SAPIENZA
UNIVERSITÀ DI ROMA

Doctoral Thesis

Integrated Wireless Access and Networking to Support Floating Car Data Collection in Vehicular Networks

Author:

Ion Turcanu

Advisor:

Prof. Andrea Baiocchi

PhD in Information and Communication Technologies

October 31, 2017

Department of Information Engineering,
Electronics and Telecommunications (DIET)
Sapienza University of Rome, Italy

Please cite as:
Ion Turcanu, "Integrated Wireless Access and Networking to Support Floating Car Data Collection in Vehicular Networks," Doctoral Thesis, Department of Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, Italy, October 2017.

To my family, for their love and support.

Abstract

Collecting data from a large number of agents scattered over a region of interest is becoming an increasingly appealing paradigm to feed big data archives that lay the ground for a vast array of applications. Vehicular Floating Car Data (FCD) collection, a major representative of this paradigm, is a key enabler for a wide range of Intelligent Transportation Systems (ITS) services and applications aiming at enhancing safety, efficiency and sustainability. Obtaining real time, high spacial and temporal resolution vehicular FCD information is becoming a reality thanks to the variety of communication platforms that are being deployed. Dedicated Short-Range Communication (DSRC) and Long Term Evolution (LTE) are the most prominent communication technologies able to support periodic and persistent FCD collection.

DSRC technology was mainly proposed for safety applications and is specifically tailored for Vehicular Ad Hoc Networks (VANETs). The first parts of this work are dedicated to assessing the suitability of DSRC to support FCD collection in real urban scenarios. We first study the basic communication paradigm that takes place in VANETs to populate vehicles' local data bases with FCD information, named beaconing, and the trade-off between the beaconing frequency and the congestion induced in the wireless shared channel used to exchange these beacons. The primary metric to measure the information freshness inside every vehicle's local data base is the Age-of-Information (AoI). We define an analytical model to evaluate the AoI of a VANET, given the connectivity graph of the vehicles, and validate the model by comparing it with realistic simulations of an urban area.

Then, we propose an integrated DSRC-based protocol that disseminates queries and collects FCD messages from vehicles roaming in a quite large city area efficiently and timely by using a single network structure, i.e., a multi-hop backbone network made up of only vehicle nodes. The proposed solution is distributed and adaptive to different traffic conditions, i.e., to different levels of vehicular traffic density. One of the main protocol advantages is that for the dissemination of queries it exploits an existing standardized data dissemination algorithm, namely the GeoNetworking Contention-Based Forwarding (CBF). The proposed protocol is evaluated with reference to a real urban environment. The main parameters are dimensioned and

an insight into the protocol operation is given. One of the main outcomes of this part of the thesis is the confirmation of the fact that DSRC is suitable to support not only safety applications, but also periodic FCD collection.

The main issue with DSRC is the low penetration rate. LTE on the other hand is pervasive and has been identified as a good candidate technology for non-safety applications. However, a high number of vehicles intermittently reporting their information via LTE can introduce a very high load on the LTE access network. The second part of this work addresses the design and performance evaluation of heterogeneous LTE-DSRC networking solutions to yield significant offloading of LTE – here, DSRC technology can support local data aggregation. We propose distributed clustering algorithms that use both LTE and DSRC networks in the forwarder selection process. We target robustness, optimizing the amount of data and the value of the collection period, keeping in mind the goals of autonomous node operation and minimal coordination effort. Our results clearly indicate that it is crucial to consider parameters drawn from both networking platforms for selecting the right forwarders. We demonstrate that our solutions are able to significantly reduce the LTE channel utilization with respect to other state-of-the-art approaches. The impact of the proposed protocols on the DSRC channels' load is evaluated and proved to be quite small, so that it does not interfere with other VANET-specific messages.

Contents

Abstract **iii**

1 Introduction **1**
1.1 Vehicular Networking for ITS 4
1.2 Contribution 7

2 Background and State-of-the-Art **9**
2.1 Dedicated Short-Range Communication 9
2.2 Long Term Evolution 13
2.3 Contention-Based Forwarding 15
2.4 Simulation Framework 17
2.5 Related Work 18

3 Minimizing the Age-of-Information in VANETs **25**
3.1 AoI Definition and Background 25
3.2 Analytical Model for the AoI Evaluation 28
3.3 Simulation Model for the AoI Evaluation 33
3.4 Numerical Results 36
3.5 Conclusion 41

4 DSRC-based FCD Collection in Vehicular Networks **43**
4.1 DISCOVER Algorithm 44
4.1.1 Dissemination Phase 46
4.1.2 Collection Phase 49
4.2 Simulation Setup 51
4.3 Dissemination Phase Evaluation 53
4.4 Collection Phase Evaluation 56
4.5 Conclusion 63

5 Heterogeneous LTE-DSRC FCD Collection in Vehicular Networks **65**
5.1 VANET-Driven LTE Off-loading 66

5.1.1 Typical FCD Collection Solutions 67
5.1.2 Proposed Solution 69
5.1.3 Simulation Models and Scenarios 72
5.1.4 Performance Metrics and Evaluation 75
5.2 On-the-Fly Distributed Clustering Formation 88
5.2.1 PureLTE Algorithm 89
5.2.2 OFC Algorithm 90
5.2.3 Baseline algorithm 92
5.2.4 OFC with Duplicate Suppression 93
5.2.5 Simulation Framework and Scenario 96
5.2.6 Evaluation Metrics 97
5.2.7 Parametric Study 99
5.2.8 Comparative Performance Evaluation 101
5.3 Conclusion 107

6 Final Conclusions and Remarks **109**

Bibliography **121**

Chapter 1

Introduction

Over the last decades vehicles have become a vital part of our society, improving significantly the quality of our lives. They are used for both private and public transportation of humans and goods, enhancing connectivity, flexibility, freedom of movement, comfort, as well as time and money saving. The benefits of private cars include on-demand and door-to-door movement, which explains the high level of private vehicle usage in modern developed countries. For instance, from Figure 1.1 we can see that on average in Europe there is one car for every two people [1], while in the United States almost every person owns a private car [2]. The economical impact brought by vehicles is undoubtedly substantial. The U.S. Department of Transportation (USDOT) estimated that in 2014 the economic contribution of transportation amounted to $692 billion [3], which is roughly 4 % of the total U.S. Gross Domestic Product (GDP).

On the other hand, the negative aspects of the continuously increasing number of vehicles on the roads are well-known. One of the major problems is the traffic

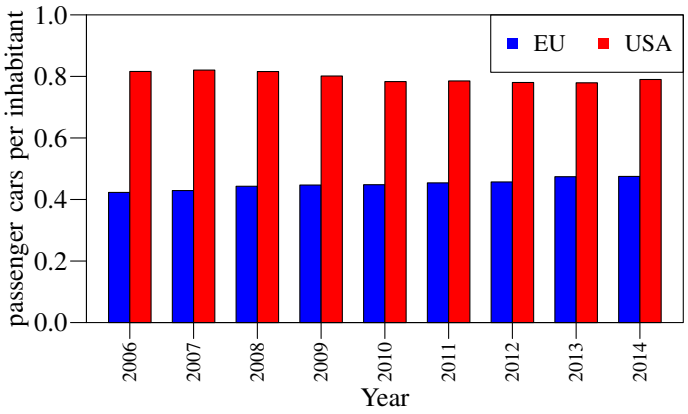


Figure 1.1 – Average number of cars per inhabitant in Europe and USA.

collisions, which cause a high number of deaths around the globe every year. Figure 1.2 points out the fact that the number of people killed in road accidents in Europe has decreased in the last years [4]. However, the same figure shows that this is not the case for the United States [5]. Moreover, the overall number of fatalities is quite high, meaning that we need to significantly improve our current transportation systems to guarantee safety on the roads.

