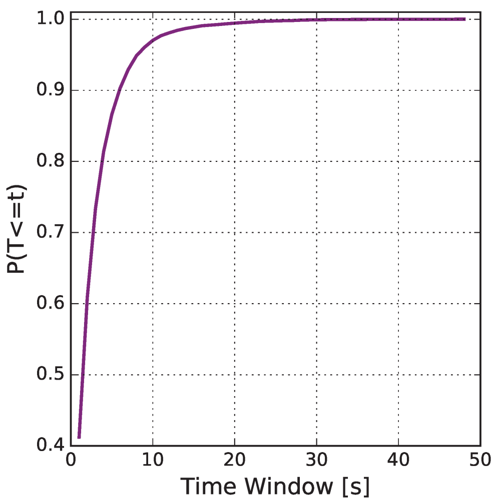
Figure 4.13(a) provides a macroscopic view of total number of the beacon packets transmit-



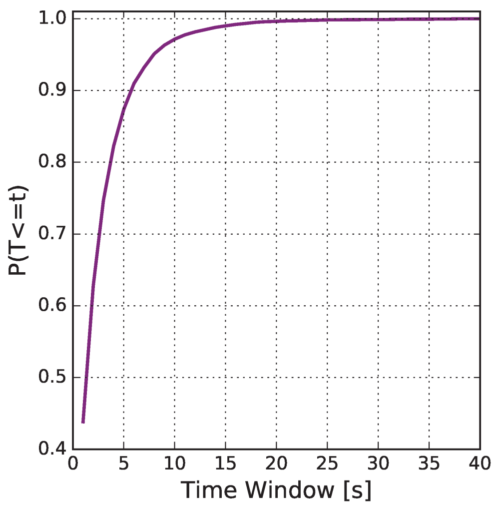
(a) 40 vehicles/km^2.



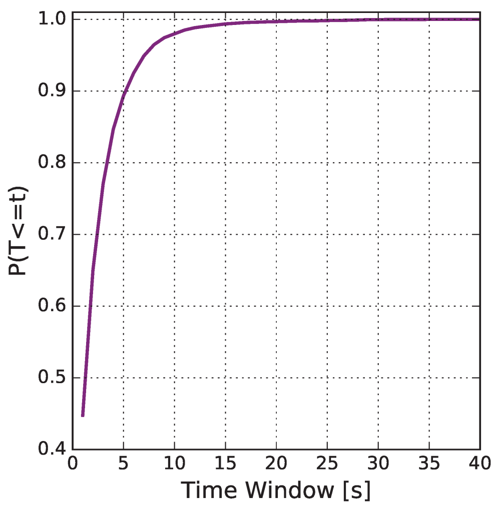
(b) 60 vehicles/km^2.



(c) 80 vehicles/km^2.



(d) 100 vehicles/km^2.



(e) 150 vehicles/km^2.

Figure 4.11: CDF of the time window duration in which there were no changes to the egocentric betweenness score in relation to the vehicle traffic densities.

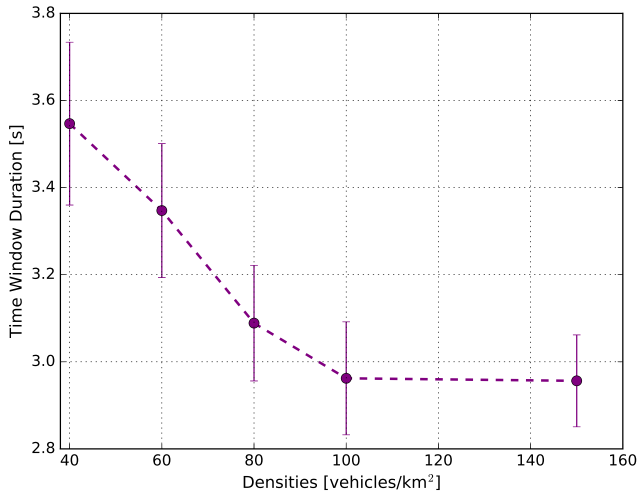
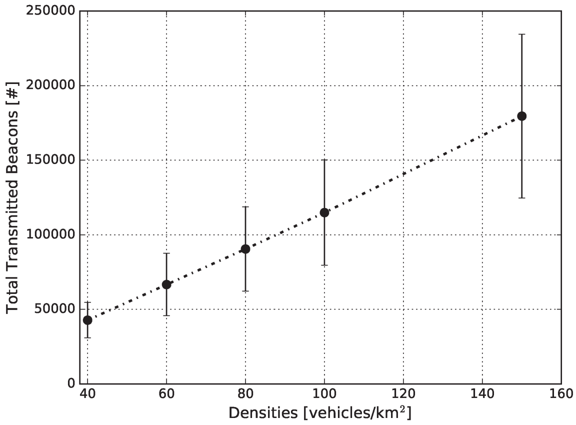


Figure 4.12: Average time window duration in which there were no changes to the egocentric betweenness scores.

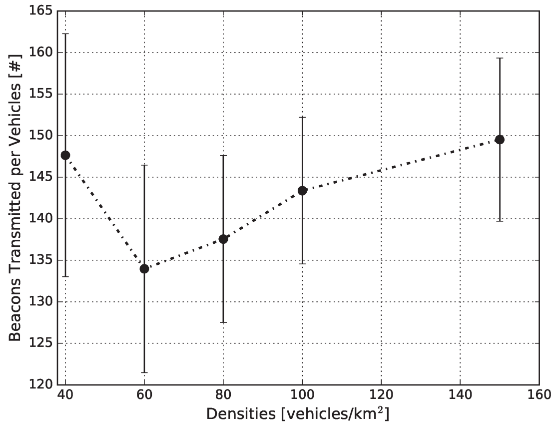
ted in each traffic density. For instance, in densities of 40, 60, 80, 100 and 150 vehicles/km^2, we had on average 49,000, 70,000, 90,000, 120,000 and 180,000 transmitted beacon packets, respectively. As can be seen, the beacon overhead increased linearly as a function of the traffic density, as expected. This expectation was well founded since as the density of vehicles increased, the higher the transmission rate of beacon packets into the network would be.

The microscopic view is depicted in Figure 4.13(b), which shows the average number of beacon packets transmitted by each vehicle in each traffic density. When the experimental scenario had a density of 40 vehicles/km^2, each vehicle, on average, transmitted around 148 beacons during the simulation time; while, in the scenarios with 60 and 80 vehicles/km^2, on average, 134 and 138 beacons were transmitted, respectively. For 100 and 150 vehicles/km^2, there were, on average, 144 and 150 beacons transmitted by each vehicle, respectively. It is easy to see that the number of beacon packets transmitted, for each vehicle, is directly related to its trip time during the simulation time. With that in mind, Figure 4.14 depicts the average trip time of the vehicles during the simulation. It is possible to observe that in both of the aforementioned figures, the same behaviour appears in all the vehicle traffic densities. For example, in Figure 4.14, for the scenarios with 40 and 150 vehicles/km^2, the average trip times are higher than all other evaluated scenarios, reaching 2.8 and 2.55 min, respectively. On the other hand, the scenario with 60 vehicles/km^2 presented the lowest average (2.0 min).

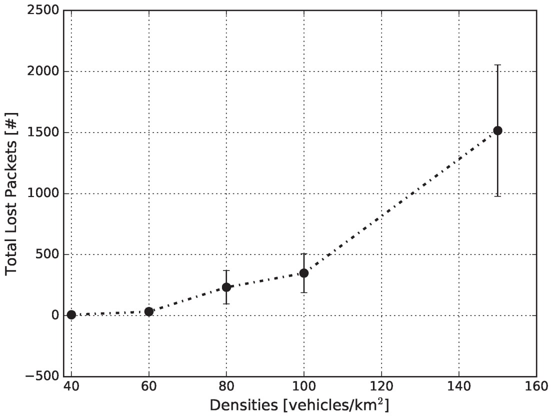
Figure 4.13(c) depicts the total number of beacon packets lost either by the fact that the communication channel was busy, or by errors in the received packets. As can be observed, the low densities (40 and 60 vehicles/km^2) presented a minimum packet loss rate. As the vehicle traffic density increased up to 150 vehicles/km^2, the total number of packets lost also increased. The observed behaviour was directly related to the channel busy ratio. Taking this into account, Figure 4.15 shows the average channel busy ratio for each vehicle traffic density. As the simulation time was set to 350 s, the calculation of the total busy time was nothing more than the channel busy ratio multiplied by the simulation time. In our case, for densities of 40 and 60 vehicles/km^2, the channel was busy for the shortest time, and as the density increased, the



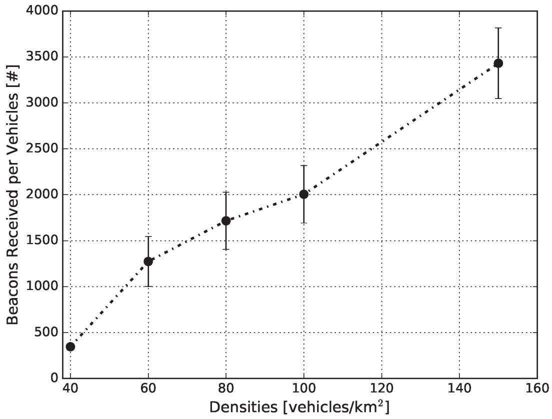
(a) Total of transmitted beacons vs. Densities.



(b) Beacons transmitted per vehicle vs. Densities.



(c) Total of lost packets vs. Densities.



(d) Beacons received per vehicle vs. Densities.

Figure 4.13: Performance evaluation of the network under different traffic densities.

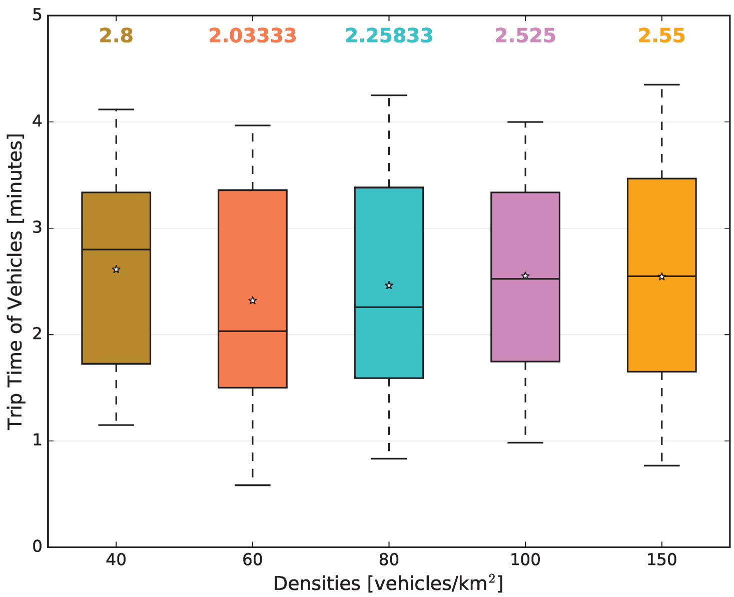


Figure 4.14: Average trip time of vehicles vs. densities.

average time also increased. Even in the density of 150 vehicles/km^2, a maximum of 35 % of channel availability was consumed. These results show that the beacon transmission frequency of 1 Hz was suitable, for this scenario, together with the mobility model applied, due to low channel utilization.

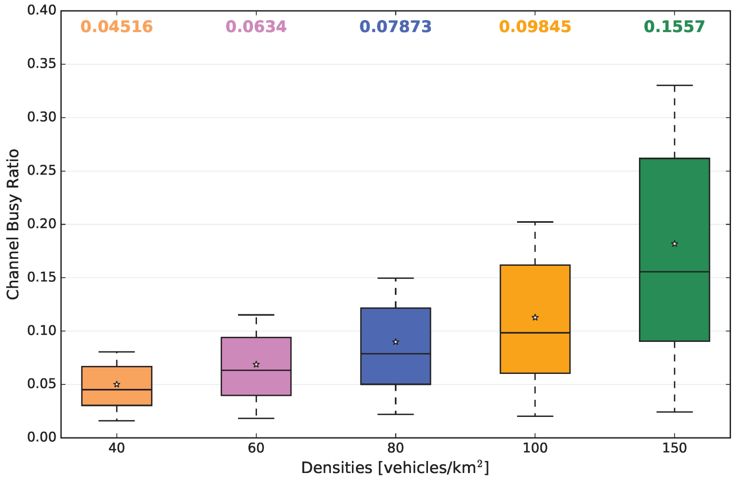


Figure 4.15: Impact on channel busy ratio vs. densities.

The number of beacon packets received per vehicle is depicted in Figure 4.13(d). This metric, combined with the channel busy ratio (Figure 4.15), can indicate if the beacon transmission frequency is adequate or not. In the same way as the total number of beacon packets transmitted, the number of beacon packets received also increased linearly as a function of the vehicle traffic density. For instance, for densities of 40, 60, 80, 100 and 150 vehicles/km^2, there were, on average, 480, 1300, 1700, 2000 and 3450 beacon packets received per vehicle, respectively. As mentioned before, the channel utilization in our approach was low; this confirmed, once again, that the beacon transmission frequency of 1 Hz was proper.

4.6 Final Remarks

In this chapter, a distributed approach to calculating egocentric betweenness scores, in VANETs, was presented. To this end, each vehicle regularly broadcasts one-hop messages about its local information among surrounding vehicles. The proposed approach only uses the locally available information to compute the egocentric betweenness score without the need for information of the entire network topology.

A set of simulation experiments has been carried out in a real urban center area in order to investigate the performance comparison of our egocentric approach against the traditional sociocentric approach in different vehicle traffic densities. The main contribution here is the demonstration that the egocentric approach has a greater similarity regarding the ranking of nodes concerning the sociocentric approach. Besides, the channel utilization in our approach was low; this confirmed that the beacon transmission frequency of 1 Hz was proper. It is important to highlight those solutions that employ the egocentric betweenness measure; it is the ranking of the nodes that matters most, rather than their absolute scores.

Chapter 5

Information Management and Knowledge Distribution in VANETs

5.1 Introduction

For many ITS applications that use VANETs, the constant sharing of local information, with one-hop communication neighbors, is essential to create awareness about vehicle traffic conditions [91, 143, 145]. This type of sharing is well-known as beaconing, and most often, the exchanges occur in the control channel with a transmission frequency generally between 1Hz and 10Hz [130]. The default information contained in the beacon package includes vehicle identification, current vehicle position, average speed, the direction of travel, among others [29]. On the other hand, the service channels are used to share all other data needed by the applications.

Several ITS that deal with local information management and knowledge distribution about vehicle traffic conditions have been proposed [91, 143, 145]. This type of system extracts knowledge, for instance, about the traffic condition of a given road, by processing the aggregated local information received from the neighbor vehicles. However, many proposed systems have the same shortcoming, the absence of a *vehicle selection mechanism* to carry out the tasks of information aggregation and knowledge generation. Without the selection mechanism, all vehicles would perform such tasks resulting in a highly redundant traffic of knowledge. In addition, other systems [91, 143] do not apply any broadcast suppression mechanism during knowledge distribution, increasing even further the network overhead.

In order to overcome the above-cited limitation, we propose TRUSTed, a dis**TR**ib**U**ted **S**ys**T**em for information management and knowl**ed**ge distribution. By means of beaconing, TRUSTed collects the local information needed to apply the vehicle selection mechanism. The result is the selection, within a subset of vehicles, of the most relevant ones in a given moment to carry out the tasks of information aggregation and knowledge generation. The *relevance* is defined as the importance of a vehicle in relation to the information flows that pass through it. In other words, it defines how important is the intermediate vehicle for the information flow continuity in the network. One of the advantages of such a mechanism is the use of the local information (egocentric measure) to perform the necessary calculation. Beyond this advantage, the work of [7] confirmed that the egocentric betweenness measure, in a highly dynamic topology, has a high correlation with the sociocentric betweenness measure. Last but not least, a broadcast suppression mechanism was applied to avoid the redundant traffic of knowledge. The goal of this case study is to prove that the mechanism can reduce bandwidth consumption, taking into account the challenges of VANETs.

The remainder of this chapter presents a brief survey of the related work. After that, the proposed solution is presented in Section 5.3. Some numerical results and analysis are given in Section 5.4. Finally, Section 5.5 concludes this chapter.

5.2 Literature Review

All proposals presented, in this section, employ a periodic exchange of local information, between one-hop communication neighbors, this allows them to create the local knowledge base. In addition, they were designed to operate only with vehicle-to-vehicle communication technology.

The work of [91] has proposed a probabilistic aggregation for knowledge generation. This approach uses a hierarchical aggregation technique called soft-state sketches. This technique is an extension of Flajolet–Martin sketches [60]. The fundamental characteristic of this approach lies in the fact that the aggregate information does not have a specific value of the monitored place, for instance, an average speed of a determined road. The aggregated information has, instead, a probabilistic value. The main benefit of this approach is the capability to combine the aggregated values, with the same context, for knowledge generation. However, this work lacks a vehicle selection mechanism to perform knowledge generation task. Therefore, all vehicles would perform such a task, thereby generating highly redundant traffic of knowledge.

Yu et al. [143] have proposed an adaptive forwarding delay control, named Catch-up, to gather aggregated local information from different sources for knowledge generation. To this goal, the forwarding speed of nearby information is dynamically adjusted. Thereby, each aggregate information can have one of the two types of adaptive delays, RUN (short) or WALK (long). The delay calculation is based on a distributed learning algorithm, in which each vehicle learns by means of local information. The main advantage of catch-up is the use of an adaptive forwarding delay for knowledge generation, as well as probabilistic aggregation. However, a disadvantage of this approach is that all vehicles can act as an information aggregator and knowledge generator, which can incur network overhead.

Another solution is the data aggregation algorithm by restricting forwarders (DARF) [145]. This algorithm concentrates mainly on the selection of the vehicles that will continue the knowledge forwarding process, which was generated in the aggregation step. In order to do that, each vehicle receives one of the two available labels (forwarder or non-forwarder) according to the neighbourhood labels. As the name says, each label defines whether the vehicles will be a forwarder, or not, of the knowledge. The vehicle will be a non-forwarder if there is a forwarding vehicle immediately in front of and behind it. One of the advantages of DARF is the broadcast suppression mechanism applied during the knowledge distribution process, which is not applied in the above-mentioned works. However, it is possible to notice that there is no vehicle selection mechanism to aggregate local information and generate the knowledge. In this way, it allows highly redundant traffic of knowledge in the network.