Another aspect that has to be considered is the environmental impact of vehicles. As shown in Figure 1.3, the transportation sector is a major source of greenhouse gas emissions in Europe [6] and in the United States [7]. The main reason is that this sector is one of the most considerable users of energy, burning a significant part of the world's petroleum. In this context, traffic congestion, a common phenomenon in highly populated urban areas, is considered to have a high contribution to air and noise pollution, acid rain, and smog. The main causes for the traffic congestion are the increasing number of vehicles and the inefficiency of the current transportation

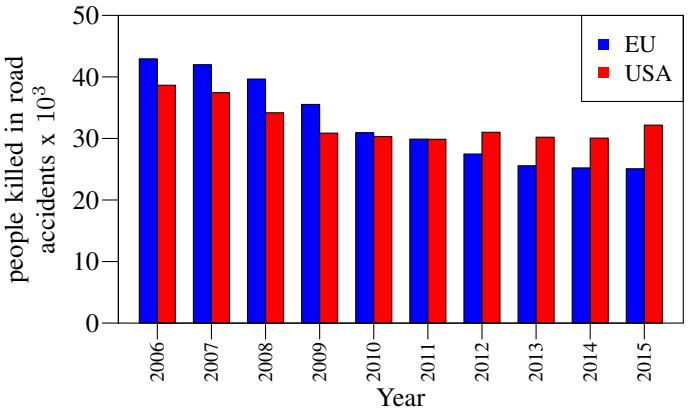


Figure 1.2 – Number of people killed in road accidents in Europe and USA.

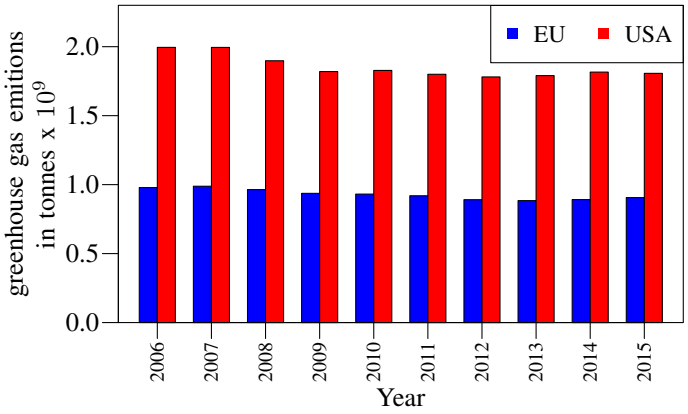


Figure 1.3 – Greenhouse gas emissions in Europe and USA.

systems. Besides the environmental impact, there are also economical aspects to account for. A recent study performed by INRIX [8] found that the congestion cost across the United Kingdom, Germany and the United States in 2016 was almost $450 billion, or $971 per capita. Same study found that in 2016 American, British, and German drivers on average spent 42 h, 32 h and 30 h respectively in congestion during peak hours.

All these facts suggest that the efficiency of the current transportation systems needs to be significantly improved. A particular attention has to be paid to providing updated information to drivers, efficiently and immediately responding to accidents and incidents, and gathering sufficient information for long-term road network planning. Nowadays, existing traffic management systems rely on conventional traffic data sources, which typically consist of intrusive technologies, such as inductive loop detectors, pneumatic road tubes, piezoelectric sensors, and/or non-intrusive technologies, such as manual count systems, microwave radar sensors, passive/active infrared, ultrasonic acoustic, video detection systems. An example of such technology can be seen in Figure 1.4. Of course, these sensing technologies provide useful data, but they have a number of drawbacks, such as limited coverage, expensive implementation and maintenance costs, sensitivity to weather conditions, all of which decrease the accuracy of the collected information.

[1] "Traffic jam detector" by Heidas is licensed under CC BY-SA 3.0

Figure 1.4 – Traffic jam detector in Germany[1].

To cope with these limitations, more recently, advanced traffic management systems started combining the existing on-road sensors with alternative data sources, such as the vehicles themselves. Nowadays, modern vehicles are becoming an essential source of information in the context of smart cities and Intelligent Transportation Systems (ITS). They can be seen as information hubs [9] due to their increasing computing and storage capacity, but also as mobile sensors due to their mobility and growing number of on-board sensors. The information provided by vehicles can contain kinematic data for traffic monitoring (e.g., car position, speed, direction of travel, time), technical and service data for vehicle monitoring, or environmental data for urban sensing. This information, known in the literature as Floating Car Data (FCD), needs to be periodically collected and reported to a remote central server for processing.

1.1 Vehicular Networking for ITS

ITS integrate Information and Communication Technologies (ICT) with transportation engineering methods to get an improved knowledge of current and future states of the transportation system and, possibly, to react to unexpected perturbations in order to keep the system near a desired state of safety, efficiency and comfort. The transportation system is traditionally represented through the interactions between its elementary components: travelers, vehicles, and infrastructure. These interactions affect and, at the same time, are affected by the external environment, both at monadic level - that is, the single vehicle - and at aggregate level, represented by the traffic system. ITS applications enhance efficiency and effectiveness of these interactions thanks to a set of sensors that monitor the environment up to a certain extent, and a set of actuators that put in practice predetermined control rules.

The concept of ITS is not new. People have been thinking about safe, efficient, and advanced transportation systems for decades. In 1939, at the New York World's Fair, one of the most popular attractions was the Futurama ride sponsored by General Motors [10]. During this presentation, visitors could see in miniature and hear the description of what the transportation might look like in 20 years. They could learn about sophisticated transportation systems, where vehicles can communicate among themselves to maintain a proper distance from one another. Of course, these ideas seemed like futuristic transportation utopia at that time. However, they paved the way for a period of remarkable development of transportation systems.

The history of the ITS development can be divided in three main phases [11]:

1. The first phase, which comprises the period between 1970 and 1980, marks the beginning of ITS research. This period is characterized by the development of such systems as Auto-fahrer Rundfunk Information (ARI) in Germany, Traffic

Responsive Capabilities (TRC) in Australia, Electronic Route Guidance System (ERGS) in the United States, or Comprehensive Automobile Control System (CACS) in Japan.

2. The second phase, that goes from 1981 until 1994, places the technological progress closer to the transportation systems. In Europe there were two main projects: the Program for a European Traffic System with Higher Efficiency and Unprecedented Safety (PROMETHEUS), an industrial project, and the Dedicated Road Infrastructure for Vehicle Safety in Europe (DRIVE), an European Community project. Another important project that was created in 1991 and still active today is the European Road Transport Telematics Implementation Coordination Organization (ERTICO), which unifies together all the European organizations interested in ITS. The same temporary phase contains the Road/Automobile Communication System (RACS) project in Japan, which is considered to be the basis for our current vehicle navigation system, and the creation of the Intelligent Vehicle Highway Systems (IVHS) group in the United States.