All systems presented here have the same shortcoming, the absence of a vehicle selection mechanism to carry out the tasks of information aggregation and knowledge generation. Without the selection mechanism, all the vehicles would perform such tasks, resulting in highly redundant traffic of knowledge in the network. This, consequently, will lead to high bandwidth consumption. Thus, the use of vehicle selection mechanism contributes to improving this issue, which has not yet been addressed in the literature.

5.3 TRUSTed

TRUSTed is a distributed system for information management and knowledge distribution related to vehicle traffic conditions in VANETs. One of the main challenges of this type of system, due to the highly dynamic topology, is the selection of the most relevant vehicle, within a subset of vehicles, to perform the tasks of information aggregation and/or knowledge generation. If a vehicle is not selected, all of them could carry out such tasks, this can overload the network with highly redundant traffic of knowledge. With this in mind, the egocentric betweenness measure was applied to select the vehicle that will carry out above-mentioned tasks.

5.3.1 Vehicle Selection Mechanism

The egocentric measure was chosen because it requires only the available local information (one-hop neighbors) to find the most relevance vehicle. This relevance is based on the information flow passing through it. The calculation of this measure is depicted in Section 4.4. In addition to egocentric betweenness measure, a radio propagation model, the two-rays ground reflected, was applied. The aim is to improve the process of data propagation, among vehicles, through a path with minimum interference in inter-vehicle communication.

$$L_{TRI}[dB] = 20log(4\pi\frac{d}{\lambda}|1 + \Gamma \exp^{\varphi}|^{-1}) \tag{5.1}$$

where λ is the wavelength, d is the Euclidean distance between two vehicles, Γ is the reflection coefficient and φ is the interfering rays. The interfering rays are given by:

$$\varphi = 2\pi\frac{d_{los} - d_{ref}}{\lambda}, \begin{cases} d_{los} = \sqrt{d^2 + (h_t - h_r)^2} \\ d_{ref} = \sqrt{d^2 + (h_t + h_r)^2} \end{cases} \tag{5.2}$$

where d_{los} and d_{ref} correspond to the line-of-sight distance and reflected path between the transmitting and receiving antennas, respectively. h_t and h_r represent the transmitter and the receiver antenna heights, respectively. In this study, the same heights applied in the test bed implementation of Sommer et al.'s work were used [129] (h_t = h_r = 149.5 cm). The value of λ was fixed at 0.051 m according to IEEE 802.11p [73]. Lastly, the reflection coefficient can be calculated as:

$$\Gamma = \frac{\sin\theta_i - \sqrt{\varepsilon - \cos\theta_i}}{\sin\theta_i + \sqrt{\varepsilon - \cos\theta_i}}, \begin{cases} \sin\theta_i = \frac{h_t + h_r}{d_{ref}} \\ \cos\theta_i = \frac{d}{d_{ref}} \end{cases} \tag{5.3}$$

where ε is the relative permittivity of the ground and θ is the angle between the ground and the reflected ray.

5.3.2 Knowledge Generation Process and Distribution

Our proposed solution periodically shares the local information, between one-hop neighbours, through beacon packets to create the local knowledge base. In order to do that, two more pieces of information were added in the beacon package: the current EBM score and the aggregated information.

The local knowledge base is built by aggregating the local information received from the neighbourhood, as well as the calculation of the weight of the roads. Once the local knowledge base is created, the next step is to share it with the most relevant neighbour vehicle, this is performed by following the selection criterion presented in Section 5.3.1.

The following representation shows an example of the fusion of two aggregated values: $A_r := \partial(A_1, A_2)$, where ∂ is the aggregation function that has two input values (A_1 and A_2). These values are combined, resulting in a new aggregated value (A_r). As the main goal of the proposed study is the generation and distribution of knowledge about the traffic condition, the aggregation function is given as follows:

$$v_{agg_i}^{avg} = \frac{v_1 n_1 + v_2 n_2}{n_1 + n_2} \tag{5.4}$$

where $v_{agg_i}^{avg}$ represents the aggregate average speed of a given road i. The parameters v_1 and v_2 are the two input values from i. n_i indicates the amount of information that contributed to the generation of the new aggregated value. Thereby, the weight of the road i (w_i) is calculated as follows:

$$w_i = 1 - \frac{v_{agg_i}^{avg}}{v_{spe_i}^{max}}, \begin{cases} w_i : \text{weight of road } i \\ v_{agg_i}^{avg} : \text{aggregate average speed of road } i \\ v_{spe_i}^{max} : \text{maximum speed of road } i \end{cases} \tag{5.5}$$

After aggregating all the local information, the vehicle with the highest EBM score classifies the weight of the roads according to Table 5.1. The levels of service and traffic classification were based on the Highway Capacity Manual (HCM) [51].

Table 5.1: Level of service and traffic classification [51].

Level of Service	**Traffic Classification**	p_i
A	Free flow	(0.0∼0.33]
B	Reasonably free flow	(0.33∼0.4]
C	Stable flow	(0.4∼0.5]
D	Approaching unstable flow	(0.5∼0.7]
E	Unstable flow	(0.7∼0.9]
F	Forced or breakdown flow	(0.9∼1.0]

After the classification step, if an event is identified (in our case, roads with the level of service D, E or F), a message (also known as knowledge), containing the identification of the roads in question is generated. Thereby, the knowledge distribution process in the service channel is started.

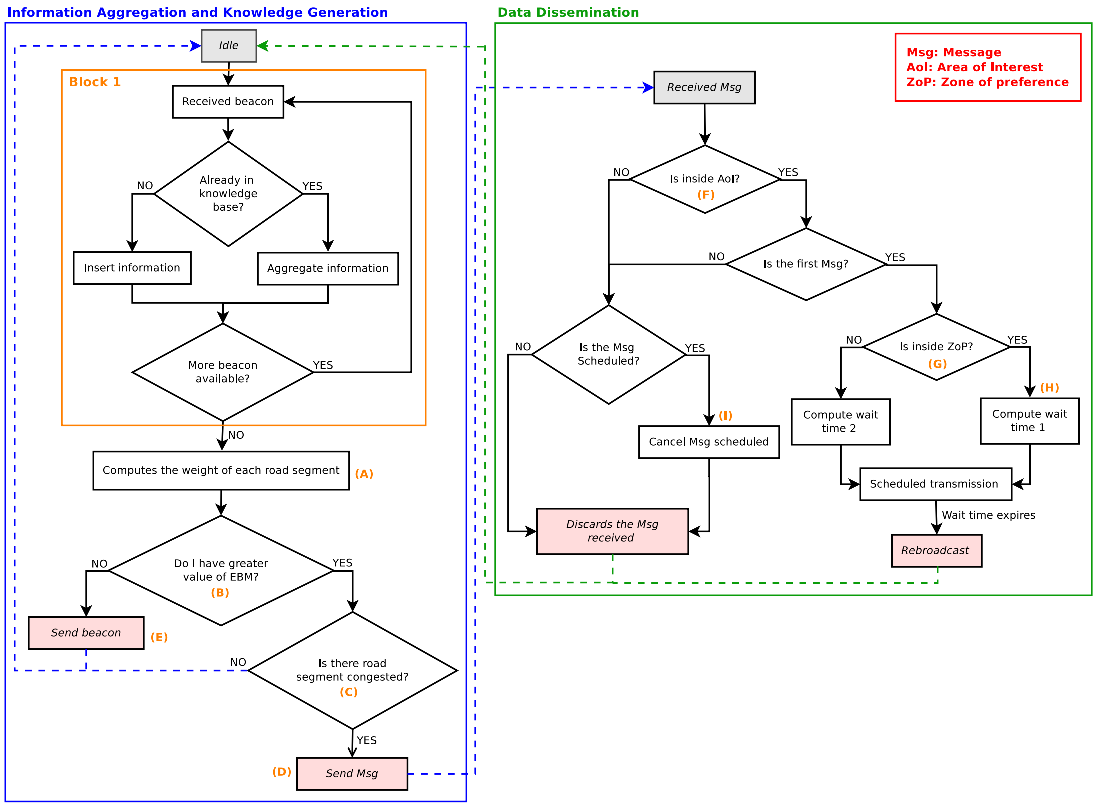


Figure 5.1: Operation flowchart of the proposed solution.

Figure 5.1 shows the operation flowchart of the proposed solution. The flowchart is divided into two phases. The first one is the information aggregation and knowledge generation, and the second is the data dissemination. In the first phase, every time the vehicle receives the local information, it either inserts or aggregates the local information into the local knowledge base (Block 1). In the next step, it calculates the weight of roads according to Equation (5.5) (Legend (A)). After this step, the vehicle with the highest EBM score (Legend (B)) classifies the weight of roads according to Table 5.1 (Legend (C)). During this process, if the selected vehicle detects some congested traffic flow, the knowledge is generated and distributed in the network (Legend (D)). On the other hand, if the vehicle does not have the highest EBM score, it selects the next most relevant vehicle and sends the aggregated local information to it (Legend (E)). The second phase (data dissemination), is responsible for informing vehicles that are inside an area of interest (AoI - Legend (F)) according to the application requirements. In addition, it also avoids the broadcast storm problem during the knowledge distribution process. Basically, to avoid this problem, a forwarder candidate suppresses the rebroadcast of low-priority candidates forwarders [148]. For this purpose, every time that a vehicle receives knowledge to be distributed, it checks if it is within the zone of preference [10] (Legend (G)), and if so, it transmits first (Legend (H)) because it has the shortest waiting time. Due to the broadcast suppression mechanism implemented (zone of preference), as soon as the neighbouring vehicles outside the zone of preference receive the same scheduled knowledge, they cancel the retransmission (Legend (I)), thereby avoiding the traffic of redundant knowledge in the network.

5.4 Evaluation Metrics

Four metrics were applied in order to evaluate the performance of the proposed solution:

- *Overhead*: measures the total amount of transmitted messages in the network;
- *Collision*: estimates the total number of packet collisions during message transmission;
- *Delay*: measures the time spent in delivering the messages to vehicles;
- *Coverage*: estimates the percentage of messages delivered to the vehicles that are within the scenario.

The simulation parameters used here are the same ones of Table 4.3, except the density of vehicles, which in this case ranges from 100–300 vehicles/km^2. Moreover, AoI has been applied with a 1-km radius from the congestion point. It is worth mentioning that the scenario used is the same as of Figure 4.6.

5.4.1 Performance Analysis and Discussion

Figure 5.2 presents the performance results of all solutions analyzed using the coverage metric. The Probabilistic solution displays the lowest coverage, reaching an average of 80 %, for all analyzed densities. These results can be justified due to the network overhead, which is caused because all vehicles perform the tasks of information aggregation, generation, and distribution of the knowledge (Figure 5.3 and Figure 5.5). In addition, during the process of knowledge distribution none broadcast suppression mechanism is applied, thus, resulting in a highly redundant traffic of knowledge, as shown in Figure 5.3. Because of this, it is possible to observe a high rate of packet collisions in the network (Figure 5.5). It is also evident the long delays in the delivery of knowledge, compared to the other systems considered (Figure 5.4). We can see a slight drop in the coverage rate as the vehicle density increases. This is due to the fact of the high network overhead and the high collision rate.

The other solution analyzed is the Catch-up system. The main strategy of this system is the insertion of an adaptive delay in the message forwarding process. This allows increasing the probability of the meeting of the aggregated information. This approach was able to decrease the total number of messages transmitted and consequently, the collisions, as shown in the Figures 5.3 and 5.5. For this reason, Catch-up achieves better results when, compared to the Probabilistic system. It was able to reduce, on average, 10 % of both transmitted messages and packet collisions. In addition to that, it increased the coverage by 5 % (Figure 5.2). In both, Probabilistic and Catch-up, there is a slight drop in the coverage rate as the vehicle density raises. In addition to this, the Catch-up system still has a higher knowledge transmission rate and packet collisions. It is known that both Probabilistic and Catch-up do not use any type of selection mechanism to chose the most relevant vehicle to perform the tasks of information aggregation, generation, and distribution of knowledge. The lack of such mechanism is translated in the delays for both systems when compared to DARF and TRUSTed. This situation is depicted in Figure 5.4.

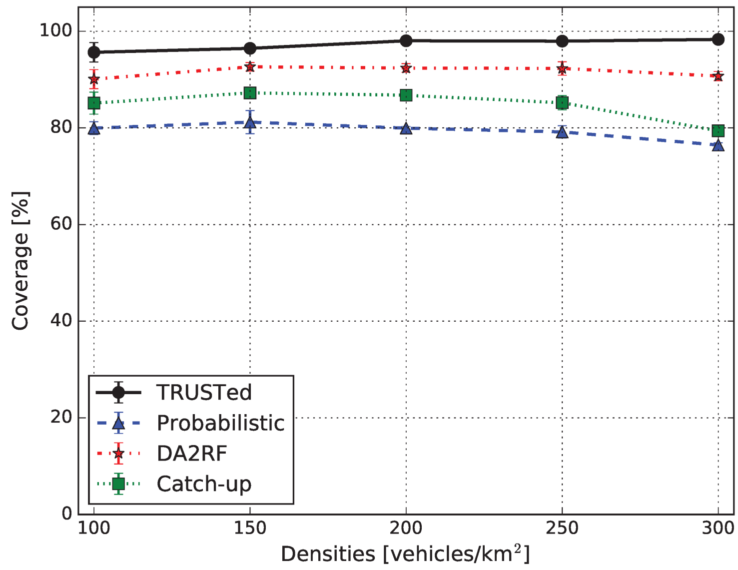


Figure 5.2: Coverage.

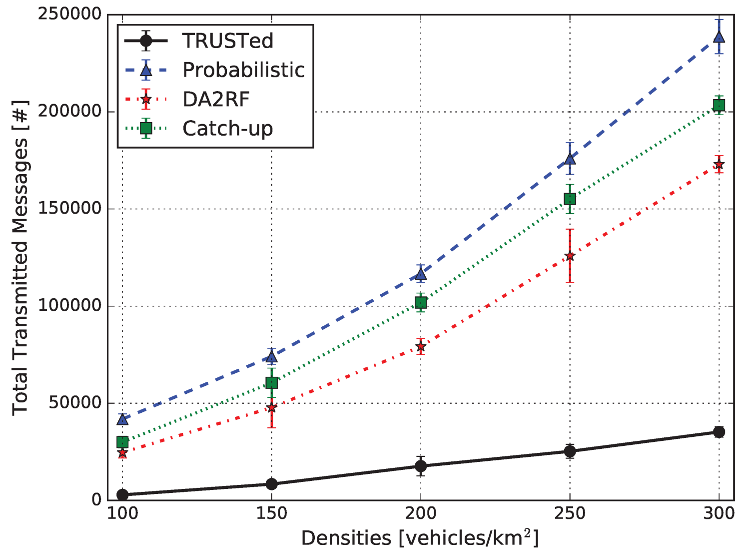


Figure 5.3: Total of transmitted messages.

The DA2RF system employs a broadcast suppression mechanism in the knowledge forwarding process.This approach, as shown in Figure 5.2, improves the coverage rate by 18 % and 15 % when compared to Probabilistic and Catch-up, respectively. By applying the suppression mechanism, it is possible to clearly see a decrease in the total number of the messages transmitted (Figure 5.3). On average, was reached a reduction of 30 % in comparison to the Probabilistic system, and 20 % fewer messages when compared to Catch-up. The same tendency was observed in regards to the packet collisions rate (Figure 5.5). On average there was a reduction of 30 % and 25 %, compared to Probabilistic and Catch-up, respectively. It is important to notice that DA2RF is implemented only with the broadcast suppression mechanism and does not have any selection mechanism. Because of this, it still introduces a delay very close to the other previously analyzed systems, as depicted in Figure 5.4.

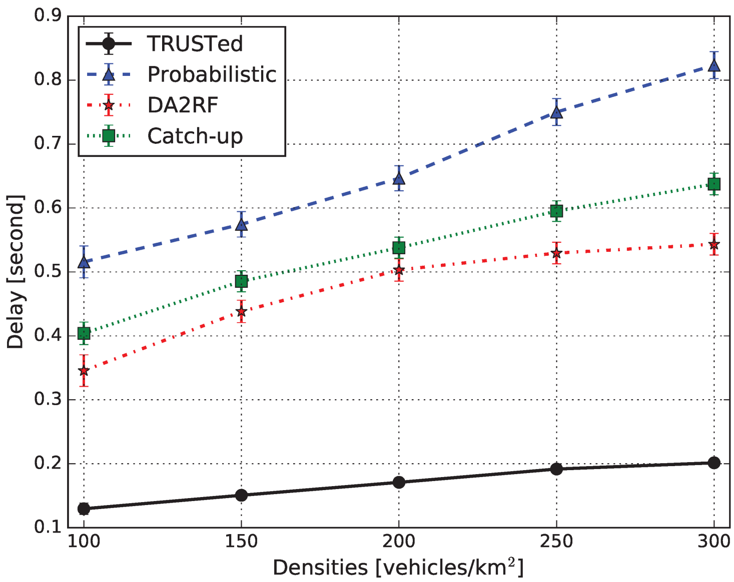


Figure 5.4: Delay.

Finally, the proposed TRUSTed system applies the egocentric betweenness measure to perform the selection of the most relevant vehicle to carry out the information aggregation and knowledge generation. In addition, it also applies the broadcast suppression mechanism in the knowledge distribution process. This combination enables it to outperform all other systems in all the metrics evaluated. TRUSTed significantly reduces the total number of messages transmitted in the network, with an average decrease of more than 85 % in comparison to Probabilistic, as well as 80 % and 70 % compared to Catch-up and DA2RF, respectively (Figure 5.3). As a consequence of this reduction, the knowledge generated can reach a larger number of vehicles in all densities analyzed, resulting in a higher coverage rate, close to 98 %, on average, as shown in Figure 5.2. Furthermore, the broadcast suppression mechanism implemented has helped reduce the number of packet collisions (Figure 5.5). The average reduction reached more than 75 %, 70 %, and 50 % compared to Probabilistic, Catch-up, and DA2RF, respectively. At the end, the TRUSTed system also presented the lowest average delay, among all systems analyzed, being around of 0.15 seconds (Figure 5.4).

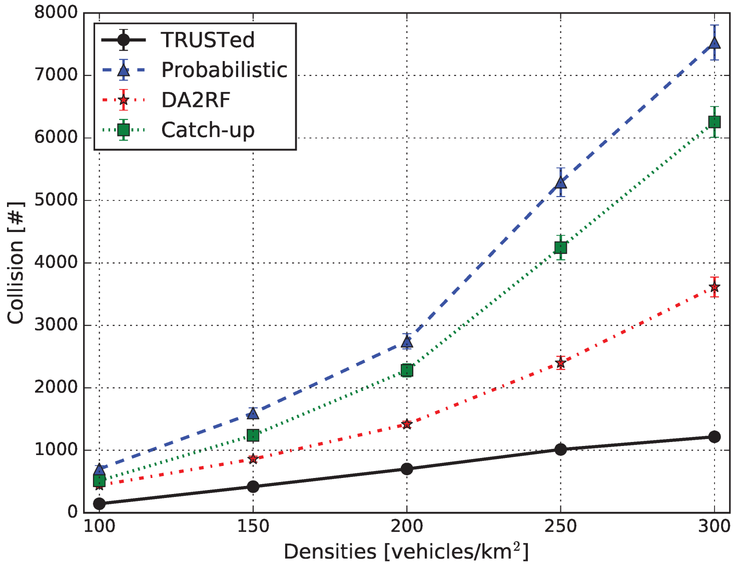


Figure 5.5: Collision.

5.5 Final Remarks

Several ITS have been proposed to deal with information management and knowledge distribution related to vehicle traffic conditions. In this type of system, the knowledge is generated from the processing of aggregated local information. However, in the systems found in the literature, all the vehicles perform the tasks of information aggregation and knowledge generation. This situation leads to overloading the network. In order to address this issue, in this chapter we proposed the TRUSTed system. TRUSTed is a distributed system for information management and knowledge distribution, which employs the egocentric betweenness measure, in order to select the most relevant vehicle to perform above-mentioned tasks. In addition, it applies the broadcast suppression mechanism during the knowledge distribution process, reducing the network overhead.

From the analysis of results, two main lessons were learned. The first one is that there is a need for a mechanism to select the most relevant vehicle in the network. Because by using this kind of mechanism it is possible to make the system scalable. The second one refers to the egocentric betweenness measure being a viable option for the selection mechanism in highly dynamic networks.

Chapter 6

Collaborative and Infrastructure-less Vehicular Traffic Management

6.1 Introduction

Over recent years, the research community in the field of communication and ad-hoc networks has been very attracted to social network analysis (SNA) and social network concepts (SNC) to design and implement new algorithms and protocols for socially aware networks, such as mobile social networks (MSNs) and vehicular social networks (VSNs). The legacy of social networks in communication networks is that all entities have a certain degree of interdependence to each other [116]. Such interdependencies can include network topology similarity, physical contact, community, and mutual interest. In addition to the interdependencies, the correlations between the entities can be explored in SNA. Social networks are a virtual group of entities that have some social interdependencies among them, and such interdependencies can be applied to improve the efficiency and effectiveness of network services [3, 57].

The VSN paradigm has emerged through the integration of the concepts of MSNs and VANETs [113, 115, 134]. As a consequence of this integration, two approaches can be explored in the vehicular environment, such as *(i)* application of the SNA [113, 115] techniques and/or *(ii)* use of the SNC [115, 134]. The first approach focuses on identifying the node importance in the network. To this end, three main measures of centrality most used in VSNs are degree, closeness, and betweenness [113, 115]. It is known that the network topology, in VSNs, is highly dynamic and consequently calculating the node centrality is a challenging task. On the other hand, once identified, it can be useful for many applications such as the management of information flow in the network. The second one, however, involves social interactions between nodes that have mutual interests in the temporal virtual community [5, 134]. In other words, such an approach provides the opportunity of vehicles to participate in a virtual vehicle community and share information of mutual interest through social interactions. Based on this idea, each vehicle can share their social information, for example, the personal route. In this way, allowing the practice of collaborative route-planning. The social interaction occurs when vehicles meet each other and share their social information through wireless communication.

ATMS integrate communication, storage, and processing technologies to collect raw data from the VSNs, to extract knowledge of vehicular traffic on roads [5]. The ATMS can provide services to improve traffic management efficiency and safety using such knowledge. For better performance, many ATMS applications require vehicles to periodically share their data (floating car data) between neighboring vehicles, a central server, and/or RSU. Through this sharing, it is possible to create awareness about vehicular traffic conditions [49, 110, 137]. This practice is known as beaconing and the data exchanged is associated with vehicle mobility.This data exchange is performed by the CCH and generally at a transmission frequency between 1 Hz and 10 Hz [130].

Different ATMS have been designed and implemented to overcome the lack of urban mobility that affects the daily life and well-being of the citizens [44, 49, 110, 137]. Several solutions implement a centralized approach [49, 110] due to the difficulty of selecting the most appropriate vehicles, in highly dynamic networks, for congestion detection and calculation of alternative routes. As a result, such solutions are not easily scalable. Another solution employs a distributed approach for congestion detection and calculation of alternative routes [44]. However, to achieve its goal, such a solution needs to segment the entire scenario into multiple sub-regions beforehand. Moreover, the alternative route is calculated selfishly, i.e., without considering the routes chosen by neighboring vehicles.

Based on the gaps found, SOPHIA, a distributed **S**ystem of urban m**O**bility management based on a collaborative a**P**proach in ve**HI**cular soci**A**l networks was designed and implemented. Inspired by the two VSN approaches mentioned above, an SNA technique to classify and select the vehicles in each clustering was applied to reduce bandwidth consumption. Two SNCs were employed (social interaction and virtual temporal community) to perform the exchange of information of common interest. This exchange of information helps in alternative route-planning in a collaborative way, thus improving urban mobility management. In brief, the focus of the SOPHIA system is to minimize the problems associated with traffic congestion, in a distributed manner, and without jeopardizing its scalability.

To address the aforementioned issues, this chapter firstly introduces a brief survey of state of the art (Section 6.2). After that, Section 6.3 describes the design of SOPHIA. Performance evaluation and results are discussed in Section 6.4. The final remarks are given in Section 6.5.

6.2 Literature Review

This section presents the related works relevant to the design and implementation of SOPHIA system. Moreover, the aspects related to dynamic clustering algorithms are discussed along with infrastructure-less and infrastructure-based for urban mobility management.

6.2.1 Dynamic Clustering Algorithms

Grouping nodes into clusters has been extensively investigated in many fields, such as wireless ad-hoc networks and mobile ad-hoc networks, by focusing mainly on energy saving [1, 37, 101]. In VANETs, due to the high topology changes, the clustering algorithms proposed for other kinds of ad-hoc networks such as mobile sensor networks are not suitable to be applied in VANETs [37].

In VANETs, clustering techniques have been proposed to improve communication efficiency and facilitate network management, by grouping vehicles in a geographical vicinity together. The advantages of clustering can be visible in highly dynamic networks, in which information aggregation and management can be performed in each network cluster [80]. Thus, clustering can increase the network scalability and decrease the communication overhead.

Hafeez et al. [66] proposed a clustering algorithm by considering speed as the main parameter to build clusters. The cluster head (CH) is elected in a distributed manner according to their relative speed and distance from their cluster members (CMs). This algorithm improves cluster stability through diffuse speed processing. Besides that, it chooses the second optimal vehicle as the temporary CH when the original one becomes unavailable.

In [120], the authors proposed a mobility-based clustering scheme according to the parameters of the vehicle's movements, such as moving direction, relative velocity, and the relative distance between vehicles. Such parameters are applied to select the CH. In mobility-based clustering, each CH is located in the geographical center of a cluster, and CMs are inside transmission range of the CH and moving in the same direction as the CH. Hassanabad et al. [69] also proposed a mobility-based clustering scheme like the aforementioned one. The difference between them is that the latter applies the Affinity Propagation algorithm, proposed by the authors, to produce clusters with high stability.

Abuashour and Kadoch [2] proposed the algorithm named CORA–Control Overhead Reduction Algorithm. The proposed algorithm aims to minimize the overhead network generated by CMs in a clustered segment scenario. The CHs are selected based on maximum lifetime among all vehicles that are located within each cluster.

6.2.2 Infrastructure-Based Urban Mobility Management

In [49], the authors proposed a centralized system for traffic management called EcoTrec. The proposed system is centralized because of congestion detection and alternative route calculation are performed by a central entity. The EcoTrec system aims to reduce CO_2 emissions without significantly increase travel time. To this end, the system was built on a three-component architecture: Vehicle Model, Road Model, and Traffic Model. The Vehicle Model collects and updates the individual information of the vehicle, as well as periodically sharing them with the Road Model. The shared information comes from Global Positioning System (GPS), accelerometer, and gyroscope embedded in vehicles. The Road Model is hosted in the RSUs which are along the roads and connected by the Traffic Model. The Traffic Model is a central server containing the characteristics and road traffic conditions. Both Road Model and Traffic Model communicate with vehicles through V2I communication. Each vehicle makes periodic requests to the server about the road traffic condition and if the route is congested, the server sends an alternate route.

In [137], the authors introduced Next Road Rerouting (NRR). The main objective is to assist drivers in choosing the next most appropriate road, to circumvent the congested areas. The proposed system operates in two-stage traffic management: *(i)* estimates only the next road for the vehicle to bypass the congested point, and thereafter, *(ii)* uses the vehicle's GPS to calculate the remainder of the alternate route to the destination. The reason for this approach lies in the fact that the calculation of the next road is less costly than the recalculation of the end-to-end route. The NRR mechanism needs a central server (Traffic Operation Center) to gather all the traffic information. In this case, NRR assumes that there is a traffic light at each intersection, to collect such information. Once the congestion is detected, the server notifies the nearest traffic light of the congested area. Thereafter, the traffic light notifies the next most appropriate road for vehicles. After that, the rest of the route is calculated with the aid of the vehicles' GPS.

Pan et al. [110], the authors proposed a hybrid urban vehicle management system named DIVERT. It is considered a hybrid approach because it requires a central server to collect information from vehicles and detect vehicular traffic condition. The alternative routes calculation is carried out by the vehicles in a collaborative manner. In the DIVERT system, the central server operates as a coordinator that receives the vehicle information (speed, location, and direction) via V2I communication. Through this information, the server can detect congested locations and inform the vehicles that are driving to such locations. In this system, the responsibility for the alternative routes calculation is given to the vehicles. Once they need to compute an alternative route, it must take into account the chosen route of the neighboring vehicles, i.e., a collaborative routing decision applies. It is important to notice that in the DIVERT system, the broadcast suppression mechanism was not applied during the message dissemination process. This can lead to a broadcast storm problem.

6.2.3 Infrastructure-Less Urban Mobility Management

In [44], the authors proposed a distributed system for vehicular traffic management, named FASTER. In the proposed system the congestion detection and alternative route calculation do not need any infrastructure. To achieve its goal, FASTER needs to previously segment the entire scenario into multiple sub-regions (or districts). This is performed to aggregate traffic information. Each district has an area equal to 1-hop communication. Each vehicle periodically collects and transmits information, such as average speed and route identification to everyone within its transmission range. The vehicle closest to the center of the district is selected to initiate the dissemination of traffic information aggregated to other vehicles. During the dissemination process, a broadcast suppression mechanism is applied to avoid network overhead. In such a system, the calculation of the alternative route is performed selfishly, based on the probabilistic k-shortest path.

Kasprzok et al. [77] presented a decentralized congestion avoidance strategy for connected vehicles. Their approach measures the vehicular traffic congestion level of a road segment using the amount of wireless network traffic generated by vehicle-to-vehicle communications. The vehicle computes an alternative path employing a modified k-shortest path algorithm whose paths are weighted using a Logit model [30] upon the congestion is detected.

In [63], the authors proposed a fully distributed congestion avoidance system which detects traffic congestion and reroutes vehicles to minimize their travel time. The proposed system does not require global traffic information to detect congested areas but rather only the local information about the traffic conditions. According to local traffic information, each vehicle computes the traffic condition in its current road segment. Hereafter, if necessary, it requests information about the alternative paths of the surrounding vehicles to make the choice that will minimize its remaining travel time. This system relies on sending information request messages whenever a vehicle desires or needs to know more about upcoming roads and traffic. This strategy was applied to reduce network overhead and increase system scalability.

On one hand, infrastructure-based vehicular traffic management systems have been most explored, due to the difficult task of selecting the most relevant vehicles within a subset for detecting congestion and calculating alternative routes. On the other hand, distributed systems cannot ignore such a task. For this, for example, in work of [44] is previously segmented the entire scenario and the most central vehicle is chosen. However, this choice is not always the most appropriate. To overcome this gap, a novel dynamic clustering approach based on SNA along with received signal strength was proposed. In addition, most of the known solutions suggest alternative routes in a selfish fashion. To overcome this gap, a novel collaborative rerouting approach based on social interaction and virtual temporal community was proposed.

6.3 Towards the Design of SOPHIA

SOPHIA is a distributed system for urban mobility management based on a collaborative approach in vehicular social networks. The aim of such a system is to improve vehicular flow on the roads without compromising the system's scalability. Taking this into consideration, the system is composed of four components: *(i)* vehicular crowdsensing/environment sensing; *(ii)* dynamic clustering approach; *(iii)* knowledge extraction and distribution; and *(iv)* collaborative route-planning/knowledge consumption, see Figure 6.1. Details of each component are presented below.

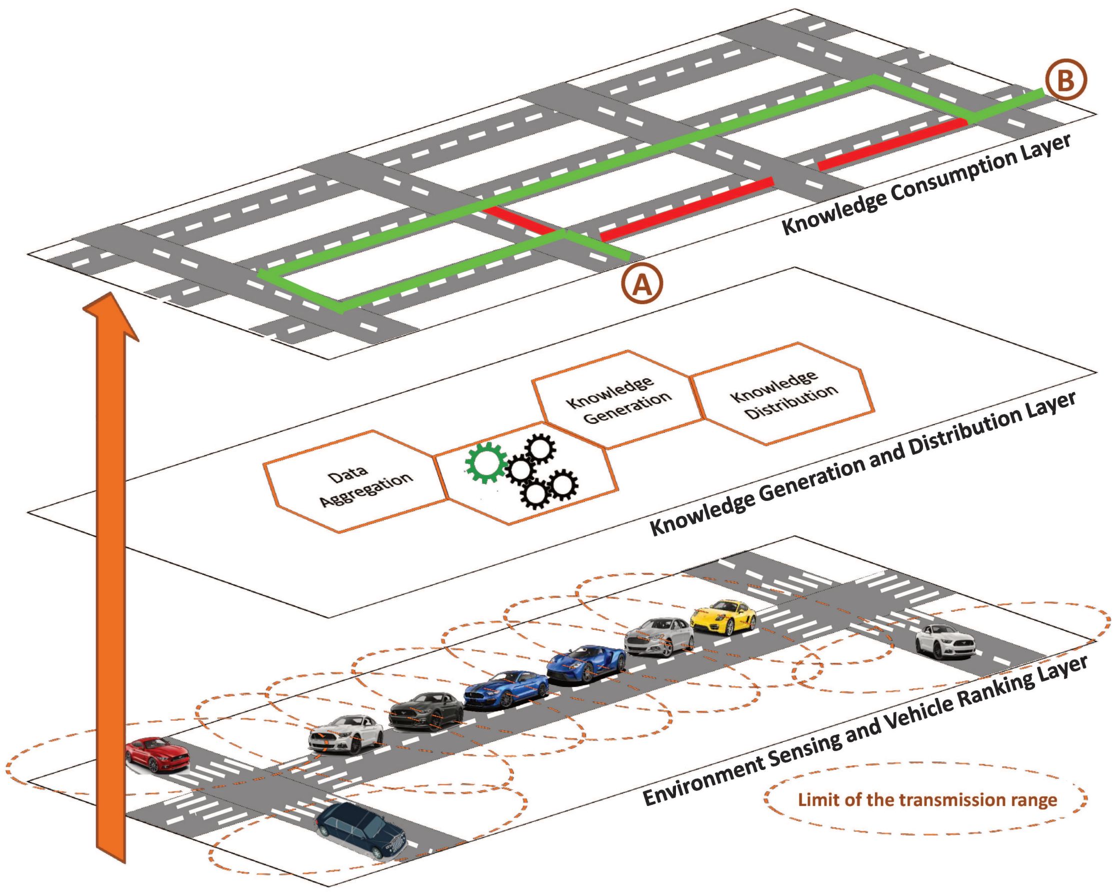


Figure 6.1: Sophia architecture.

6.3.1 Vehicular Crowdsensing

The mobile crowdsensing paradigm (MCS) employs the concept of ubiquitous computing in the collection and sharing of data [22, 138]. In other words, this paradigm aims to incentivize participants to efficiently and effectively contribute to a common goal to use context-related sensing data from their mobile devices in solving a specific problem in a collaborative manner [62]. In addition, by aggregating the crowd-generated local data, it is possible to create cooperative local awareness. Such awareness can lead to improvements in numerous large-scale applications, such as air pollution monitoring and traffic congestion warnings. Since vehicles are equipped with wireless communication technologies along with smart sensors in VSNs, that enables the vehicle crowdsensing (VCS) paradigm [138]. This paradigm, in turn, enables the monitoring of dynamic and large-scale phenomena [18].

The motivation for using VCS lies in the fact that the participants of the networks can solve problems in cooperation. For example, VSN participants can jointly improve urban mobility by sharing data collected about traffic conditions. In doing so, VSNs' systems can aggregate the collected data and extract knowledge (local awareness) about real-time traffic conditions. Thus, the knowledge extracted can assist in urban mobility management.

In this work, the VCS paradigm was applied to create the local traffic awareness, in which vehicles cooperate to sense and collect urban data requested by the system. For this purpose, it was assumed that each vehicle (n) periodically generates a packet (b_n) containing some data collected from onboard units, such as current speed (s_n), location (p_n), time stamp (t_n), and vehicle score (v_{esc_n}), as described in Equation (6.1). The v_{esc_n} will be used in the dynamic clustering mechanism which will be explained later.

$$b_n = (p_n, s_n, v_{esc_n}, t_n) \tag{6.1}$$

6.3.2 Dynamic Clustering Approach

One of the great challenges in highly dynamic networks is to select the most appropriate nodes within a subset to perform a given task [37]. A straight solution for this problem is to employ an infrastructural approach, for example, RSUs and/or a central server [49, 110, 137], thus eliminating the difficult task of selecting vehicles. To overcome this challenge in an infrastructure-less approach, the proposed work adopts a dynamic clustering technique. Unlike the FASTER [44] system, SOPHIA does not need to segment scenario to select the most appropriate vehicle that will perform the congestion detection task.