3. The third phase, which started since 1994, is mainly characterized by the general acceptance and recognition of ITS as a major topic of research. Many research and development programs, as well as standardization entities have been established, with the goal of integrating and implementing the ICT in ITS. The same phase recognizes the benefits of the connected vehicles in the context of ITS, with the allocation of a 75 MHz frequency band in the United States (by the U.S. Federal Communications Commission) and 50 MHz in Europe in the 5.9 GHz frequency range, specifically assigned for vehicular communications.

Different reports and studies have proved the benefits that connected vehicles can bring [12], [13]. In fact, Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) safety systems may address up to 83 % of collision-based accidents [12]. At the same time, the estimated positive impact of the connected vehicles applications is significant in terms of mobility (e.g., travel time, speed variations) and environment (e.g., fuel saving, greenhouse gas emissions) [13]. Vehicular connectivity has been prioritized by the USDOT Joint Program Office in the ITS Strategic Plan [14].

Collecting massive data from a vast set of agents scattered over a Region of Interest (ROI) is one of the key features of next generation ITS and of a number of increasingly popular applications. While having agents equipped with possibly multiple sensors gathering data that are periodically reported to some central facility is a well-established and widely investigated paradigm, there are evolving features that are undermining currently available solutions. Today, real-time FCD collection consists in detecting the vehicles via mobile phones or Global Positioning System

(GPS) and extracting useful information, such as vehicle location, velocity, and travel direction. This basically means that all vehicles connected to the cellular network and equipped with a GPS device are able to contribute to the FCD collection process and act as a mobile sensor for the road network.

On the bright side, the steeply growing amount of data that is made available for collection by the increasing number of Internet of Things (IoT) devices [15] connected to the cellular network, which comprises also modern vehicles, provide a potential boost of new applications. On the other side, the unprecedented number of individual terminals that could access the network concurrently and quite often will challenge the current setup of the cellular networks. This last point marks a definite break with the paradigm of few broadband users that has lead the rush to higher network capacities for broadband applications. As a matter of example, it has been observed [16]–[19] that Long Term Evolution (LTE) [20] is definitely inefficient when accommodating a large population of agents that need to send limited amount of data periodically. The main reason is the heavy procedural overhead associated with obtaining and configuring radio resources for carrying data. Those procedures are warranted when a large amount of data is to be transferred, but they become an unsustainable burden for intermittent sources that need to send relatively small data chunks.

In this context, the research community is considering the feasibility of Vehicular Ad Hoc Networks (VANETs) [21], [22] to support the timely and periodical collection of FCD information [23]–[26]. Originally, the main motivation behind VANETs was to enable and to support vehicular safety applications. VANETs exploit the Dedicated Short-Range Communication (DSRC) [22] wireless technology to enable V2V and V2I communication, as well as information sharing between vehicles and pedestrians, bicyclists, also known as Vehicle-to-Everything (V2X) communication. In general, vehicular networks are especially interesting for several reasons:

- automotive applications, ranging from safety to traffic efficiency and infotainment, have witnessed rapid growth over the last years and are on the brink of an explosive spread and impact;

- the density of vehicles in urban areas and the huge amount of data collected by on-board sensors make a major case of big data collection over time and space;

- vehicles have a number of features that suggest specific directions to be pursued for an effective data collection: they can be connected, they have no severe energy constraints (on the communication equipment), they can afford a relevant amount of processing and memory space.

However, a number of issues makes it quite challenging to exploit the DSRC-based VANETs: the highly dynamic nature of these type of networks, the broadcast-based best-effort communication paradigm, the impact of obstacles (e.g., buildings, trees, other traffic participants) on the radio communication channels, the lack of a centralized congestion control mechanism. All these issues require a significant effort from the research community to fully take advantage of the benefits that VANETs and the DSRC technology can bring in the context of massive FCD collection.

1.2 Contribution

The main focus of this thesis is to explore and evaluate new solutions for frequent and intermittent FCD collection in the context of ITS. To do so, we consider the possibility of exploiting the DSRC technology, both as a standalone solution, as well as in conjunction with LTE cellular networks.

One of the main contributions of this dissertation is a quantitative assessment of the benefits brought about by a cooperation between VANETs based on the DSRC technology and the LTE cellular network to tackle the issue of massive spatio-temporal data collection. While 5G is expected to improve substantially the efficiency of Machine-to-Machine (M2M) communications, it does not seem reasonable to give up exploiting the potential of the VANET, both as regards its support of safety messages (which can be deemed to have the highest priority) and to support signaling procedures to coordinate the nodes for an efficient usage of the cellular radio resources.

Another important contribution is the evaluation of the feasibility of the DSRC technology as an independent solution to support FCD collection. Although DSRC was initially proposed for vehicular safety applications, we show that DSRC-based VANETs are a good candidate for timely collection of vehicular information needed to enable a broad range of traffic efficiency applications.

The thesis is divided in several chapters, which are organized as follows:

- Chapter 2 gives an overview of the existing standards and technologies for vehicular networking, describes the simulation tools and models that are being used to evaluate the various proposed solutions, and, finally, it presents an exhaustive review of the literature and of the state-of-the-art FCD collection schemes and algorithms.

- Chapter 3 defines an analytical model to evaluate the Age-of-Information (AoI) of DSRC-based VANETs, given the connectivity graph of the vehicles. Most of the FCD collection mechanisms that exploit the DSRC technology are based on existing information stored by each vehicle in a local database. These databases contain data from other neighboring vehicles and are periodically

updated by a beaconing mechanism. Optimizing the AoI metric is important for the freshness of the gathered information by the FCD collection algorithm. The analytical model described in this chapter provides a handy tool to optimize the AoI trade-off.

- In Chapter 4 we propose an integrated DSRC-based FCD collection protocol that disseminates and collects data of interest in a quite large city area efficiently and timely by using a single VANET-based network structure, i.e., a multi-hop backbone made up of elected relaying vehicle nodes. The relay node selection mechanism is based on a state-of-the-art data dissemination protocol, namely the GeoNetworking Contention-Based Forwarding (CBF) algorithm [27].

- In Chapter 5 we propose and evaluate different heterogeneous LTE/DSRC-based FCD collection mechanisms. First, we show how the DSRC technology can help to significantly offload the LTE cellular network while periodically collecting FCD information from vehicles roaming an area of interest, by exploiting the VANET and the CBF algorithm to reduce the number of vehicles uploading their collected information via LTE. Then, we identify a key LTE parameter that has a substantial impact on the LTE channel utilization and propose a distributed clustering algorithm that considers relevant parameters drawn from both DSRC and LTE networking platforms to properly select the vehicles uploading information via LTE.

- Finally, in Chapter 6 we draw the main conclusions and identify the key directions for possible future developments of this work.

Chapter 2

Background and State-of-the-Art

Allowing vehicles to communicate and exchange information among themselves and with other traffic participants makes them an invaluable resource, which is why intense work has been carried out in the last years by the research community and the automotive industry to enable vehicular networking [28]. Although different technologies have been proposed to enable vehicular networking, such as Wi-Fi [29], [30], Millimeter Wave (mmWave) [31], Visible Light Communication (VLC) [32], [33], in this thesis we focus on the two most prominent communication technologies, namely DSRC [22], based on the IEEE 802.11p Wireless Access for Vehicular Environments (WAVE) standard [21], and LTE [20].