Network clustering is the division of a graph into a set of subgraphs, called clusters. Each cluster elects one node leader (CH), according to some rules, that works as a local management entity. In addition to that, CMs are all nodes from CH's 1-hop neighbor set. A 1-hop cluster is a clustering such that every node in the network can communicate in 1-hop with the CH of the cluster it belongs to. The cluster is composed of two levels of communications [120]. The first one is intra-cluster communication, where CMs can directly communicate with its CH or nearby CMs within the same cluster. The second one is when a CH communicates with nearby CHs or roadside infrastructures, which is known as inter-cluster communication.

As a general procedure in cluster formation, the nodes participating in, or seeking to join in one, will typically carry out some or all the steps described below [37]:

1. *Neighborhood discovery*: a node generally announces its existence to its neighbors through a periodic short-message transmission, while simultaneously gathering the same message from its neighbors;

2. *CH selection*: after collecting data about the local environment, each node will compute, based on some rule, to find the most appropriate node to act as its CH. In this step, the node can also consider its suitability to be a CH;

3. *Affiliation*: the node will contact the neighbor node that was chosen as the appropriate CH and seek to become a CM of that cluster;

4. *Announcement*: the most appropriate CH may then send an announcement message to its neighbors to initiate the process of cluster formation;

5. *Maintenance*: this step is divided into two parts:

   (a) *As a CH*: if a CH loses all connections with its CMs, the cluster is assumed to be dead, and the procedure is started again (Step 1). On the other hand, a cluster can merge with another one and become a larger cluster. In this case, the node will execute the Step 5(b);

   (b) *As a CM*: the node periodically evaluates the link to its CH. If the link fails it will return to Step 1. If the node receives an affiliation request from a node that does not belong to its group, it can start the CH selection again (Step 2) to choose the next appropriate CH.

In SOPHIA, each cluster is associated with a set of vehicles called CMs and a representative of CH, as shown in Figure 6.2. The vehicles depicted by the labels A and B represent the CHs of the clusters 1 and 2, respectively, while the other vehicles portray the CMs. The vehicle label as 1 will be used in an example afterward. The CH is the vehicle temporarily selected with the responsibility of gathering and forwarding the information on behalf of the CMs. The vehicle with the highest score (v_{esc_n}) is selected as CH, the details of the scoring computation are given below. By means of the dynamic clustering approach, it is possible to overcome the following challenges: *(i)* selecting the most appropriate vehicle in a distributed manner; *(ii)* minimizing the network overhead; *(iii)* increasing the scalability of the system; and *(iv)* facilitating the data flow within network. It is noteworthy that in congested areas, fatally, there will be vehicles in multiple clusters and this particularity was explored to improve the flow of data on the network, otherwise, the information flow would be interrupted.

Our dynamic clustering algorithm procedure only takes into consideration Steps 1 and 2 of the aforementioned general procedure. The idea here is to explore the social properties of nodes to select the CH to improve data flow in the network. This improvement can be done by a path with minimal interference in communication along with the social properties of nodes. To achieve this goal, each vehicle autonomously calculates its score according to neighborhood communication links. This calculation is performed together with a received signal strength indicator, as shown in Equation (6.2).

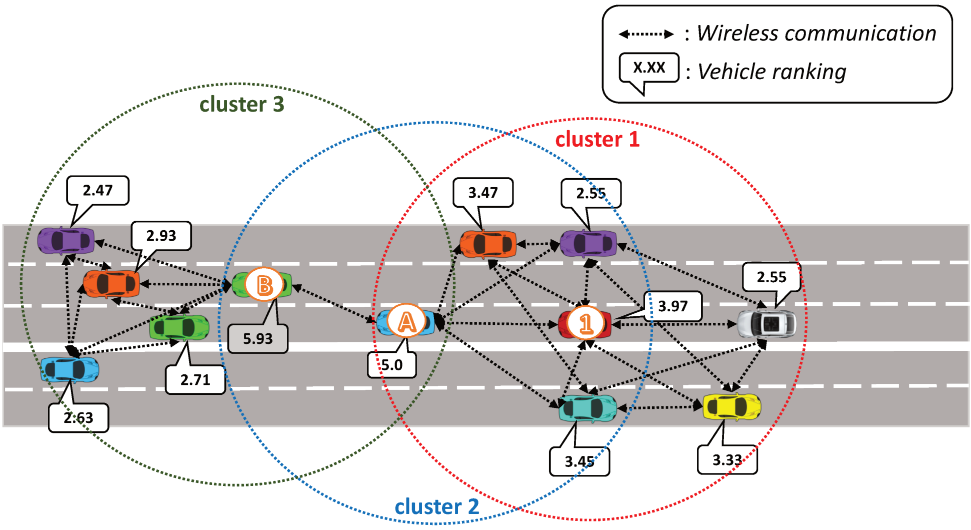


Figure 6.2: Example of clustering, the labels A and B represent the temporary CHs of groups 2 and 3.

$$v_{esc_n} = \sum_{M_n(i,j)\neq 0, i<j} \frac{1}{M_n^2[1-M_n]_{i,j}} + (PL(d_0) + 10\alpha \log(\frac{d}{d_0}) + \Psi_{\sigma_{single}}), d \geq d_0 \qquad (6.2)$$

For simplicity's sake, initially, let us focus on only the first half of the equation. This part of the equation describes an egocentric network metric. One advantage of this metric is that it applies only the locally available topology information. More specifically, the egocentric betweenness metric (EBM) [7, 4]. EBM aims to indicate the relevance of the node for the information flow continuity in the network. It is known that an adjacency matrix, $(M_{k\times k})$, can represent the intercommunication links between the nodes, in which k is the number of 1-hop communication. The EBM calculation is given by the inverse sum of the equation $(M_n^2[1 - M_n]_{i,j})$, where M_n denotes the adjacency matrix of the vehicle n, M_n^2 represents the geodesic distance between the pairs of vehicles i and j, and finally, 1 in the expression corresponds to a matrix with all elements equal to 1.

The second half of the equation refers to the received signal strength indicator. The log-distance path-loss [103] was the model applied. d is the Euclidean distance between vehicles, d_0 is the distance from a reference point to the emitter, $PL(d_0)$ is the power of the reference point to the sender, α describes the path-loss exponent (it varies according to the environment), and $\Psi_{\sigma_{single}}$ is a variable that describes the attenuation of the communication signal. In brief, the power of the received signal fades logarithmically with the distance between the vehicles.

For each change in the local topology, the vehicle's score should be updated. Algorithm 3 describes the procedure of our proposed dynamic clustering. For every change in the network topology (Lines 2 to 4), which corresponds to Step 1 of the general procedure in cluster formation, the vehicle score is recalculated (Line 7), which matches the Step 2. Thereafter, the value is added to v_{esc_n} and transmitted in the subsequent beacon package (b_n), Line 10.

Algorithm 3: Vehicle score calculation.

inputs : $N = \{n_1, n_2, ..., n_n\}$ the set of all vehicles that are currently within the transmission range

output: Vehicle score (v_{esc_n})

```
foreach N_i, i ∈ [1, n] do
    if isNew(n_i) then
        M = updateAdjacencyMatrix(n_i);
if wasUpdate(M) then
    v_esc_n = computeVehicleScore(Equation(6.2));
updateAllBeaconData();
sendBeacon();
```

6.3.3 Knowledge Extraction and Distribution

To better understand the details of the aggregation functions for knowledge extraction, a formal definition of the road network topology is required.

Definition 4 *The road topology can be represented through a directed graph $G = (V, E, W)$, where V corresponds to a set of intersections (v), whereas E denotes to a set of segments (e, where $e \in E \subseteq V^2$). In addition to that, a set of weight ($\rho \in W$) is attributed to each road segment. This weight indicates the level of service and will be explained in detail later on. Finally, a route between two points A and B, $r(A, B)$, is a sequence of intersections $(v_1, \ldots, v_n)$ such that $v_1 = A$, $v_n = B$ and all pairs of consecutive intersections are connected by a road segment, i.e., for all $i = 1, \ldots, n-1$ exists $(v_i, v_{i+1}) \in E$.*

To extract the knowledge about the vehicular traffic condition, two different aggregation functions are required, i.e., *(i)* aggregation of beacons received from the neighborhood—local awareness (Equation (6.3)) and *(ii)* aggregation of local awareness—knowledge of the traffic condition (Equation (6.4)).

$$\Lambda := (E', \Upsilon, \Omega) \tag{6.3}$$

where $E' = \{e_1, \ldots, e_n\} \mid E' \in E(G)$. The parameters Υ and Ω are $\{t_1, \ldots, t_n\}$ and $\{v_{m_1}, \ldots, v_{m_n}\}$, i.e., the current time and average speed of each element of E'.

$$\Lambda_{r,s} := \sum_{1 \leq r,s \leq n} \sigma\Lambda_r + (1-\sigma)\Lambda_s \ , \ \begin{cases} t_r > t_s \\ s_r \, , s_s \neq 0 \end{cases} \tag{6.4}$$

where σ is the weighting factor. The purpose of such a factor is to consider the most current information in the information aggregation process ($t_r > t_s$).

Considering again the example of the Figure 6.2, assuming that the vehicle 1 starts the process of extracting local awareness. After finishing the initial process, it forwards the local awareness to the CH (vehicle A) of its cluster. After that, the CH performs the aggregation of the beacons of its neighborhood (Equation (6.3)) and the aggregate information received from the vehicle 1 (Equation (6.4)). The result of that will be forward to the subsequent CH (vehicle

B) until it reaches the vehicle with the highest score. In this example, the vehicle B has the highest score temporarily, therefore, such a vehicle is responsible for computing the weight of each road segment according to Equation (6.5).

$$\rho_k = v_{agr_k}^{avg} \times (1 - v_{e_k}^{lim})^{-1} \mid \forall e_k \in E' \tag{6.5}$$

where the parameters $v_{agr_k}^{avg}$ and $v_{e_k}^{lim}$ correspond the average aggregate speed and the maximum speed allowed on the road, k, respectively.

After this step, the vehicle B classifies the weight of the road segment according to the level of service (LOS) according to Table 6.1. This table shows the traffic classification for each service level according to the weight (ρ) calculated by Equation (6.5). Each service level depicts a traffic condition. If during the classification process, the LOS D, E, and F are found, a message containing identification about these roads segment is generated and the dissemination process begins. To avoid the problem of the broadcast storm during the data dissemination process, the concept of preference zone (ZoP) [10] was applied. ZoP is a region within the transmission range, whose vehicles within it are most proper to continue the dissemination process. The ZoP concept is based on the delay, this means that the vehicles within it have lower delay (or priority) than the vehicles outside it. Thus, vehicles outside the ZoP receive redundant messages and cancel the scheduled transmission.

Table 6.1: Level of service and traffic classification [51].

Level of Service	**Traffic Classification**	p_i
A	Free flow	(0.0∼0.33]
B	Reasonably free flow	(0.33∼0.4]
C	Stable flow	(0.4∼0.5]
D	Approaching unstable flow	(0.5∼0.7]
E	Unstable flow	(0.7∼0.9]
F	Forced or breakdown flow	(0.9∼1.0]

6.3.4 Collaborative Route-Planning

As mentioned earlier, VSNs involve social interactions (also known as social object relationship–SOR [16]) within a temporal virtual community of vehicles based on common interests or mutual goals [115, 134]. The common interests applied in this work is the alternative routes chosen neighborhood vehicles. Inspired by this idea, it was proposed the collaborative route-planning employing two SNC concepts, such as temporal virtual community and social interactions, as shown in Figure 6.3. Therefore, all vehicles within the temporal virtual community area are considered participants of such a community. The social interactions between community participants are realized through V2V communication and the information of common interest exchanged are the alternative routes chosen. It is worth mentioning that the area covered by the temporal virtual community depends on the circumference radius defined by the application, and the location of the congestion point was defined the central point of the community area.

The main goal in this step is to route vehicles away from the current congestion point, without creating secondary congestion points.

For the sake of clarity, Algorithm 4 is introduced, which describes the procedure of collaborative route-planning. During the route-planning phase, vehicles within the temporal virtual community and closest to the congestion point have priority in choosing an alternative route, i.e., they have the shortest waiting time in choosing an alternative route. This time is directly proportional to the distance between vehicle and congestion point (Line 1). Before calculating an alternative route, the vehicle computes the road popularity (pop) according to the alternative routes chosen by the neighborhood vehicles (Line 3). The pop indicates the most popular roads chosen by vehicles to bypass congestion areas. Thus, road popularity (v) is given by Equation (6.6).

$$pop_v = num_v \times (\frac{len_{(v)}}{lin_{(v)}}) \tag{6.6}$$

where num_v, $len_{(v)}$ and $lin_{(v)}$ represent number of vehicles, road length, and lines on the road surface, respectively.

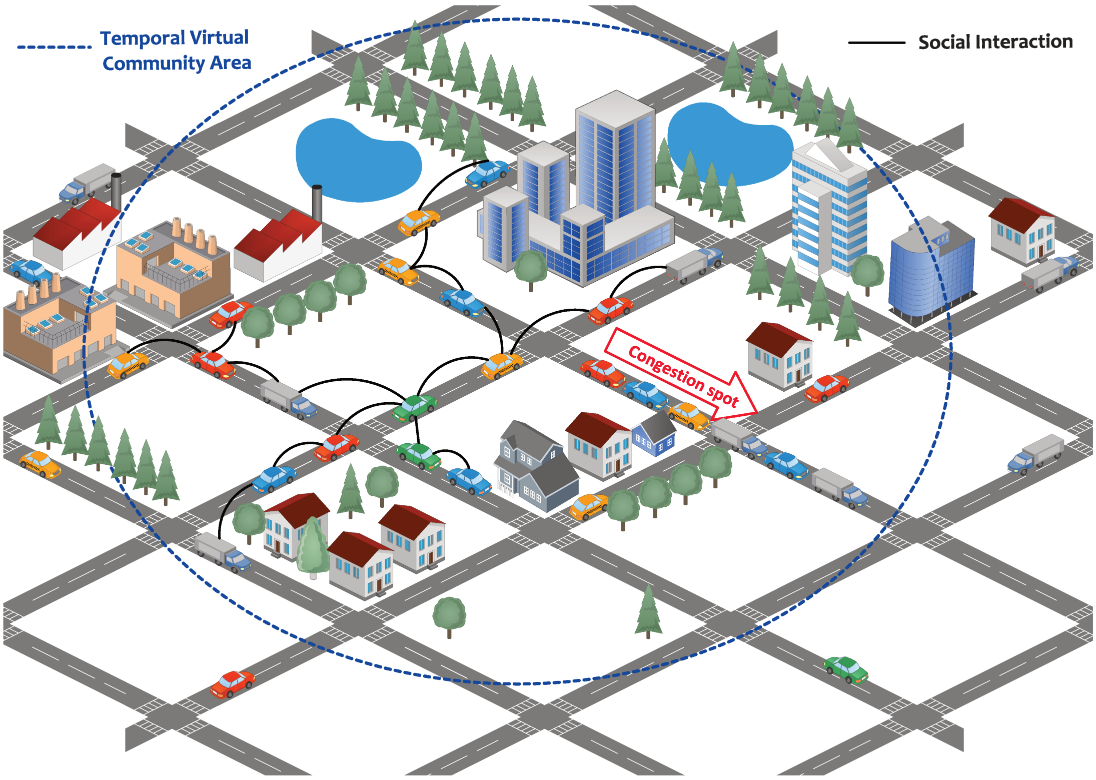


Figure 6.3: Temporal virtual community and social interactions area in VSNs.

Algorithm 4: Collaborative route-planning for vehicles that are moving toward the congested road.

inputs : msg—warning message, which contains the coordinates of the traffic congestion point (s_x, s_y). (r_x, r_y) depicts the coordinates of the receiving vehicle
output: r - the alternative route chosen

1 $waitingTime(ms) = \sqrt{(s_x - r_x)^2 + (s_y - r_y)^2}$;
2 **if** $hasExpired(waitingTime)$ **then**
3 $\quad pop = computeRoadsPopularity(Equation(6.6))$;
4 $\quad r = \text{leastPopularRoute} \parallel r^* \parallel$;
5 $send(r)$;

Now suppose that $r^*(p_{cur}, d_{est})$ denotes the set of all possible alternative routes from the current position (p_{cur}) to the destination (d_{est}). Thus, the choice of an alternative route is given by Equation (6.7), in other words, the vehicle selects the least popular route (r) among all possible routes (Line 4) and shares it through social interaction (Line 5). In this way, reducing the possibility of generating congestion points in another place in the near future.

$$r = \underset{r \in R(p_{cur}, d_{est})}{\text{leastPopularRoute}} \parallel r^* \parallel \tag{6.7}$$

6.4 Performance Evaluation and Results

This section shows the performance assessment of SOPHIA and compares it to FASTER [44], DIVERT [110], and EcoTrec [49] systems. In addition, the EcoTrec system is going to be used as a baseline due to its simplicity. It is worth mentioning that SOPHIA's aim is to make the most of public roads without compromising the system's scalability. For a better presentation, this section was divided into four subsections: simulation setup is shown in Section 6.4.1 and the results and analysis of simulations were divided into: control channel assessment–Section 6.4.2, scalability assessment–Section 6.4.3, and traffic management assessment–Section 6.4.4.

6.4.1 Simulation Setup

The TAPASCologne project[1] of the Institute of Transportation Systems at the German Aerospace Center (ITS-DLR) was adopted in the simulation process. This project aims to reproduce the vehicle traffic, with the highest possible level of realism, in a large-scale scenario of the city of Köln, Germany, see Figure 6.4.

We chose the dataset that contains traffic data traces from 6:00 am to 8:00 am, representing more than 250.000 vehicle routes. However, only a central submap was chosen for the simulation experiments because it displays a higher incidence of traffic congestion (LOS D, E, and F heat bar), as shown in Figure 6.4. With the traffic demand of the submap, it was constructed a new dataset (containing more than 46.000 vehicles routes) and divided into five different vehicle insertion rates, namely 20 %, 40 %, 60 %, 80 %, and 100 %. For example, 20 % means that only 20 % of the total vehicles are inserted in the scenario for the simulation experiments, and

[1] http://kolntrace.project.citi-lab.fr/

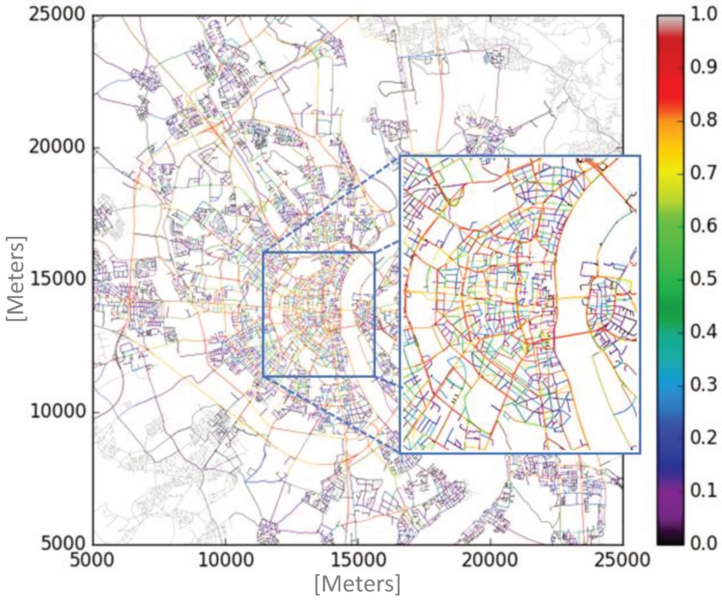


Figure 6.4: Road network of Cologne used in the simulation.

so on. All the experimental results of this work were conducted with a confidence interval of 95 %. Table 6.2 summarizes the simulation parameter settings.

Table 6.2: Simulation parameters settings.

Parameter	Value
Vehicle Insertion Rate	20% to 100%
MAC layer	IEEE 802.11p PHY
Bandwidth	10 MHz
NIC Bitrate	6 Mbps
NIC TX power	20 mW
NIC Sensitivity	-82 dBm
Transmission range	287 m
Beacon transmission rate	1 Hz
Confidence interval	95 %

Additionally, nine metrics were used to evaluate the performance of the SOPHIA system. These metrics were divided into three perspectives (or assessments), which are described in detail below.

1. Control channel assessment

    - *Channel busy ratio*: indicates the interference level. This is estimated as the fraction of the time in which the channel is identified as busy due to packet collisions or successful transmissions;

2. Scalability assessment

    - *Overhead*: measures the total amount of transmitted messages by the vehicles;

- *Latency*: demonstrates the time spent to deliver the messages to the vehicles;
- *Packet loss*: shows the total number of lost packets during the message transmissions;
- *Coverage*: indicates the percentage of messages successfully delivered.

3. Traffic management assessment

- *Travel time*: indicates the average travel time in relation to all vehicles;
- *Travel Time Index*: measures the level of urban traffic congestion [124]. This index is calculated by the ratio of the total travel time to the free-flow travel time;
- *Congestion time loss*: describes the average time spent on congestion;
- *CO_2 emission*: gives the average CO_2 emission of all vehicles.

6.4.2 Control Channel Assessment

As all the solutions apply the beaconing approach in their solution to achieve the goals, and the channel used for that purpose is the control channel. Then, the assessment of the control channel is necessary to analyze. In the experiments, the beacon transmission rate of 1Hz was set to all systems.

Figure 6.5 shows the performance result of the control channel in relation to the vehicle insertion rate. The table (top of figure) depicts the channel busy ratio while the bar chart (bottom) depicts the gain over EcoTrec. As expected, the channel busy ratio increases with the vehicle insertion rate because of the number of vehicles in the neighborhood increases, thereby raising the competition for control channel access. Among all the analyzed solutions, SOPHIA has the lowest average channel busy ratio for all vehicle insertion rates. The reason for this behavior is due to the system's ability to perform better vehicular traffic management. In a few words, SOPHIA distributes vehicular traffic to make the most of the availability of public roads. As a result, the homogeneous distribution of vehicular traffic on the roads reduces the consume on the control channel bandwidth. In addition, we can observe that SOPHIA, FASTER, and DIVERT have a gain, on average, of 19 %, 15 %, and 11.17 %, respectively, over EcoTrec in all vehicle insertion rates. It is important to notice that, on average, SOPHIA had 27 % better result in comparison to FASTER and a 70% improvement in comparison to DIVERT.

6.4.3 Scalability Assessment

This subsection analyzes the scalability results of SOPHIA against the FASTER, DIVERT, and EcoTrec systems in terms of overhead, packet loss, latency, and coverage metrics. Each figure is composed of two bar charts. The top one represents the numerical value of the assessed metric and the bottom one represents the gain with respect to EcoTrec.

Figure 6.6 displays the performance results of all the evaluated systems according to the overhead metric. Both systems, EcoTrec and DIVERT, constantly need to exchange messages between the vehicles and the central server to reach their purposes. Due to this strategy, it is possible to observe that both have a higher average rate of messages transmitted in relation to FASTER and SOPHIA. Another determining factor for this high rate, for both systems, is the

Channel Busy Ratio

	20	40	60	80	100
SOPHIA	0.1539	0.2069	0.2553	0.3111	0.3623
FASTER	0.1631	0.218	0.2686	0.325	0.3763
DIVERT	0.1699	0.2317	0.2827	0.3369	0.3884
EcoTrec	0.2115	0.2589	0.3068	0.3651	0.4319

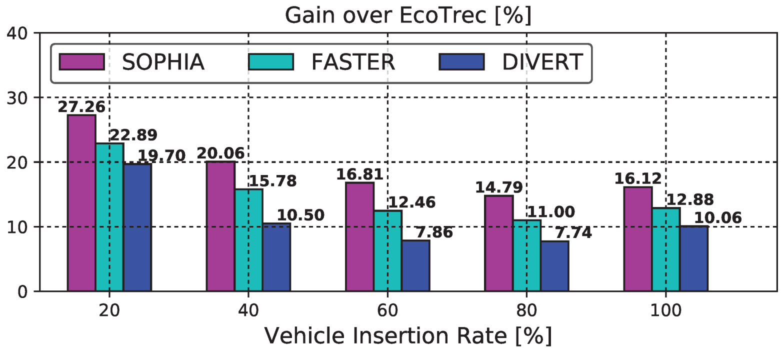


Figure 6.5: Control Channel Assessment.

absence of a broadcast suppression mechanism during the message distribution process. By examining carefully, it is possible to notice that DIVERT has a slightly higher transmission rate than EcoTrec. This is because DIVERT, in addition to communicating with the central server, implements a collaborative routing mechanism when choosing an alternative route. It is worth mentioning that such a mechanism contributes to vehicular traffic management and this contribution will be discussed in the following subsection. Both FASTER and SOPHIA apply vehicle selection techniques for the extraction of knowledge. FASTER segments the scenario into several sub-regions and in each of them one vehicle for knowledge extraction is selected. However, SOPHIA applies a dynamic clustering approach to select the most appropriate vehicle. The dynamic clustering is more appropriate, in this case, as it does not need to segment the scenario for the vehicle selection. It should also be mentioned that both FASTER and SOPHIA applies a mechanism to deal with the broadcast storm problem. Additionally, both have similar performance and they can drastically reduce the total amount of transmitted messages, more than 91 % decrease in comparison with DIVERT and EcoTrec, as shown in Figure 6.6 (bottom).

Figure 6.7 shows the number of packet loss according to the vehicle insertion rate. Since it is known that EcoTrec and DIVERT systems have the highest network overhead compared to FASTER and SOPHIA (Figure 6.6), it is expected that both also have similar results in relation to the packet loss metric. This expectation is confirmed in Figure 6.7. It shows that solutions that have higher transmission rates also have a greater amount of packet loss. Since FASTER and SOPHIA have the lowest network overhead among its competitors, consequently they also have lower packet loss rates. Another factor that causes the rising of packet loss is the intermittent connection between vehicles. According to Figure 6.7 (bottom), the percentage reduction achieved for FASTER and SOPHIA is around 70.8 % and 74.2 % for all vehicle insertion rates, compared to EcoTrec and DIVERT, respectively.

Another metric evaluated is the transmission latency in relation to the vehicle insertion rate,

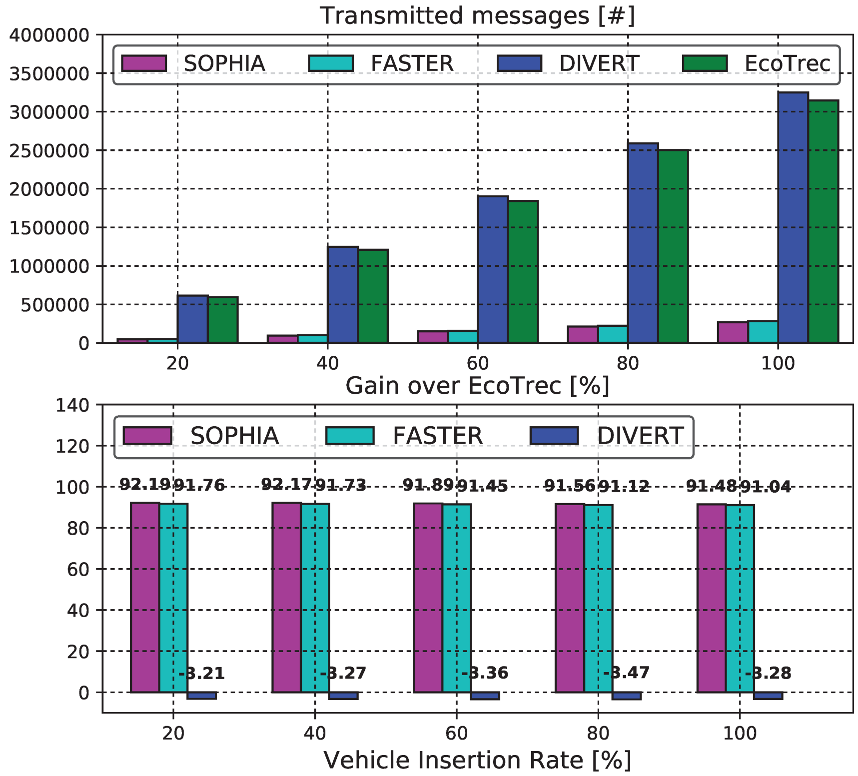


Figure 6.6: Total of transmitted messages.

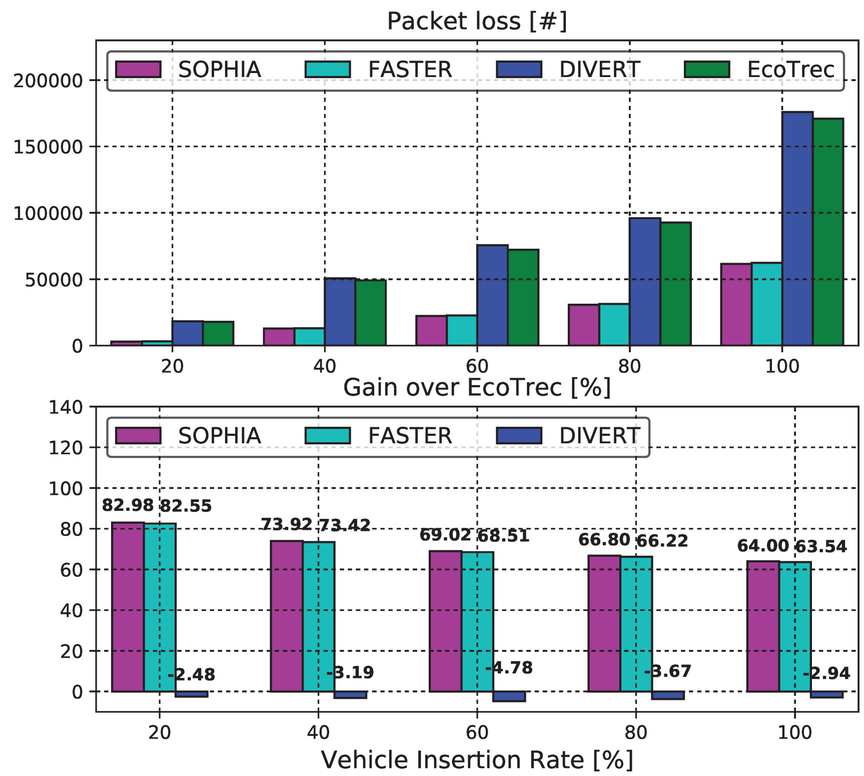


Figure 6.7: Packet loss.

Figure 6.8. In both, the infrastructural and distributed approaches, as vehicle insertion rates increase the latency also increases, as expected. This is because raising the number of vehicles in the simulations increases the network overhead caused by the exchange of messages. However, FASTER and SOPHIA have the lowest latencies compared to other systems analyzed. Comparing numerically, the mean delay of the SOPHIA, FASTER, DIVERT, and EcoTrec systems is around 0.48, 0.42, 1.93, and 1.87 s, respectively. Comparing SOPHIA and FASTER systems between each other, we can observe that the FASTER system has a slight reduction in latency. This is because the knowledge is extracted in several sub-regions, thus delivering it more rapidly to vehicles. Both SOPHIA and FASTER have an average reduction above 74 % compared to the EcoTrec and DIVERT systems, as shown in the Figure 6.8 (bottom).

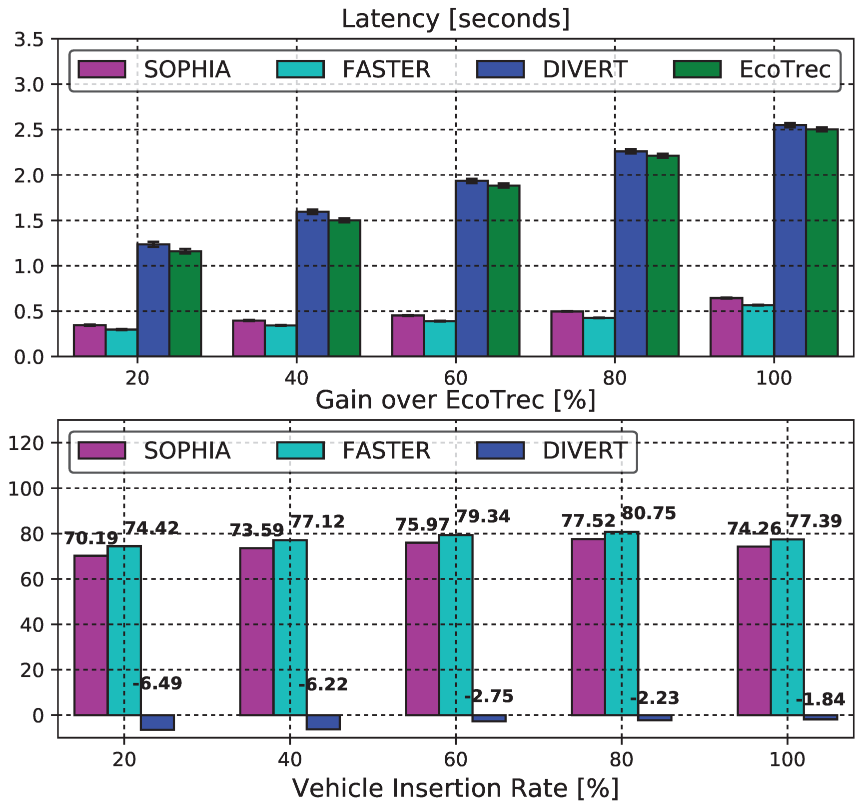


Figure 6.8: Latency.

Figure 6.9 shows the coverage achieved as a function of vehicle insertion rate. EcoTrec has a coverage slightly larger than DIVERT because it has a lower network overhead when compared with its opponent, as depicted in Figures 6.6 and 6.7. On the other hand, since FASTER and SOPHIA have lower network overloads, compared with their competitors, the knowledge extracted can reach a larger number of vehicles at all analyzed insertion rates. FASTER presents a slightly higher result compared to SOPHIA, because knowledge is extracted in several subregions, thus reaching coverage of 1.8 % higher, see Figure 6.9 (bottom). There are two observations that should be considered about the development of the SOPHIA system in relation to FASTER that segments the entire scenario previously are: *(i)* slightly lower coverage and *(ii)* slightly higher latency. However, these two observations do not compromise the system's scalability.

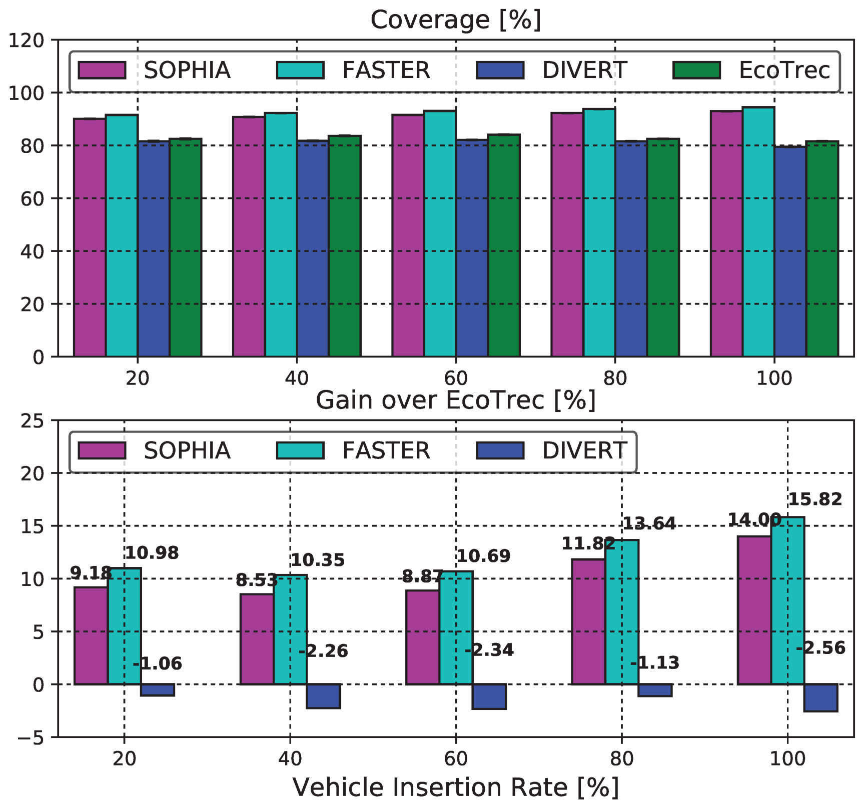


Figure 6.9: Coverage.

6.4.4 Traffic Management Assessment

This section analyzes the urban mobility management of the SOPHIA system as a function of travel time, travel time index, CO_2 emission, and congestion time loss. Each figure is also composed of two bar charts. The top one represents the numerical value of the metric assessed and the bottom one represents the gain with respect to EcoTrec.