In this chapter we first give a detailed description of the DSRC and LTE communication technologies. Second, we illustrate the operation of a standardized networking protocol proposed for the dissemination of information in VANETs, that we often refer to throughout the thesis. Third, we talk about the main vehicular simulation frameworks that were used in our studies. Finally, we present an exhaustive literature review of the current solutions and algorithms for collecting FCD information in vehicular networks.

2.1 Dedicated Short-Range Communication

The main motivation for DSRC deployment is to enable vehicular safety applications, which mainly depend on the information exchange among vehicles, as well as between vehicles and the road infrastructure. The communication paradigm consists in each DSRC-equipped vehicle periodically broadcasting basic state information, such as current speed and acceleration, direction of travel, geographical position. A vehicle receiving such information, either stores it in a local data base, if the information is new, or it updates the existing records. This data is used by the safety applications to extract useful information, like the trajectory of every neighboring

vehicle, which can be compared with the vehicle's own predicted path, so that to prevent potential collisions. Other examples of V2V safety applications based on DSRC include Forward Collision Warning (warns about an imminent collision with a vehicle ahead), Emergency Electronic Brake Light (notifies the driver about a hard-braking vehicle ahead), Control Loss Warning (informs surrounding vehicles about a control loss event), Do-not-pass / Blind-spot / Lane-change Warning, etc. Vehicles are also able to exchange information with roadside infrastructure equipment, also known as Road Side Units (RSUs), being informed about the traffic light state in an intersection, about a railroad crossing, or current traffic situation.

Different standard organization groups are currently working on the standardization of the DSRC protocol stack, both at a global level, like the Institute of Electrical and Electronics Engineers (IEEE), as well as at a regional level, such as the Society of Automotive Engineers (SAE) in the United States, the European Telecommunications Standards Institute (ETSI) in Europe, and the Association of Radio Industries and Businesses (ARIB) in Japan. In this work we describe only the European and US standards, since they are more similar, at least from the MAC and PHY layers' perspective. Also, for the evaluation part of this thesis, these are the two standards that have been considered. The DSRC protocol stack is shown in Figure 2.1, as proposed by IEEE WAVE [21] and ETSI ITS-G5 [34], and it is compared with the typical TCP/IP model. We can notice that the LLC, MAC, and PHY layers of the IEEE WAVE are embedded into a single Access layer in the ETSI ITS-G5 architecture. At the transport and network layers DSRC supports the use of Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and Internet Protocol version 6 (IPv6), as well as a set of standard dependent protocols, which will be described bellow. At the higher layers ETSI ITS-G5 defines a set of facilities to support ITS applications, such as specific data structures and messages [35]–[37]. In US similar functionalities are specified by the SAE J2735 standard [38].

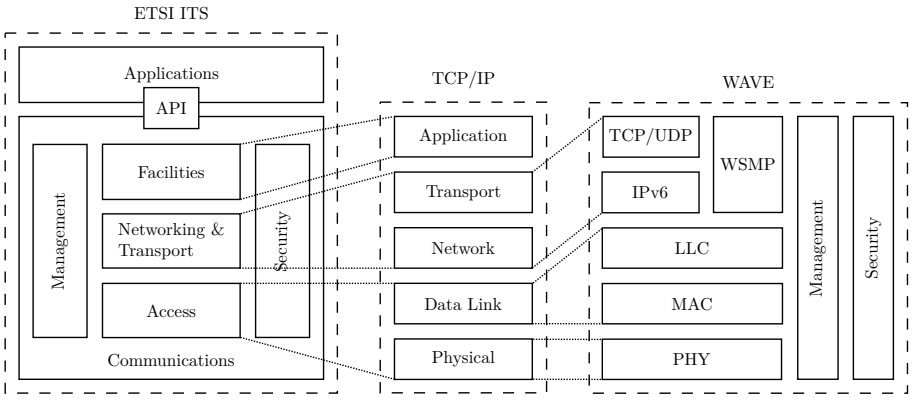


Figure 2.1 – Relationship among protocol stack architectures.

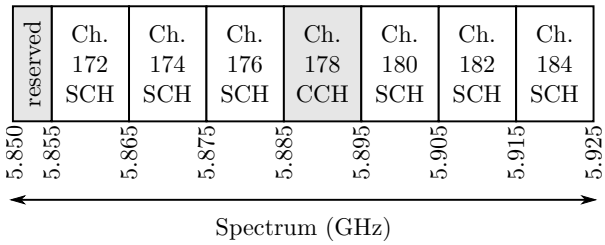


Figure 2.2 – DSRC spectrum allocation.

DSRC operates in a dedicated spectrum in the 5.9 GHz frequency band, which consists of 75 MHz of spectrum divided in 7 channels of 10 MHz each, with a 5 MHz guard band at the low end, as shown in Figure 2.2. From these 7 channels, 1 is the Control Channel (CCH), and the other 6 are Service Channels (SCHs). The PHY layer of DSRC utilizes the IEEE 802.11p standard, a modified version of the well-known IEEE 802.11 (Wi-Fi) standard, which uses the Orthogonal Frequency Division Multiplexing (OFDM) technique. In IEEE 802.11, the OFDM protocol was designed for three different channel widths: 20 MHz, 10 MHz and 5 MHz. Normally, DSRC is using 10 MHz OFDM channels, whose main parameters are shown in Table 2.1.

The IEEE 802.11p MAC layer is similar to the one of IEEE 802.11 when it comes for the rules that govern the frame-by-frame individual transmission. The medium access paradigm is Carrier Sense Multiple Access / Collision Avoidance (CSMA/CA), that comprises also the Enhanced Distributed Channel Access (EDCA) mechanism to ensure the Quality of Service (QoS). The main differences come from the session-based rules: while the IEEE 802.11 standard operates in a Basic Service Set (BSS) context, meaning that the users who want to exchange information have to go first through a synchronization and/or setup procedure, IEEE 802.11p's operation mode is Outside the Context of a BSS (OCB). This is a set of lightweight procedures defined for highly dynamic vehicular environment, meaning that users do not have to belong to the same BSS to be able to communicate among each other. In particular, the OCB operation mode does not require authentication, association. and synchronization.

Parameter	Value
Data rate (Mbit/s)	3, 4.5, 6, 9, 12, 18, 24 and 27
Modulation	BPSK, QPSK, 16-QAM, 64-QAM
Coding rate	1/2, 2/3, 3/4
Number of data subcarriers	48
Number of pilot subcarriers	4
Total number of subcarriers	52
Subcarrier frequency spacing	156.25 kHz
Guard interval	1.6 µs
Symbol interval	8 µs

Table 2.1 – IEEE 802.11 OFDM basic PHY parameters for a 10 MHz channel

The frames that are sent OCB have the Basic Service Set Identifier (BSSID) field set to all 1, which allows a receiver to ignore all other frames that are not sent OCB.

The middle layers standardization of the DSRC protocol stack in the United States is led by the IEEE 1609 working group. They cover three main areas: security services (IEEE 1609.2 [39]), networking services (IEEE 1609.3 [40]), and multi-channel operation (IEEE 1609.4 [41]). As it can be seen from Figure 2.1, the DSRC network layer splits into two branches: the first one is based on the typical Internet protocols, such as TCP, UDP, and IPv6, while the second one uses the WAVE Short Message Protocol (WSMP) defined in IEEE 1609.3 [40]. One of the main services that IPv6 offers is the connectivity, which is guaranteed by a set of well-known routing protocols. The minimum packet overhead of a typical UDP/IPv6 packet is 52 B. However, the dynamic structure of the vehicular networks makes it quite challenging to maintain a proper and reliable path from a sender to a receiver for more than few seconds. Moreover, most of the safety applications are based on 1-hop message broadcasts. This was the main motivation for the definition of the WSMP protocol, that is based on the exchange of WAVE Short Messages (WSMs), whose packet overhead ranges between 5 B and 20 B. The European standard, besides the TCP, UDP, and IPv6, implements the GeoNetworking [27] protocol and the Basic Transport Protocol (BTP) [42], which provide data delivery among DSRC devices and between DSRC devices and other network nodes. Other services include protocols for multi-hop dissemination of information in geographical areas, like CBF [27].