Figure 6.10 shows the result of the average travel time for all insertion rates. From the figure, it is possible to notice that the higher the vehicle insertion rate, the longer the average travel time for all solutions analyzed. This behavior is expected since, at high rates the roads become denser, leading to the occurrence of congestion. Among all solutions analyzed, EcoTrec system has the longest average travel time, around 22 min. It is known that the choice of an alternative route within it is given by the path that emits the lowest CO_2 rate until the trip destination. Differently, the FASTER system selects a selfish route based on the probabilistic k-shortest path. This strategy has a gain of 6.71 % agains EcoTrec. Another approach is taken by the DIVERT system, where the vehicles calculate an alternative route collaboratively. In this approach, it is possible to notice a reduction in the mean travel time around 15 % and 8.3 %, compared to EcoTrec and FASTER, respectively. The SOPHIA system applies collaborative routing, such as DIVERT. Even so, it overcomes DIVERT in this metric, due to the low network overhead. As mentioned before, DIVERT has a higher overhead, so many messages arrive corrupted at the recipients. Analyzing numerically, SOPHIA achieves a mean reduction of 6.46 %, 14.75 %, and 21.46 % compared to DIVERT, FASTER, and EcoTrec, respectively, see Figure 6.11 (bottom).

Figure 6.11 indicates the level of traffic congestion as a function of vehicle insertion rate. It is observed that the results of this metric show a behavior similar to the average travel time

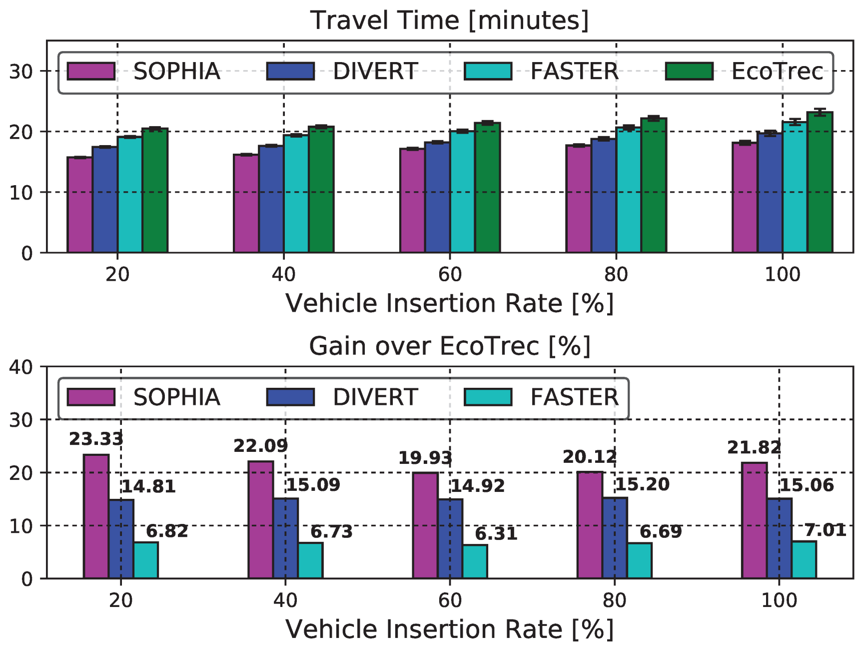


Figure 6.10: Average travel time.

metric (Figure 6.10). This is because both metrics take into account the average travel time. As discussed earlier, the DIVERT system has a slightly higher overhead on the network compared to EcoTrec, as there are exchanges of information on alternative routes chosen by neighboring vehicles. However, this slightly higher overhead causes DIVERT to outpace its competitors (except SOPHIA) in travel time, trip time index, and two other metrics (congestion time loss and CO_2 emission) that will be explained in more detail below.

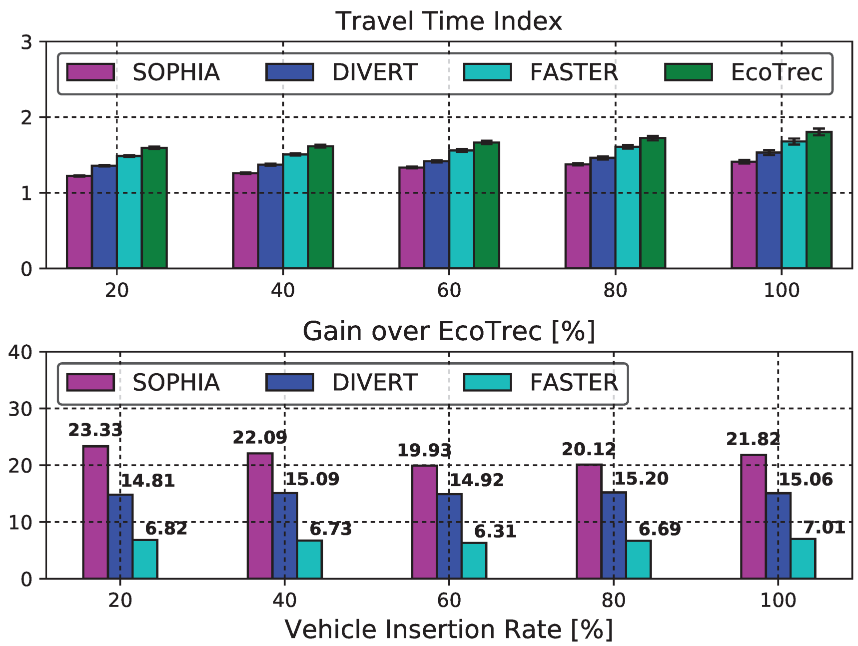


Figure 6.11: Travel time index.

Another important metric to be evaluated is the time lost in congestion, Figure 6.12. All evaluated systems apply some vehicle rerouting mechanism after congestion detection. It is important to emphasize that systems that implement collaborative routing outperform the selfish one. This can be observed in Figures 6.10 and 6.11. To demonstrate them numerically, DIVERT achieves a time reduction of approximately 7.87 % and 15.14 % over FASTER and EcoTrec, respectively. While SOPHIA reaches approximately 21.92 % and 29.18 % compared to FASTER and EcoTrec, respectively, see Figure 6.12 (bottom). As mentioned earlier, the SOPHIA system has a lower overhead compared to DIVERT. Therefore, this fact contributes to the information reaching the largest number of participants thus contributing to improving traffic management efficiency.

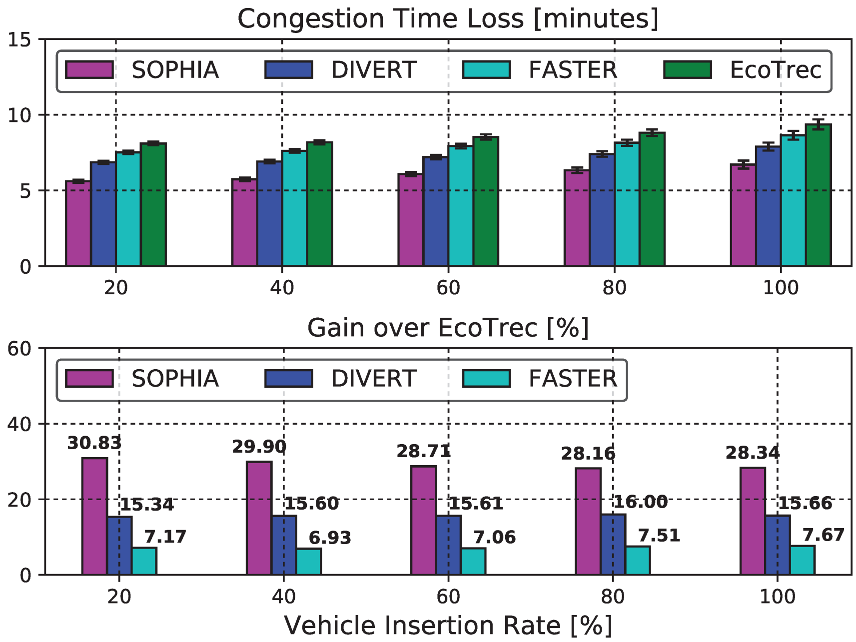


Figure 6.12: Congestion time loss.

Figure 6.13 demonstrates the CO_2 emission in relation to the vehicle insertion rate. As expected, EcoTrec presents the highest CO_2 emission at all the analyzed insertion rates, since it has the highest travel time index (Figure 6.11) and also the highest time lost in congestion (Figure 6.12). By analyzing this metric, it is possible to observe that the most efficient systems, in the urban mobility management, present a smaller amount of CO_2 emission. In this case, the most efficient ones are DIVERT and SOPHIA. This happens because both implement collaborative routing. Analyzing numerically, SOPHIA, DIVERT and FASTER presented a mean reduction in CO_2 emission, against to EcoTrec, of approximately 25.92 %, 13.15 %, and 5.9 %, respectively, see Figure 6.13 (bottom).

6.5 Final Remarks

There is an increasing need for efficient urban mobility management systems to improve vehicular traffic management. To meet this demand, in this chapter, we proposed SOPHIA system, a distributed system of urban mobility management based on a collaborative approach in vehicular social networks. The main advantage of SOPHIA is the combined use of two approaches of vehicular social networks, such as social network concepts and analysis. A metric of social

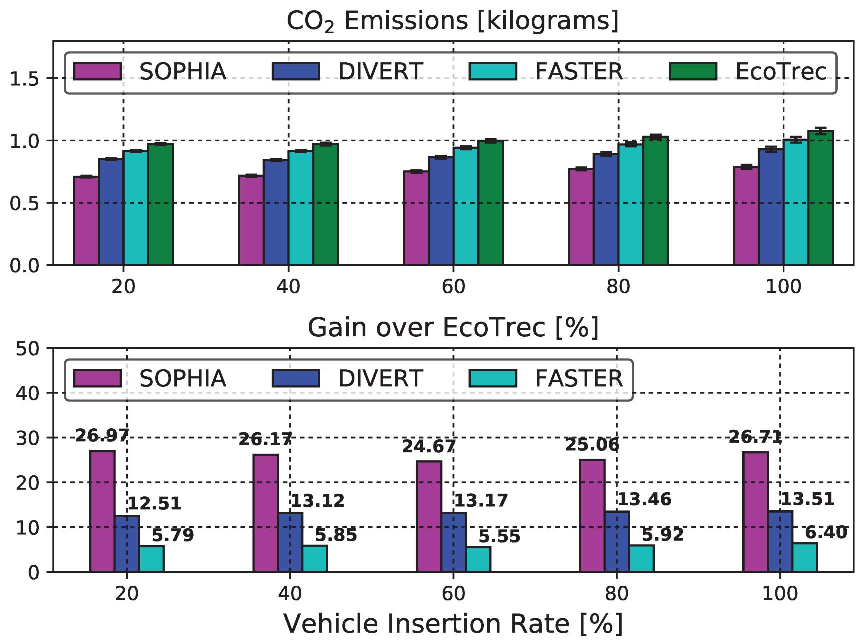


Figure 6.13: CO_2 emissions.

network analysis, more specifically, the egocentric betweenness metric was employed to compute the vehicle ranking. In addition, two social network concepts were employed, i.e., social interaction and temporal virtual community for the collaborative route-planning. Experimental results showed that the difficulty of selecting the most appropriate vehicle, in highly dynamic networks, can be overcome by the proposed dynamic clustering approach. The main advantage of this approach is to rely only on local knowledge of the network topology to achieve its goals. Furthermore, it was proven that SOPHIA was able to do that without jeopardizing system scalability. Another observation presented is that collaborative decision making is more efficient than selfish in alternative routes planning. In summary, the distributed solutions analyzed tend to be more scalable than the infrastructures and those that use the collaborative routing strategy are most efficient in urban mobility management.

Chapter 7

Final Remarks

This chapter summarizes this thesis and discusses directions for future research. The objective is to highlight our main contributions and point out some possible directions to proceed with the research to address the drawbacks of the proposed solutions. In this context, we first present the thesis conclusions in Section 7.1. Then, in Section 7.2, we present future directions of this work. Finally, in Section 7.3, we present the publications related to this thesis.

7.1 Conclusions

Traffic congestion is a daily occurrence for citizens living in large cities around the world. This problem tends to worsen with economic and population growth in urban centers. The increasing vehicular traffic demand may overwhelm the existing transport infrastructure, especially during rush hour. To overcome this issue, two immediate solutions come to mind: *(i)* the expansion of road infrastructure; or *(ii)* the amendment of the traffic management system. In the former solution, the cost of road infrastructure expansion is often impractical, due to financial and/or physical-space constraints. The latter solution, on the other hand, allows the use of already existing technologies, along with the new ones, to improve the efficiency of the vehicular traffic management system. This thesis has been directed toward the second one.

To this end, we proposed the collaborative and infrastructure-less vehicular traffic rerouting as an alternative to improving ITS applications. However, before we reach this primary goal, four research questions (Section 1.2) had to be answered such as:

- **Research Question 1**: *How can we obtain a global view of road network topology without exchanging data between vehicles and the central server for traffic management purposes?*

  To deal with this question, we proposed a people-centric approach for vehicular traffic management in urban centers, called APOLO. The purpose is to use the historical mobility to obtain a global view of the road network and avoid the constant data exchange between the vehicle and the central server. The main idea of APOLO is to periodically analyze the spatial and temporal parameters of mobility patterns of drivers to manage vehicular traffic flow in urban centers.

- **Research Question 2**: *How can we dynamically identify the best-located vehicle among the candidate ones, in a distributed manner, to perform a given application task?*

  To answer this question, we firstly evaluated the betweenness measure based on two approaches: *(i)* local topology information (egocentric) and *(ii)* global topology information (sociocentric). The egocentric network has the benefit of using only locally available knowledge to evaluate the importance of a node. The idea here is to assess whether the egocentric metric, applied in VANETs, has a high degree of similarity compared to the sociocentric. The evaluation results showed that the egocentric betweenness measure has a high degree of similarity. We then designed and implemented TRUSTed, a distributed system for information management and knowledge distribution. In this system, each vehicle autonomously ranks themselves based on the betweenness measure (one-hop link structure). Therefore, best-ranked vehicles are selected to carry out the tasks of information aggregation and knowledge generation. From the obtained results, it is clear that such a measure makes VANETs applications more scalable and leads to more efficient use of network resources.

- **Research Questions 3 and 4**: *Can collaborative route planning help effectively minimize traffic congestions without compromising scalability? Can infrastructure-less vehicular traffic management systems be as efficient as infrastructure approaches and also scalable and cost-effective?*

  To deal with these questions, we proposed and assessed a distributed system of urban mobility management based on a collaborative approach in vehicular social networks (VSNs), called SOPHIA. The main advantage of SOPHIA is the combined use of two approaches of VSNs, such as social network concepts and analysis. A metric of social network analysis, more specifically, the egocentric betweenness metric, was employed to compute the vehicle ranking. Also, two social network concepts were employed for the collaborative route-planning, i.e., social interaction and temporal virtual community. Simulation results confirmed that SHOPIA has excellent potential in increasing system scalability, as well as improving urban mobility management efficiency.

7.2 Future Research Directions

During the development of this research, new ideas have emerged to advance the state-of-the-art in urban mobility management systems. Such ideas, listed below as future works, may guide new research projects in addition to complementing this thesis:

- In our solution, the area covered by the temporal virtual community depends directly on the circumference radius defined by the application. In this regard, a solution that goes one step further is to propose a community detection algorithm in highly dynamic topology. Where each community, Figure 7.1, can perform the exchange of information of common interest. Thus the solution does not depend on a static delimitation of community;

- Current solutions for urban mobility management that use the Fog Computing infrastructure (for example, RSU), require all vehicles send their floating car data (FCD) directly

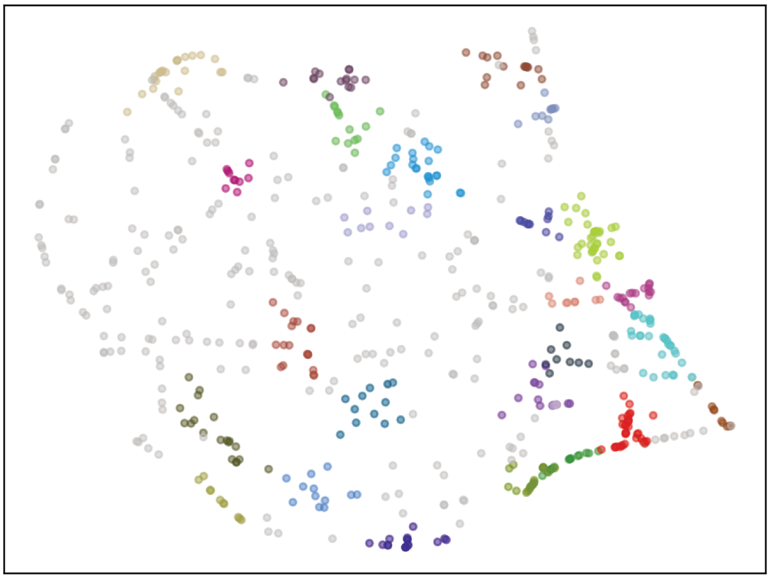
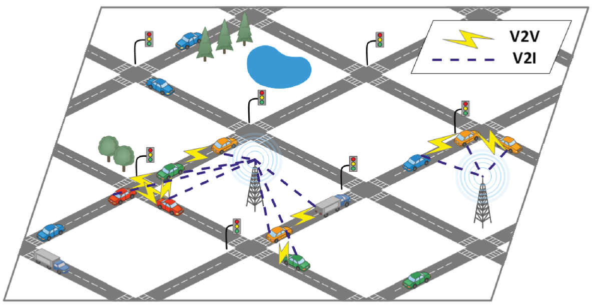
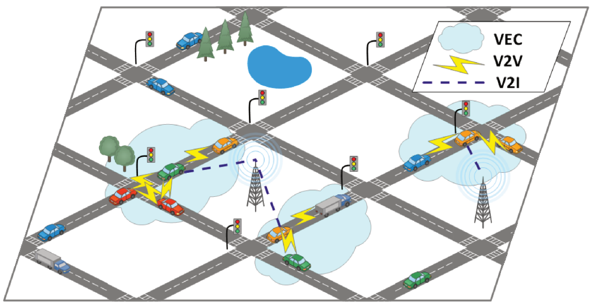


Figure 7.1: Snapshot of communities.

to the RSU (Figure 7.2(a)), consequently overloading the network and increasing the cost of storage and processing. Thus, one way to minimize this communication cost is to process the data locally partially, before sending it to such an infrastructure. Local processing can be performed by a vehicular edge computing (VEC) paradigm [70, 71]. This paradigm uses a set of vehicles as infrastructure to make the best use of computational resources of vehicles autonomously [70, 71, 90, 111, 112], Figure 7.2(b). Fog Computing and VEC have some features in common such as low latency communications and wide geographic distribution. The most distinguishing feature that sets these two paradigms, VEC and Fog, apart is that in the first, the neighboring vehicles can collaboratively participate in addressing a problem. The clustering mechanism proposed in this thesis can be adapted for VEC formation;



(a) A traditional example.



(b) An example applying VEC.

Figure 7.2: Left side represents a traditional example and right side represents an example applying VEC.

- In the context of urban mobility management, multiple heterogeneous data sources can be used to answer the query, retrieve information, and extract and present useful information to the user [118]. In this case, they could be used as new entry data, for example, environmental monitoring data (earthquake, tsunami, and weather) and open access data (criminal event and festival), in addition to FCD to improve the urban mobility management system. The idea here is to propose an infrastructure-based urban mobility management system, where the collection of the FCD is the responsibility of VECs, and other data is collected opportunistically by the Fog infrastructure;

- The infrastructure-based urban mobility management system needs a mechanism for data orchestration, identifying which part of the data load will be handled by VEC resources and which part will be handled by the Fog infrastructure [90, 111]. In this context, another aspect that could be investigated is the implementation of the mechanism for data orchestration, in order to provide low latency, real-time computing, and autonomy to decide data processing and storage.

7.3 Publications from the Thesis

In the following sections, we present a list of publications produced at the moment of this thesis writing. The list is divided into journals, conference papers, and book chapter:

7.3.1 Journals

1. **Akabane, A. T.**, Immich, R., Madeira, E.R. and Villas, L.A. (2019). *"Handling Dynamic Community Structures for Intelligent Traffic Management System with Support of VANETs"*. Elsevier Ad Hoc Networks (Impact Factor: 3.490 and QUALIS A2). Under review;

2. **Akabane, A. T.**, Immich, R., Bittencourt, L. F., Madeira, E.R. and Villas, L.A. (2019). *"Towards a Distributed and Infrastructure-less Vehicular Traffic Management System"*. Elsevier Computer Communications (Impact Factor: 2.766 and QUALIS A2). Accepted;

3. **Akabane, A. T.**, Immich, R., Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2019). *"Exploiting Vehicular Social Networks and Dynamic Clustering to Enhance Urban Mobility Management"*. Sensors (19)16, 3558. (Impact Factor: 3.031 and QUALIS A1);

4. **Akabane, A. T.**, Immich, R., Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2018). *"Distributed Egocentric Betweenness Measure as a Vehicle Selection Mechanism in VANETs: A Performance Evaluation Study"*. Sensors (18)8, 2731. (Impact Factor: 3.031 and QUALIS A1).

7.3.2 Conferences

1. **Akabane, A. T.**, Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2019). *"Aplicando Redes Sociais Veiculares para Aprimorar o Gerenciamento da Mobilidade Urbana"*. In the 37th The Brazilian Symposium on Computer Networks and Distributed Systems (SBRC). (QUALIS - A4);

2. **Akabane, A. T.**, Immich, R., Madeira, E.R. and Villas, L.A. (2018). *"iMOB: An Intelligent Urban Mobility Management System Based on Vehicular Social Networks"*. In IEEE Vehicular Networking Conference (VNC). (QUALIS A2);

3. **Akabane, A. T.**, Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2018). *"Sistema Distribuído para Gerenciamento de Informação e Distribuição de Conhecimento em Redes Veiculares"*. In the 36th The Brazilian Symposium on Computer Networks and Distributed Systems (SBRC). (QUALIS - A4);

4. **Akabane, A. T.**, Immich, R., Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2018). *"TRUSTed: A Distributed System for Information Management and Knowledge Distribution in VANETs"*. In the 23th IEEE Symposium on Computers and Communications (ISCC). (QUALIS - A3);

5. **Akabane, A. T.**, Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2017). *"Modeling and Prediction of Vehicle Routes Based on Hidden Markov Model"*. In IEEE 86th Vehicular Technology Conference (VTC-Fall). (QUALIS - A1);

6. Nikolovski T., **Akabane, A. T.**, Villas, L. A., Pazzi R. W. (2017). *"Efficient encounter-based event dissemination protocol (e-bed) for urban and highway vehicular ad hoc networks"*. In the 21th IEEE Symposium on Computers and Communications (ISCC). (QUALIS - A3);

7. **Akabane, A. T.**, Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2017). *"Applying ego-centric betweenness measure in vehicular ad hoc networks"*. In IEEE 16th Network Computing and Applications (NCA). (QUALIS - B1);

8. **Akabane, A. T.**, Gomes, R.L., Pazzi, R.W., Madeira, E.R. and Villas, L.A. (2017). *"Apolo: A mobility pattern analysis approach to improve urban mobility"*. In IEEE 60th Global Communications Conference (GLOBECOM). (QUALIS - A1);

9. Santos, A. F., **Akabane, A. T.**, Yokoyama, R.S., Loureiro, A. A. F., Villas, L. A. (2016). *"A roadside unit-based localization scheme to improve positioning for vehicular networks"*. In IEEE 84th Vehicular Technology Conference (VTC-Fall). (QUALIS - A1).

7.3.3 Book Chapter

1. Rodrigues, D. O.; Santos, F. A.; Rocha Filho, G. P.; **Akabane, A.T.**; Cabral, R.; Immich, R. ; Lobato Junior, W.; Cunha, F. D.; Guidoni, D. L.; Silva, T. H.; Rosario, D.; Cerqueira, E.; Loureiro, A. A. F.; Villas, L. A. (2019). *"Computação Urbana da Teoria à Prática: Fundamentos, Aplicações e Desafios"*. (Presented in SBRC 2019).

Bibliography

[1] Abdelrahman Abuarqoub, Mohammad Hammoudeh, Bamidele Adebisi, Sohail Jabbar, Ahcène Bounceur, and Hashem Al-Bashar. Dynamic clustering and management of mobile wireless sensor networks. *Computer Networks*, 117:62–75, 2017.

[2] Ahmad Abuashour and Michel Kadoch. Performance improvement of cluster-based routing protocol in vanet. *IEEE Access*, 5:15354–15371, 2017.

[3] Ahmedin Mohammed Ahmed, Tie Qiu, Feng Xia, Behrouz Jedari, and Saeid Abolfazli. Event-based mobile social networks: Services, technologies, and applications. *IEEE Access*, 2:500–513, 2014.

[4] Ademar Akabane, Roger Immich, Richard Pazzi, Edmundo Madeira, and Leandro Villas. Distributed egocentric betweenness measure as a vehicle selection mechanism in vanets: A performance evaluation study. *Sensors*, 18(8):2731, 2018.

[5] Ademar T Akabane, Roger Immich, Edmundo RM Madeira, and Leandro A Villas. imob: An intelligent urban mobility management system based on vehicular social networks. In *2018 IEEE Vehicular Networking Conference (VNC)*, pages 1–8. IEEE, 2018.

[6] Ademar T Akabane, Richard W Pazzi, Edmundo RM Madeira, and Leandro A Villas. Carro: A context-awareness protocol for data dissemination in urban and highway scenarios. In *2016 8th IEEE Latin-American Conference on Communications (LATINCOM)*, pages 1–6. IEEE, 2016.

[7] Ademar T Akabane, Richard W Pazzi, Edmundo RM Madeira, and Leandro A Villas. Applying egocentric betweenness measure in vehicular ad hoc networks. In *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)*, pages 1–4. IEEE, 2017.

[8] Ademar Takeo Akabane, Roger Immich, Richard Wenner Pazzi, Edmundo Roberto Mauro Madeira, and Leandro Aparecido Villas. Exploiting vehicular social networks and dynamic clustering to enhance urban mobility management. *Sensors*, 19(16):3558, 2019.

[9] Ademar Takeo Akabane, Leandro Aparecido Villas, and Edmundo Roberto Mauro Madeira. Gto: A broadcast protocol for highway environments over diverse traffic conditions. In *2014 IEEE 13th international symposium on network computing and applications*, pages 37–40. IEEE, 2014.

[10] Ademar Takeo Akabane, Leandro Aparecido Villas, and Edmundo Roberto Mauro Madeira. An adaptive solution for data dissemination under diverse road traffic conditions in urban scenarios. In *2015 IEEE wireless communications and networking conference (WCNC)*, pages 1654–1659. IEEE, 2015.

[11] Ahmad Alaiad and Lina Zhou. Patients' adoption of wsn-based smart home healthcare systems: An integrated model of facilitators and barriers. *IEEE Transactions on Professional Communication*, 60(1):4–23, 2017.

[12] Muhammad Alam, Joaquim Ferreira, and José Fonseca. Introduction to intelligent transportation systems. In *Intelligent Transportation Systems*, pages 1–17. Springer, 2016.

[13] Mouhannad Ali, Fadi Al Machot, Ahmad Haj Mosa, and Kyandoghere Kyamakya. Cnn based subject-independent driver emotion recognition system involving physiological signals for adas. In *Advanced Microsystems for Automotive Applications 2016*, pages 125–138. Springer, 2016.

[14] Mohammad S Almalag, Michele C Weigle, and Stephan Olariu. Mac protocols for vanet. *Mobile ad hoc networking: Cutting edge directions*, pages 599–618, 2013.

[15] Marica Amadeo, Claudia Campolo, and Antonella Molinaro. Enhancing ieee 802.11 p/wave to provide infotainment applications in vanets. *Ad Hoc Networks*, 10(2):253–269, 2012.

[16] Luigi Atzori, Alessandro Floris, Roberto Girau, Michele Nitti, and Giovanni Pau. Towards the implementation of the social internet of vehicles. *Computer Networks*, 147:132–145, 2018.

[17] Ashley Auer, Shelley Feese, Stephen Lockwood, and Booz Allen Hamilton. History of intelligent transportation systems. Technical report, United States. Department of Transportation. Intelligent Transportation ..., 2016.

[18] Miguel Baguena, Carlos T Calafate, Juan-Carlos Cano, and Pietro Manzoni. An adaptive anycasting solution for crowd sensing in vehicular environments. *IEEE Transactions on Industrial Electronics*, 62(12):7911–7919, 2015.

[19] Varsha Bapat, Prasad Kale, Vijaykumar Shinde, Neha Deshpande, and Arvind Shaligram. Wsn application for crop protection to divert animal intrusions in the agricultural land. *Computers and electronics in agriculture*, 133:88–96, 2017.

[20] Alain Barrat, Marc Barthelemy, Romualdo Pastor-Satorras, and Alessandro Vespignani. The architecture of complex weighted networks. *Proceedings of the national academy of sciences*, 101(11):3747–3752, 2004.

[21] Ramon Bauza and Javier Gozálvez. Traffic congestion detection in large-scale scenarios using vehicle-to-vehicle communications. *Journal of Network and Computer Applications*, 36(5):1295–1307, 2013.

[22] Alessandro Bazzi and Alberto Zanella. Position based routing in crowd sensing vehicular networks. *Ad Hoc Networks*, 36:409–424, 2016.

[23] Stephen P Borgatti and Martin G Everett. A graph-theoretic perspective on centrality. *Social networks*, 28(4):466–484, 2006.

[24] Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2):163–177, 2001.

[25] Ulrik Brandes. On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30(2):136–145, 2008.

[26] Karel A Brookhuis, Dick De Waard, and Wiel H Janssen. Behavioural impacts of advanced driver assistance systems–an overview. *European Journal of Transport and Infrastructure Research*, 1(3), 2019.

[27] Gamini Bulumulle and Ladislau Bölöni. A study of the automobile blind-spots' spatial dimensions and angle of orientation on side-sweep accidents. In *Proceedings of the Symposium on Theory of Modeling & Simulation*, page 18. Society for Computer Simulation International, 2016.

[28] Claudia Campolo and Antonella Molinaro. Multichannel communications in vehicular ad hoc networks: a survey. *IEEE Communications Magazine*, 51(5):158–169, 2013.

[29] Claudia Campolo, Antonella Molinaro, and Alexey Vinel. Understanding the performance of short-lived control broadcast packets in 802.11 p/wave vehicular networks. In *2011 IEEE Vehicular Networking Conference (VNC)*, pages 102–108. IEEE, 2011.

[30] Ennio Cascetta, Francesco Russo, and Antonino Vitetta. Stochastic user equilibrium assignment with explicit path enumeration: comparison of models and algorithms. *IFAC Proceedings Volumes*, 30(8):1031–1037, 1997.

[31] German Castignani, Thierry Derrmann, Raphaël Frank, and Thomas Engel. Driver behavior profiling using smartphones: A low-cost platform for driver monitoring. *IEEE Intelligent Transportation Systems Magazine*, 7(1):91–102, 2015.

[32] Lei Chen and Cristofer Englund. Cooperative intersection management: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 17(2):570–586, 2015.

[33] Lien-Wu Chen, Yu-Chee Tseng, and Kun-Ze Syue. Surveillance on-the-road: Vehicular tracking and reporting by v2v communications. *Computer Networks*, 67:154–163, 2014.

[34] Lin Cheng and Sooksan Panichpapiboon. Effects of intervehicle spacing distributions on connectivity of vanet: a case study from measured highway traffic. *IEEE Communications Magazine*, 50(10):90–97, 2012.

[35] Eunjoon Cho, Seth A Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090. ACM, 2011.

[36] Federal Communications Commission et al. Amendment of the commission's rules regarding dedicated short-range communication service in the 5.850-5.925 ghz band, 2006.

[37] Craig Cooper, Daniel Franklin, Montserrat Ros, Farzad Safaei, and Mehran Abolhasan. A comparative survey of vanet clustering techniques. *IEEE Communications Surveys & Tutorials*, 19(1):657–681, 2016.

[38] Felipe Cunha, Leandro Villas, Azzedine Boukerche, Guilherme Maia, Aline Viana, Raquel AF Mini, and Antonio AF Loureiro. Data communication in vanets: Protocols, applications and challenges. *Ad Hoc Networks*, 44:90–103, 2016.

[39] Alfredo Cuzzocrea, Alexis Papadimitriou, Dimitrios Katsaros, and Yannis Manolopoulos. Edge betweenness centrality: A novel algorithm for qos-based topology control over wireless sensor networks. *Journal of Network and Computer Applications*, 35(4):1210–1217, 2012.

[40] Felipe Domingos Da Cunha, Azzedine Boukerche, Leandro Villas, Aline Carneiro Viana, and Antonio AF Loureiro. Data communication in vanets: a survey, challenges and applications. *[Research Report] RR-8498, INRIA Saclay; INRIA. 2014. hal-00981126v4*, 2014.

[41] Elizabeth M Daly and Mads Haahr. Social network analysis for routing in disconnected delay-tolerant manets. In *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, pages 32–40. ACM, 2007.

[42] Tasneem Darwish and K Abu Bakar. Traffic density estimation in vehicular ad hoc networks: A review. *Ad Hoc Networks*, 24:337–351, 2015.

[43] Maurício Braga de Paula and Cláudio Rosito Jung. Automatic detection and classification of road lane markings using onboard vehicular cameras. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):3160–3169, 2015.

[44] Allan Mariano de Souza and Leandro Aparecido Villas. A fully-distributed traffic management system to improve the overall traffic efficiency. In *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 19–26. ACM, 2016.

[45] George Dimitrakopoulos and Panagiotis Demestichas. Intelligent transportation systems. *IEEE Vehicular Technology Magazine*, 5(1):77–84, 2010.

[46] Xingjian Ding, Guodong Sun, Gaoxiang Yang, and Xinna Shang. Link investigation of ieee 802.15. 4 wireless sensor networks in forests. *Sensors*, 16(7):987, 2016.

[47] Soufiene Djahel, Ronan Doolan, Gabriel-Miro Muntean, and John Murphy. A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches. *IEEE Communications Surveys & Tutorials*, 17(1):125–151, 2014.

[48] Ronan Doolan and Gabriel-Miro Muntean. Vanet-enabled eco-friendly road characteristics-aware routing for vehicular traffic. In *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2013.

[49] Ronan Doolan and Gabriel-Miro Muntean. Ecotrec—a novel vanet-based approach to reducing vehicle emissions. *IEEE Transactions on Intelligent Transportation Systems*, 18(3):608–620, 2017.

[50] Antoine Dutot, Frédéric Guinand, Damien Olivier, and Yoann Pigné. Graphstream: A tool for bridging the gap between complex systems and dynamic graphs. In *Emergent Properties in Natural and Artificial Complex Systems. Satellite Conference within the 4th European Conference on Complex Systems (ECCS'2007)*, 2007.

[51] Lily A Elefteriadou. The highway capacity manual: A guide for multimodal mobility analysis. *TR News*, (306), 2010.

[52] Cristofer Englund, Lei Chen, Alexey Vinel, and Shih Yang Lin. Future applications of vanets. In *Vehicular ad hoc Networks*, pages 525–544. Springer, 2015.

[53] European Parliament (EP). Directive 2010/40/eu of the european parliament and of the council of 7 july 2010 on the framework for the deployment of intelligent transport systems in the field of road transport and for interfaces with other modes of transport, 2010.

[54] EN ETSI. 302 637-2 v1. 3.2 (2014-11) intelligent transport systems (its); vehicular communications; basic set of applications; part 2: Specification of cooperative awareness basic service, 2014.

[55] TS ETSI. 102 724 (v1. 1.1)(10-2012):" intelligent transport systems (its). *Harmonized Channel Specifications for Intelligent Transport Systems operating in the 5GHz Frequency Band, Channel Specifications 5 GHz*, 5, 2012.

[56] Martin Everett and Stephen P Borgatti. Ego network betweenness. *Social networks*, 27(1):31–38, 2005.

[57] Howard Falk. Applications, architectures, and protocol design issues for mobile social networks: A survey. *Proceedings of the IEEE*, 99(12):2125–2129, 2011.

[58] YunFei Fang, Feng Chu, Saïd Mammar, and MengChu Zhou. Optimal lane reservation in transportation network. *IEEE Transactions on Intelligent Transportation Systems*, 13(2):482–491, 2012.

[59] Juan Angel Ferreiro-Lage, Cristina Pereiro Gestoso, Oscar Rubiños, and Fernando Aguado Agelet. Analysis of unicast routing protocols for vanets. In *2009 Fifth International Conference on Networking and Services*, pages 518–521. IEEE, 2009.

[60] Philippe Flajolet and G Nigel Martin. Probabilistic counting algorithms for data base applications. *Journal of computer and system sciences*, 31(2):182–209, 1985.

[61] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.