On top of the Network and Transport layers, the American standard is based on SAE J2735 [38], which defines a set of message types aiming to support different kinds of ITS applications. One of the most important message types is the Basic Safety Message (BSM), that contains vehicle safety-related information that is periodically broadcast to surrounding vehicles. BSM structure is made of two parts: the first part, which is 39 B long, contains only the most critical information about the sending vehicle, while the second part allows other optional information to be included. According to the SAE J2735 standard, the BSM messages have to be transmitted on either an event basis, or periodically with a sending frequency ≤ 10 Hz. The European standard has a dedicated Facility layer, as shown in Figure 2.1, that provides ITS-specific message handling and support, as well as data structures to store and maintain different types of information needed by the ITS applications. In particular, ETSI ITS-G5 defines a message type to be transmitted periodically, named Cooperative Awareness Message (CAM) [36], as well as an event-based message type, named Decentralized Environmental Notification Message (DENM) [37]. CAMs have to be exchanged periodically, with a generation period $0.1\,\text{s} \leq T_{\text{gen}} \leq 1\,\text{s}$, to ensure cooperative awareness, while DENMs are triggered by an exceptional event, like a road hazard or an abnormal traffic condition. CAM and DENM messages provide similar functionalities to BSM. The information contained in CAM and DENM

messages is usually stored and maintained in a specific data structure defined by ETSI, named Local Dynamic Map (LDM) [35]. The LDM resides in every ITS vehicle and stores different kinds of information relevant to safety and traffic efficiency applications. The information contained inside an LDM can arrive from different data sources, such as vehicles (i.e., by means of CAM and DENM messages), RSUs, on-board sensors, etc. The applications can access necessary information from the LDM, but also can store data into the LDM.

2.2 Long Term Evolution

The LTE development was started in 2004 by the Third Generation Partnership Project (3GPP)[2] as an evolution of the Universal Mobile Telecommunication System (UMTS). The main motivation was the increasing usage of mobile data and the emergence of a wide range of new applications demanding high bandwidth capacities. LTE is designed to support packet-switched traffic, seamless connectivity, and good QoS.

There are three main components that constitute the high-level architecture of the LTE technology: the User Equipment (UE), which represents the mobile equipment at the user side, the Evolved Packet Core (EPC), which is basically the core network of LTE, and the Evolved UMTS Terrestrial Radio Access Network (E-UTRAN), that manages the communications between the UEs and the EPC and whose main component is the eNodeB. The EPC architecture and operation is out of the scope of this work.

The E-UTRAN protocol stack [43] is shown in Figure 2.3. At the physical layer, LTE uses OFDM in downlink (i.e., data transmission flow from the eNodeB to the UE) and Single Carrier Frequency Division Multiple Access (SC-FDMA) in uplink (i.e., data transmission flow from the UE to the eNodeB). The OFDM symbols are grouped into Resource Blocks (RBs), which can be represented as a time-frequency

[2] http://www.3gpp.org/

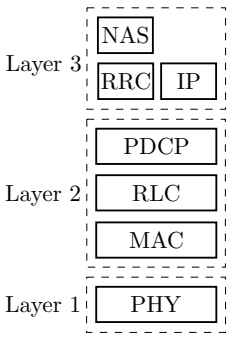


Figure 2.3 – E-UTRAN protocol stack.

grid. In the frequency domain, an RB has a total size of 180 kHz, while in time domain it is 0.5 ms long. The RB allocation is decided by the eNodeB scheduling mechanism based on local policies and on the current channel conditions for each UE. The basic LTE PHY parameters are summarized in Table 2.2.

The LTE Layer 2 is split into the following sublayers: Medium Access Control (MAC), Radio Link Control (RLC) and Packet Data Convergence Protocol (PDCP). Among the main functionalities of the MAC sublayer we can count the mapping between the logical channels and transport channels, scheduling information reporting, error correction through Hybrid Automatic Repeat Request (HARQ), transport format selection. The services of the RLC sublayer include the transfer of upper layer Protocol Data Units (PDUs), concatenation, segmentation and reassembly of RLC Service Data Units (SDUs). PDCP is responsible for header compression and decompression, transfer of user data, in-sequence delivery of upper layer PDUs at PDCP re-establishment of lower layers. The Radio Resource Control (RRC) sublayer of the LTE Layer 3 is in charge of broadcasting system information related to the Access Stratum (AS) and Non-Access Stratum (NAS), paging, security functions including key management, mobility and QoS management functions.

LTE Channel Quality Indicator (CQI)

One of the key features of LTE is the possibility of selecting the downlink/uplink transmission configuration and related parameters depending on the current channel condition, including the interference situation [20]. The instantaneous channel quality, namely CQI, is provided periodically or aperiodically by the terminals to the eNodeB. The eNodeB makes up decisions on resource allocation based on the terminal CQI information. *Periodic* CQI reports can be transmitted on the Physical Uplink Control Channel (PUCCH) or Physical Uplink Service Channel (PUSCH), while *aperiodic* reports can be transmitted only on PUSCH.

In LTE, CQI provides quantized indication of the highest modulation and coding scheme that, if used by the eNodeB, lets the UE demodulate and decode the transmitted downlink data with a maximum block error rate of 10 %. However, the CQI is only a recommendation, meaning that the eNodeB does not need to necessarily

Parameter	Value
Peak data rate (Mbit/s)	75 (uplink), 300 (downlink)
Duplexing	FDD, TDD, half-duplex FDD
Channel coding	Turbo code
Channel Bandwidth (MHz)	1.4, 3, 5, 10, 15 and 20
Available resource block configuration	6, 15, 25, 50, 75 and 100
Modulation schemes	QPSK, 16-QAM, 64-QAM
Multiple access schemes	SC-FDMA (uplink), OFDM (downlink)

Table 2.2 – Main LTE parameters.

use it. The reason is that the eNodeB has to consider also other information when allocating resources. For instance, if the UE needs to transmit only a small amount of data, then there is no need to select a very high data rate, because a small number of RBs with robust modulation is sufficient. There are 15 different CQI values, ranging from 1 to 15. The higher the CQI value reported by the UE, the richer the modulation scheme (from QPSK to 64QAM) and the bigger the coding rate used by the eNodeB to improve the efficiency as much as possible.

There is no explicit description in the standard documents of the mechanism by which the CQI is calculated, but it is known that the Signal to Noise Ratio (SNR) and/or Signal to Interference plus Noise Ratio (SINR) factors play important roles in the CQI computation. How these factors should be used and whether there are any other factors that should be involved is not well defined. Our estimation of the CQI is based on the work of Virdis et al. [44], which uses a mapping table of measured block errors to determine the CQI based on a given SINR value.[3] The SINR is computed as

$$\text{SINR} = \frac{P_s}{\sum_i P_i + N} \tag{2.1}$$

where P_s is the power received from the serving eNodeB, P_i is the power received from the interfering eNodeB i, while N is background Gaussian noise. The received power P is computed as

$$P[\text{dBm}] = P^{\text{tx}}[\text{dBm}] - L_{\text{P}}[\text{dB}] - L_{\text{S}}[\text{dB}] - L_{\text{F}}[\text{dB}] \tag{2.2}$$

where P^{tx} is the transmit power, L_{P} is the path loss [45], while L_{S} and L_{F} represent the attenuation due to slow and fast fading, respectively.

2.3 Contention-Based Forwarding

As discussed in Section 2.1, ETSI ITS-G5 proposes the GeoNetworking protocol at the core of its protocol stack, that implements transport and networking functionalities to support information transfer and exchange among ITS vehicle nodes. The GeoNetworking design supports the communication among single ITS stations, as well as the dissemination of information in a geographical area of interest. Besides the typical 1-hop communication paradigm, the standard defines different modes of data transport: GeoUnicast (i.e., from one node to another), GeoBroadcast (i.e., from one node to all nodes in a geographic area), and GeoAnycast (i.e., from one node to any node inside a geographic area). One of the main features of the GeoNetworking protocol is that it is designed to manage the dynamic characteristics of the VANETs, being able to meet the requirements of different types of applications. Moreover,

[3] `http://github.com/inet-framework/simulte`

GeoNetworking is not technology dependent, meaning that it can operate on top of different access technologies for vehicular communications.

Efficient geographic routing and data dissemination schemes for VANETs have been widely studied in the last years [46]–[48]. The most widespread techniques to disseminate data are broadcast-based, because of the advantage given by eliminating the complexity of route discovery, address resolution and topology management. A recent survey on dissemination protocols in vehicular networks is provided in [46]. One of the most used techniques to implement an efficient data dissemination scheme is to identify only a small subset of vehicles responsible for re-broadcasting the information. We can find this idea in [47], where a new dissemination protocol, named Vehicular Backbone Network (VBN), has been proposed. Here the messages sent out by the RSU are forwarded only by those vehicles that are situated closest to nominal relaying positions, that are spaced out by a range D. The distance D is chosen so as to provide each receiving relay node with a SINR level that can support the intended packet transmission rate. Viriyasitavat et al. [48] face the problem of data dissemination in urban scenarios in the presence of disconnected networks. They propose Urban Vehicular Broadcast (UV-CAST), a completely distributed protocol which utilizes both direct relays through multi-hop transmissions and indirect packet relays through the *store-carry-forward* mechanism.

The main motivation for the above mentioned data dissemination schemes is to mitigate the well-known broadcast storm problem [49]. The GeoNetworking standard addresses this problem with several broadcast-based multi-hop forwarding algorithms: Simple GeoBroadcast, Contention-Based Forwarding (CBF), and Advanced Forwarding (AF). According to the performance evaluation of all these forwarding schemes given in [50], the best trade-off in terms of data traffic overhead, node coverage ratio and dissemination delay is given by CBF.

The CBF data forwarding algorithm [27] assumes that every vehicle has a local data base (e.g., an LDM) that maintains basic information about 1-hop neighboring vehicles (i.e., by means of a CAM/BSM exchange process). The algorithm itself is based on timers triggered at the reception of a message. An illustrative representation of the CBF operation is given in Figure 2.4. Suppose a generic vehicle sends in

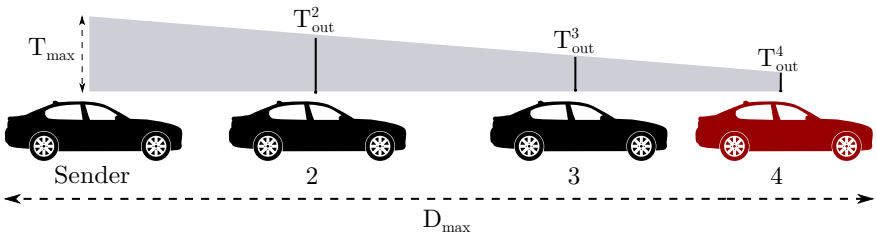


Figure 2.4 – CBF algorithm operation.

broadcast a message. All the vehicles that receive this message compute a local timeout based on the distance from the sender, as shown in Equation (2.3).

$$T_{\text{out}} = \begin{cases} T_{\text{max}} + \frac{T_{\text{min}} - T_{\text{max}}}{D_{\text{max}}} D & \text{if } D \leq D_{\text{max}} \\ T_{\text{min}} & \text{if } D > D_{\text{max}} \end{cases} \tag{2.3}$$

Here T_{out} represents the timeout, T_{min} and T_{max} are the minimum and maximum duration a message should be buffered, D is the distance between the sending and receiving vehicles, and D_{max} is a parameter denoting the theoretical maximum communication range of the wireless access technology. According to this timer setting, vehicles having a larger distance from the sender will have a shorter timeout. When the timeout expires, the message is retransmitted in broadcast. If a vehicle waiting for its timeout to expire receives a second copy of the message, then it cancels the timer and discards the message.

2.4 Simulation Framework

When studying FCD collection protocols, as well as ITS applications in general, one of the main challenges is to properly evaluate the proposed solutions. Since large scale experiments are often infeasible, the research community rely on computer simulations. In this context, having realistic simulation tools is fundamental.

From an ITS perspective, there are two main aspects that have to be considered: simulation of the vehicular mobility [51] and simulation of the vehicular networking [52]. Moreover, the dynamic interaction between these two sides is most of the time very important. There is a large number of available VANET simulation models and tools in the literature that combine together vehicular mobility and networking [52]. However, different simulators provide different features with various degree of realism. The simulation framework that we chose to use is represented in Figure 2.5 and has three main software components: Objective Modular Network Testbed in

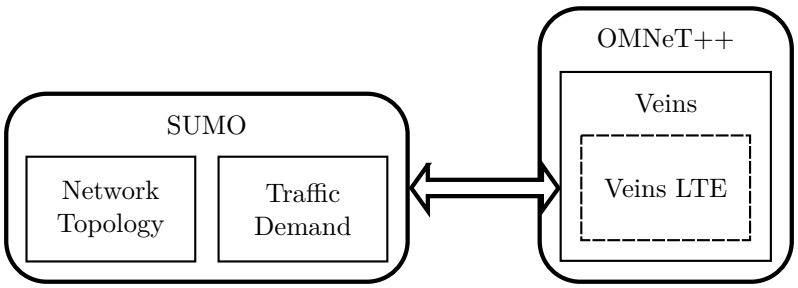


Figure 2.5 – An overview of the simulation framework.

C++ (OMNeT++)[53], Simulation of Urban MObility (SUMO) [54], and Vehicles in Network Simulation (Veins) [55].

OMNeT++ is an open-source, component-based C++ simulation framework that is normally used to build network simulators. The main advantages of OMNeT++ are that it is extensible and modular, allowing users to easily use and/or modify existing modules and libraries, or implement new ones. The main architecture is implemented in C++, while the network topology is written in Network Description Language (NED). OMNeT++ also provides an extensive GUI support.