[62] Raghu K Ganti, Fan Ye, and Hui Lei. Mobile crowdsensing: current state and future challenges. *IEEE Communications Magazine*, 49(11):32–39, 2011.

[63] Mevlut Turker Garip, Mehmet Emre Gursoy, Peter Reiher, and Mario Gerla. Scalable reactive vehicle-to-vehicle congestion avoidance mechanism. In *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, pages 943–948. IEEE, 2015.

[64] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *nature*, 453(7196):779, 2008.

[65] Juan Guerrero-Ibáñez, Sherali Zeadally, and Juan Contreras-Castillo. Sensor technologies for intelligent transportation systems. *Sensors*, 18(4):1212, 2018.

[66] Khalid Abdel Hafeez, Lian Zhao, Zaiyi Liao, and Bobby Ngok-Wah Ma. A fuzzy-logic-based cluster head selection algorithm in vanets. In *2012 IEEE international conference on communications (ICC)*, pages 203–207. IEEE, 2012.

[67] Hannes Hartenstein and Kenneth Laberteaux. *VANET: vehicular applications and inter-networking technologies*, volume 1. Wiley Online Library, 2010.

[68] Hannes Hartenstein and LP Laberteaux. A tutorial survey on vehicular ad hoc networks. *IEEE Communications magazine*, 46(6):164–171, 2008.

[69] Behnam Hassanabadi, Christine Shea, L Zhang, and Shahrokh Valaee. Clustering in vehicular ad hoc networks using affinity propagation. *Ad Hoc Networks*, 13:535–548, 2014.

[70] Xumin Huang, Peichun Li, and Rong Yu. Social welfare maximization in container-based task scheduling for parked vehicle edge computing. *IEEE Communications Letters*, 2019.

[71] Xumin Huang, Rong Yu, Jianqi Liu, and Lei Shu. Parked vehicle edge computing: Exploiting opportunistic resources for distributed mobile applications. *IEEE Access*, 6:66649–66663, 2018.

[72] PB Hunt, DI Robertson, RD Bretherton, and M Cr Royle. The scoot on-line traffic signal optimisation technique. *Traffic Engineering & Control*, 23(4), 1982.

[73] Standards Association IEEE. Ieee guide for wireless access in vehicular environments (wave) architecture. *IEEE Std*, pages 1609–0, 2013.

[74] INRIX Research. Inrix 2018 global traffic scorecard, 2018. `https://static.poder360.com.br/2019/02/INRIX_2018_Global_Traffic_Scorecard_Report_final`, Last accessed on 21/08/2019.

[75] SAE International. Dedicated short range communications (dsrc) message set dictionary, 2016.

[76] Boangoat Jarupan and Eylem Ekici. A survey of cross-layer design for vanets. *Ad Hoc Networks*, 9(5):966–983, 2011.

[77] Andreas Kasprzok, Beshah Ayalew, and Chad Lau. Decentralized traffic rerouting using minimalist communications. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–7. IEEE, 2017.

[78] John B Kenney. Dedicated short-range communications (dsrc) standards in the united states. *Proceedings of the IEEE*, 99(7):1162–1182, 2011.

[79] John B Kenney. Dedicated short-range communications (dsrc) standards in the united states. *Proceedings of the IEEE*, 99(7):1162–1182, 2011.

[80] Sanaz Khakpour, Richard W Pazzi, and Khalil El-Khatib. Using clustering for target tracking in vehicular ad hoc networks. *Vehicular Communications*, 9:83–96, 2017.

[81] Seon-Geol Kim, Jai-Eun Kim, Kang Yi, and Kyeong-Hoon Jung. Detection and tracking of overtaking vehicle in blind spot area at night time. In *2017 IEEE International Conference on Consumer Electronics (ICCE)*, pages 47–48. IEEE, 2017.

[82] David Knoke and Song Yang. *Social network analysis*. Number 154. Sage, 2008.

[83] Stefan Krauß. *Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics*. PhD thesis, 1998.

[84] Rakesh Kumar and Mayank Dave. A comparative study of various routing protocols in vanet. *arXiv preprint arXiv:1108.2094*, 2011.

[85] Daniel T Larose. K-nearest neighbor algorithm. *Discovering Knowledge in Data: An Introduction to Data Mining*, 90, 2005.

[86] Jure Leskovec and Julian J Mcauley. Learning to discover social circles in ego networks. In *Advances in neural information processing systems*, pages 539–547, 2012.

[87] Ruimin Li and Min Guo. Effects of odd–even traffic restriction on travel speed and traffic volume: Evidence from beijing olympic games. *Journal of traffic and transportation engineering (English edition)*, 3(1):71–81, 2016.

[88] Jiun-Ren Lin, Timothy Talty, and Ozan K Tonguz. A blind zone alert system based on intra-vehicular wireless sensor networks. *IEEE Transactions on Industrial Informatics*, 11(2):476–484, 2015.

[89] Lei Liu, Chen Chen, Tie Qiu, Mengyuan Zhang, Siyu Li, and Bin Zhou. A data dissemination scheme based on clustering and probabilistic broadcasting in vanets. *Vehicular Communications*, 13:78–88, 2018.

[90] Shaoshan Liu, Liangkai Liu, Jie Tang, Bo Yu, Yifan Wang, and Weisong Shi. Edge computing for autonomous driving: Opportunities and challenges. *Proceedings of the IEEE*, 2019.

[91] Christian Lochert, Björn Scheuermann, and Martin Mauve. A probabilistic method for cooperative hierarchical aggregation of data in vanets. *Ad Hoc Networks*, 8(5):518–530, 2010.

[92] Guilherme Maia, Cristiano Rezende, Leandro A Villas, Azzedine Boukerche, Aline C Viana, Andre L Aquino, and Antonio A Loureiro. Traffic aware video dissemination over vehicular ad hoc networks. In *Proceedings of the 16th ACM international conference on Modeling, analysis & simulation of wireless and mobile systems*, pages 419–426. ACM, 2013.

[93] Guilherme Maia, Leandro A Villas, Azzedine Boukerche, Aline C Viana, Andre LL Aquino, and Antonio AF Loureiro. Data dissemination in urban vehicular ad hoc networks with diverse traffic conditions. In *2013 IEEE Symposium on Computers and Communications (ISCC)*, pages 000459–000464. IEEE, 2013.

[94] Athanasios Maimaris and George Papageorgiou. A review of intelligent transportation systems from a communications technology perspective. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 54–59. IEEE, 2016.

[95] Highway Capacity Manual. Highway capacity manual. *Washington, DC*, 2, 2000.

[96] Peter V Marsden. Egocentric and sociocentric measures of network centrality. *Social networks*, 24(4):407–422, 2002.

[97] Abbas Mehrabi and Kiseon Kim. Using a mobile vehicle for road condition surveillance by energy harvesting sensor nodes. In *2015 IEEE 40th Conference on Local Computer Networks (LCN)*, pages 189–192. IEEE, 2015.

[98] Prerna Mohit, Ruhul Amin, and GP Biswas. Design of authentication protocol for wireless sensor network-based smart vehicular system. *Vehicular Communications*, 9:64–71, 2017.

[99] Antonio Molina-Pico, David Cuesta-Frau, Alvaro Araujo, Javier Alejandre, and Alba Rozas. Forest monitoring and wildland early fire detection by a hierarchical wireless sensor network. *Journal of Sensors*, 2016, 2016.

[100] Nazmus S Nafi, Reduan H Khan, Jamil Y Khan, and Mark Gregory. A predictive road traffic management system based on vehicular ad-hoc network. In *2014 Australasian Telecommunication Networks and Applications Conference (ATNAC)*, pages 135–140. IEEE, 2014.

[101] Padmalaya Nayak and Anurag Devulapalli. A fuzzy logic-based clustering algorithm for wsn to extend the network lifetime. *IEEE sensors journal*, 16(1):137–144, 2015.

[102] Mark EJ Newman. A measure of betweenness centrality based on random walks. *Social networks*, 27(1):39–54, 2005.

[103] Mikael G Nilsson, Carl Gustafson, Taimoor Abbas, and Fredrik Tufvesson. A measurement-based multilink shadowing model for v2v network simulations of highway scenarios. *IEEE Transactions on Vehicular Technology*, 66(10):8632–8643, 2017.

[104] Diana Vonk Noordegraaf, Jan Anne Annema, and Bert van Wee. Policy implementation lessons from six road pricing cases. *Transportation Research Part A: Policy and Practice*, 59:172–191, 2014.

[105] Tamoghna Ojha, Sudip Misra, and Narendra Singh Raghuwanshi. Sensing-cloud: Leveraging the benefits for agricultural applications. *Computers and electronics in agriculture*, 135:96–107, 2017.

[106] Rene Oliveira, Carlos Montez, Azzedine Boukerche, and Michelle S Wangham. Reliable data dissemination protocol for vanet traffic safety applications. *Ad Hoc Networks*, 63:30–44, 2017.

[107] Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social networks*, 32(3):245–251, 2010.

[108] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

[109] Juan Pan, Iulian Sandu Popa, Karine Zeitouni, and Cristian Borcea. Proactive vehicular traffic rerouting for lower travel time. *IEEE Transactions on vehicular technology*, 62(8):3551–3568, 2013.

[110] Juan Susan Pan, Iulian Sandu Popa, and Cristian Borcea. Divert: A distributed vehicular traffic re-routing system for congestion avoidance. *IEEE Transactions on Mobile Computing*, 16(1):58–72, 2017.

[111] Qi Qi and Zhanyu Ma. Vehicular edge computing via deep reinforcement learning. *arXiv preprint arXiv:1901.04290*, 2018.

[112] Qi Qi, Jingyu Wang, Zhanyu Ma, Haifeng Sun, Yufei Cao, Lingxin Zhang, and Jianxin Liao. Knowledge-driven service offloading decision for vehicular edge computing: A deep reinforcement learning approach. *IEEE Transactions on Vehicular Technology*, 68(5):4192–4203, 2019.

[113] Jun Qin, Hongzi Zhu, Yanmin Zhu, Li Lu, Guangtao Xue, and Minglu Li. Post: Exploiting dynamic sociality for mobile advertising in vehicular networks. *IEEE Transactions on Parallel and Distributed Systems*, 27(6):1770–1782, 2016.

[114] Carlos Quadros, Eduardo Cerqueira, Aldri Santos, Jaehan Lim, and Mario Gerla. Beacon-less video streaming management for vanets based on qoe and link-quality. In *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, pages 191–198. IEEE, 2015.

[115] Azizur Rahim, Xiangjie Kong, Feng Xia, Zhaolong Ning, Noor Ullah, Jinzhong Wang, and Sajal K Das. Vehicular social networks: A survey. *Pervasive and Mobile Computing*, 2017.

[116] Azizur Rahim, Tie Qiu, Zhaolong Ning, Jinzhong Wang, Noor Ullah, Amr Tolba, and Feng Xia. Social acquaintance based routing in vehicular social networks. *Future Generation Computer Systems*, 2017.

[117] Amir-Mohammad Rahmani, Nanda Kumar Thanigaivelan, Tuan Nguyen Gia, Jose Granados, Behailu Negash, Pasi Liljeberg, and Hannu Tenhunen. Smart e-health gateway: Bringing intelligence to internet-of-things based ubiquitous healthcare systems. In *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, pages 826–834. IEEE, 2015.

[118] M Mazhar Rathore, Awais Ahmad, Anand Paul, and Seungmin Rho. Urban planning and building smart cities based on the internet of things using big data analytics. *Computer Networks*, 101:63–80, 2016.

[119] Osama M Hussain Rehman, Hadj Bourdoucen, and Mohamed Ould-Khaoua. Forward link quality estimation in vanets for sender-oriented alert messages broadcast. *Journal of Network and Computer Applications*, 58:23–41, 2015.

[120] Mengying Ren, Lyes Khoukhi, Houda Labiod, Jun Zhang, and Veronique Veque. A mobility-based scheme for dynamic clustering in vehicular ad-hoc networks (vanets). *Vehicular Communications*, 9:233–241, 2017.

[121] Pierpaolo Salvo, Francesca Cuomo, Andrea Baiocchi, and Izhak Rubin. Probabilistic relay selection in timer-based dissemination protocols for vanets. In *2014 IEEE International Conference on Communications (ICC)*, pages 2725–2730. IEEE, 2014.

[122] Robert K Schmidt, Tim Leinmuller, Elmar Schoch, Frank Kargl, and Gunter Schafer. Exploration of adaptive beaconing for efficient intervehicle safety communication. *IEEE network*, 24(1):14–19, 2010.

[123] Elmar Schoch, Frank Kargl, Michael Weber, and Tim Leinmuller. Communication patterns in vanets. *IEEE Communications Magazine*, 46(11):119–125, 2008.

[124] David Schrank, Bill Eisele, and Tim Lomax. Tti's 2012 urban mobility report: Powered by inrix traffic data. college station, texas: Texas a&m transportation institute, texas a&m university system. accessed 18 april 2012, 2012.

[125] Kaveh Shafiee and Victor CM Leung. Connectivity-aware minimum-delay geographic routing with vehicle tracking in vanets. *Ad Hoc Networks*, 9(2):131–141, 2011.

[126] Steven E Shladover, Christopher Nowakowski, Xiao-Yun Lu, and Robert Ferlis. Cooperative adaptive cruise control: Definitions and operating concepts. *Transportation Research Record*, 2489(1):145–152, 2015.

[127] Arthur G Sims and Kenneth W Dobinson. The sydney coordinated adaptive traffic (scat) system philosophy and benefits. *IEEE Transactions on vehicular technology*, 29(2):130–137, 1980.

[128] Antonios Skordylis and Niki Trigoni. Efficient data propagation in traffic-monitoring vehicular networks. *IEEE Transactions on Intelligent Transportation Systems*, 12(3):680–694, 2011.

[129] Christoph Sommer, Stefan Joerer, and Falko Dressler. On the applicability of two-ray path loss models for vehicular network simulation. In *2012 IEEE Vehicular Networking Conference (VNC)*, pages 64–69. IEEE, 2012.

[130] Christoph Sommer, Stefan Joerer, Michele Segata, Ozan K Tonguz, Renato Lo Cigno, and Falko Dressler. How shadowing hurts vehicular communications and how dynamic beaconing can help. *IEEE Transactions on Mobile Computing*, 14(7):1411–1421, 2015.

[131] Nasser Torabi and Behrouz Shahgholi Ghahfarokhi. Survey of medium access control schemes for inter-vehicle communications. *Computers & Electrical Engineering*, 64:450–472, 2017.

[132] United Nations: DESA/Population Division. World urbanization prospects: Highlights, 2018. `https://population.un.org/wup/Publications/Files/WUP2018-Highlights`, Last accessed on 21/08/2019.

[133] Andrés Vázquez-Rodas and J Luis. A centrality-based topology control protocol for wireless mesh networks. *Ad Hoc Networks*, 24:34–54, 2015.

[134] Anna Maria Vegni and Valeria Loscri. A survey on vehicular social networks. *IEEE Communications Surveys & Tutorials*, 17(4):2397–2419, 2015.

[135] Wantanee Viriyasitavat, Fan Bai, and Ozan K Tonguz. Toward end-to-end control in vanets. In *2011 IEEE Vehicular Networking Conference (VNC)*, pages 78–85. IEEE, 2011.

[136] Hong Wang, Chi Zhang, Tianwei Shi, Fuwang Wang, and Shujun Ma. Real-time eeg-based detection of fatigue driving danger for accident prediction. *International journal of neural systems*, 25(02):1550002, 2015.

[137] Shen Wang, Soufiene Djahel, Zonghua Zhang, and Jennifer McManis. Next road rerouting: A multiagent system for mitigating unexpected urban traffic congestion. *IEEE Transactions on Intelligent Transportation Systems*, 17(10):2888–2899, 2016.

[138] Xiaojie Wang, Zhaolong Ning, Xiping Hu, Edith C-H Ngai, Lei Wang, Bin Hu, and Ricky YK Kwok. A city-wide real-time traffic management system: Enabling crowdsensing in social internet of vehicles. *IEEE Communications Magazine*, 56(9):19–25, 2018.

[139] Stanley Wasserman, Katherine Faust, et al. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.

[140] Qing Yang, Alvin Lim, Shuang Li, Jian Fang, and Prathima Agrawal. Acar: Adaptive connectivity aware routing for vehicular ad hoc networks in city scenarios. *Mobile Networks and Applications*, 15(1):36–60, 2010.

[141] Yang Yang and T-SP Yum. Delay distributions of slotted aloha and csma. *IEEE Transactions on Communications*, 51(11):1846–1857, 2003.

[142] Mervyn VM Yeo, Xiaoping Li, Kaiquan Shen, and Einar PV Wilder-Smith. Can svm be used for automatic eeg detection of drowsiness during car driving? *Safety Science*, 47(1):115–124, 2009.

[143] Bo Yu, Cheng-Zhong Xu, and Minyi Guo. Adaptive forwarding delay control for vanet data aggregation. *IEEE Transactions on Parallel and Distributed systems*, 23(1):11–18, 2012.

[144] Xi Yu, Huaqun Guo, and Wai-Choong Wong. A reliable routing protocol for vanet communications. In *2011 7th International Wireless Communications and Mobile Computing Conference*, pages 1748–1753. IEEE, 2011.

[145] Yuan Yuan, Jie Luo, Wei Yan, Tong Zhao, and Shanbo Lu. Da2rf: A data aggregation algorithm by restricting forwarders for vanets. In *2014 International Conference on Computing, Networking and Communications (ICNC)*, pages 393–397. IEEE, 2014.

[146] Junping Zhang, Fei-Yue Wang, Kunfeng Wang, Wei-Hua Lin, Xin Xu, and Cheng Chen. Data-driven intelligent transportation systems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 12(4):1624–1639, 2011.

[147] Wenhui Zhang. Analysis of packet forwarding in vanets using probabilistic channel model. In *VTC Spring 2009-IEEE 69th Vehicular Technology Conference*, pages 1–5. IEEE, 2009.

[148] Xinming Zhang, Long Yan, Hui Zhang, and Dan Keun Sung. A concurrent transmission based broadcast scheme for urban vanets. *IEEE Transactions on Mobile Computing*, 2018.

[149] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from gps trajectories. In *Proceedings of the 18th international conference on World wide web*, pages 791–800. ACM, 2009.

[150] Huan Zhou, Shouzhi Xu, Dong Ren, Chungming Huang, and Heng Zhang. Analysis of event-driven warning message propagation in vehicular ad hoc networks. *Ad Hoc Networks*, 55:87–96, 2017.