SUMO is an open-source micro-mobility road traffic simulation suite, which allows the modeling and interaction between different traffic systems including road vehicles, public transport and pedestrians. It is a microscopic simulator in the sense that each vehicle is modeled explicitly. SUMO provides a broad range of supporting tools for network import and demand modeling. The road networks are represented as graphs, where intersections are the nodes of the graph and roads are modeled as edges. SUMO also includes traffic lights and connections between roads at intersections. The simulator allows to either generate road networks manually, or to import them from different formats, such as VISUM[4], OSM[5], and XML-Descriptions. Vehicles' speed in SUMO is computed according to the so-called *car-following* models. The default model is an extension of the Krauss car-following model [56], which is the one that we will use throughout this thesis. SUMO also provides the possibility to interact with external applications via an API called TraCI.

Veins is an open-source simulation framework for vehicular networking. It's main purpose is to provide a basic platform for writing and simulating vehicular networking related applications. Veins couples the SUMO and OMNeT++ simulators in a bidirectional manner. This means that vehicular mobility can affect the networking part, but also the vehicular networking can change the mobility of vehicles. Veins contains different simulation models related to vehicular networks. In this thesis we make use of the lower layers of the IEEE 802.11p/WAVE protocol stack implemented in Veins, and build our own applications on top of that. Veins also provides an extension for simulating heterogeneous LTE-DSRC vehicular networks. This extended version, named Veins LTE [57], integrates Veins with SimuLTE [44], a simulation framework for LTE networks.

2.5 Related Work

DSRC has been proposed as the main technology for Inter-Vehicle Communication (IVC). The primary motivation is to ensure safety on the roads by enabling V2V communication and cooperative awareness. The latter is usually obtained

[4] `http://vision-traffic.ptvgroup.com/en-us/products/ptv-visum/`
[5] `http://www.openstreetmap.org`

through periodic exchange of beacon messages (e.g., CAM, BSM). Among the advantages that DSRC offers we can identify dedicated spectrum, low message delays, decentralized architecture, and localized network load. However, to support non-safety applications, DSRC needs additional hardware and infrastructure deployment, like RSUs. Moreover, the technology currently is not yet widely deployed, meaning that at least in the initial stage DSRC needs to be supported by other existing communication technologies.

There are only few FCD collection algorithms in the literature that are based only on the DSRC technology. Brik et al. [58] propose a Token-based Clustered Data Gathering Protocol (TCDGP). This protocol divides the interested road in two types of segments: *collection segments*, where the V2V communication is allowed, and *silent segments*, where V2V communication is prohibited in order to avoid collisions between adjacent segments. A cluster head vehicle is elected for every collection segment, considering the vehicle's distance from the center of the segment and the time traveled on the middle lane. The cluster head vehicle is in charge of collecting data inside the collection segment. The downside part of this solution is that it was designed to collect data only on highways and requires the knowledge of the road network topology and segmentation beforehand. A similar approach to TCDGP was proposed by Chang et al. [59], who introduce TrafficGather. This protocol also divides the road into a series of contiguous clusters and for each cluster it elects a cluster head vehicle. Once the clusters are created, each vehicle sends its information to its own cluster head according to a TDMA Access Control Algorithm (TACA) proposed by the authors, in order to avoid collisions. In the data retrieval phase the cluster head vehicles send their collected data towards the initiating vehicle using a flooding strategy, which generates a very large overhead.

A different approach for data collection is found in [60], where the authors propose the Adaptive Data Collection Protocol using Reinforcement Learning (ADOPEL). This protocol is based on a data collection technique designed to collect data on highways while making the collection operation more reactive to nodes mobility and topology changes. ADOPEL is based on a distributed Qlearning technique, where a reward function is provided and defined to take into account the delay and the number of packets to aggregate. The collect operation is periodically started by a randomly selected node called *initiator*, which has to collect the traffic data from vehicles and deliver it to a remote traffic control center. The selection of the best next relay is based on the Qvalue, determined by the Qlearning algorithm. Zhu et al. [61] face the problem of delay-constrained data aggregation in VANETs. They propose a centralized and a distributed approach for their protocol named *aTree*. Their solution first constructs a data aggregation tree using a flooding approach, and then assigns a waiting time budget for each node on the tree. Their basic idea is to assign larger waiting timers to nodes closer to the collection node, in this way allowing child

nodes to send their information earlier. Nodes aggregate all the data received from their child nodes in the tree before transmitting it towards the collection node.

All the above mentioned algorithms are either designed for very specific scenarios, or they make use of flooding-based approaches, which are known to be the main cause for the broadcast storm problem [49]. What we aim for is a general FCD collection algorithm that operates independently from the considered scenario, especially in the most challenging ones, like urban areas, and that requires no a priory information about the road network topology. In Section 2.3 we described a standardized solution for disseminating data in VANETs, named CBF, that handles the broadcast storm problem by inhibiting most of the vehicles and selecting only a subset of them to forward the information. The FCD collection algorithm that we propose in Chapter 4 exploits this protocol and the ad-hoc structure formed by these forwarding vehicles to periodically collect FCD messages using only the DSRC technology.

LTE has been identified as another potential access technology able to support vehicular communications [62]–[64]. There are several reasons why LTE is suitable. First of all, it has the benefit of an already pre-deployed infrastructure, which offers wide area coverage and supports high mobility. Secondly, the market penetration rate of LTE is expected to be higher compared to other communication technologies, since the LTE technology is already integrated in common user devices, like smartphones, tablets, and smartwatches. Moreover, many vehicular applications can migrate to these devices. Araniti et al. [62] provide an extensive survey on the state of the art of LTE and its capability to support vehicular applications. Mangel et al. [63] analyze the usability of LTE for vehicular safety communication at intersections, comparing them with DSRC. They conclude that even if LTE seems able to support periodic delivery of beacon messages, its performance in terms of awareness update rate and latency is inferior with respect to DSRC. On the other hand, the latency and reliability requirements are not so strict for non-safety applications. Yet, the information generated by a high number of vehicles can heavily load the uplink channel, preventing the normal operation of traditional human-to-human traffic. Ide et al. [64] propose a channel sensitive probabilistic transmission scheme in order to reduce the LTE channel load. Their algorithm reduces the number of forwarders, but does not guarantee an exhaustive collection of data.

However, LTE technology has several drawbacks. First of all, it operates in a licensed spectrum, meaning that its performance and availability is highly dependent on the mobile and network operators. Also, in high density urban scenarios the periodic data transmissions from many vehicles can use a significant part of the LTE channels, possibly degrading the normal operation of traditional applications. In order to support the increasing amount of data traffic, LTE needs further upgrades,

like decreasing the cell sizes, or adding more spectrum [65]. All these upgrades are not for free, requiring additional investments from the network operators.

To cope with the limitations that both LTE and DSRC have, the research community is shifting towards heterogeneous vehicular networking approaches [23]–[25], [66]–[69]. The idea is to deploy both technologies to vehicles and road, and to exploit the benefits from each technology. A common paradigm is to use the cooperative awareness enabled by DSRC to create clusters of vehicles having common features (e.g., proximity, travel direction, speed, connectivity), and to designate cluster head vehicles in charge of aggregating the information inside their local clusters and sending it via LTE to a remote central unit. A complete taxonomy on clustering in vehicular networks is proposed by Bali et al. [70]. They provide a comprehensive analysis of existing proposals in literature, as well as a detailed discussion for each category of clustering, including challenges and future directions.

Remy et al. [68] propose a framework that uses a centralized clustering mechanism to collect the FCD information. In particular, eNodeBs are responsible to organize vehicles into clusters, to elect cluster heads, and broadcast the clusters topology to the vehicles. According to this solution, each cluster head vehicle collects data inside its own cluster using V2V communication, aggregates this information and sends it to the eNodeB. A similar approach is adopted by Jia et al. [71], who study the impact of the FCD collection in an LTE network. The proposed cluster head selection process is managed by the eNodeB. The authors show that such a system is able to reduce the negative impact of FCD load on the quality of the transport service obtained by conventional LTE traffic. D'Orey et al. [72] propose a centralized system for creating clusters and for electing cluster heads. The clustering process is performed here by a remote server, assuming it to have a much wider regional view of the system, when compared with the limited scope available to a single eNodeB. The main issue with the three previously mentioned solutions is that they rely on an initial centralized phase where the whole vision of the network has to be known to a central server, which will then be in charge of creating clusters and selecting the cluster heads. This means that in the initial phase every vehicle has to communicate its own information via LTE. We propose a possible solution to this problem in Section 5.1, where we exploit the VANET already in the initial setup phase by selecting a subset of vehicles in charge of sending the information from the entire VANET. The selection process is based on the standard CBF algorithm described in Section 2.3.

Non-safety applications usually require periodic collection of data from vehicles inside a target area. Various applications have different requirements in terms of accuracy of the reported information. For instance, Ide et al. [66] focus on a traffic forecast application where neighboring vehicles have similar information. Based on this assumption, elected forwarding vehicles perform local aggregation and

compression before sending the information to the remote server via LTE. The upper bound of the amount of compressed data is modeled as a square root function of the number of uncompressed data units. However, in many non-safety applications the information cannot be compressed, meaning that data from every single vehicle must be gathered. In this case, which is also the case that we consider in this thesis, the aggregation consists of concatenating the payloads gathered from the DSRC neighboring vehicles. El Mouna Zhioua et al. [67] propose the Fuzzy QoS-balancing Gateway Selection (FQGwS) algorithm for choosing gateway vehicles to communicate with the LTE infrastructure. However, this solution assumes that the vehicles are already organized in clusters, and cluster head vehicles are already elected. Bazzi et al. [69] also consider the DSRC technology to offload data from LTE. Their protocol is based on the assumption that RSUs are installed around the map and their position is known to all vehicles. The algorithm idea is that every vehicle identifies among its DSRC neighbors the one closest to an RSU to whom it sends its own information.

The target application has a strong impact on the decision of what parameters to consider in the clustering mechanism. Many applications aim at obtaining cluster stability, meaning that the vehicles' position, speed, and driving direction are the most critical parameters. Other applications focus on minimizing the LTE channel utilization while periodically collecting data from vehicles. In this case DSRC connectivity becomes predominant, since the main objective is to collect data from the whole vehicle network, while minimizing the number of forwarders and maximizing the local aggregation.

Stanica et al. [24] identify this as to be equivalent to the Minimum Dominating Set problem in graph theory, known to be NP-complete. They propose three heuristics for the election of Cluster Head vehicles in a heterogeneous LTE-DSRC vehicular network: Degree-Based (*DB*), Degree-Based with Confirmation (*DB-C*), and Reservation-Based (*RB*). *DB* basically uses the safety beacons exchanged over the DSRC network to compute the number of neighbors for each vehicle. This information is used by the cluster head election mechanism, in which a vehicle becomes a cluster head with a probability equal to k/D, where D is the number of neighbors and k is a parameter for the trade-off between coverage and offloading gain. Although it is very simple, this algorithm does not provide any guarantee on the coverage of the entire area. *DB-C* copes with this issue by extending the previous approach with a simple confirmation mechanism in order to obtain the total coverage. The idea is that whenever a vehicle chooses to be a cluster head it informs its neighbors by broadcasting a notification message. If during a collection period a vehicle does not become a cluster head and does not receive any notification message, it deems to be disconnected and sends its own information via LTE. With *RB* each vehicle, at the beginning of every collection period, selects a transmission slot among N_s available

and waits for its slot to transmit. Whenever a vehicle transmits the data on LTE, it becomes *dominator* and informs its neighbors, who cancel their waiting times and become *dominated*. A crucial point in this mechanism is the choice of N_s.

These algorithms are evaluated in terms of *system gain*, defined as the fraction of vehicles that do not have to access the cellular infrastructure when data is offloaded through DSRC communication. However, this metric does not directly measure the utilization of the LTE channel. In this article we are actually focusing on measuring the RB utilization in the LTE network. Moreover, Stanica et al. [24] assume a simple unit disk model for IVC connectivity, where two vehicles can communicate whenever their distance is below a threshold R, which is a non-realistic assumption. Also, most of the heuristics presented above are trying to minimize the number of forwarders by relying only on the DSRC connectivity parameter. In Section 5.2 we prove such choice to be suboptimal and propose two heterogeneous LTE-DSRC distributed FCD collection algorithms that use parameters drawn from both communication technologies in the cluster head selection mechanism.

Chapter 3

Minimizing the Age-of-Information in VANETs

Beaconing is a basic communication process taking place in VANETs to achieve cooperative awareness among vehicles on the road. It is actually a paradigm of information spreading among peer-agents, where each node of a network periodically sends broadcast messages containing information collected by the node itself. A trade-off arises between the update frequency of the broadcast information and the congestion induced in the wireless shared channel used to send the messages, which is based on the IEEE 802.11p standard in case of a VANET. For periodic updates, the primary metric is the AoI, i.e., the age of the latest update received by neighboring nodes.

Often, FCD collection algorithms rely on clustering mechanisms to reduce the load on the wireless communication channel. These mechanisms usually consist in the selection of a subset of vehicles that are in charge of sending the information contained in their local data bases towards a central collection unit. In this context, the AoI metric measures the freshness of the collected information contained in these local data bases.

In this chapter, which is based on our article published in *International Teletraffic Congress* [73], we define an analytical model to evaluate the AoI of a VANET, given the connectivity graph of the vehicles. Analytical results are compared to simulation to assess the accuracy of the model. The model provides a handy tool to optimize the AoI trade-off.

3.1 AoI Definition and Background

A common messaging paradigm in the context of the IoT consists of a network of peer-agents sending broadcast update messages to one another. These messages

carry state variables of sending nodes and/or information collected locally by sending nodes, which is spread to their neighbors. Each node maintains a map of the latest updates or possibly of the most updated processing results obtained from the stream of data it receives continually from its neighbors.

A specific major example of this communication paradigm is a VANET, where vehicle nodes exchange beacon messages to maintain cooperative awareness of one another. In this context the DSRC standard [22] has been defined, based on IEEE 802.11p PHY and MAC layers, fully compliant with ETSI ITS-G5 [34] standard. On top of this CSMA-based communication technology, safety and traffic information or efficiency applications have been defined, such as CAMs [36] and DENMs [37]. The former ones are sent out periodically by each equipped vehicle to inform about its type, position, direction of movement, speed and other optional features. The latter ones are intended to provide early warning of potentially critical road and traffic events, aiming ultimately to enhance travel safety.

Examples of message sending in different contexts include wireless sensor networks [74] (of whom the VANET is one major, special example), gossiping algorithms [75], [76], distributed consensus algorithms [77], [78], network of automata [79], and synchronization of coupled nodes through a network [80]. The common point of all these contexts is that some kind of dynamical process evolves over a network, possibly large and time-varying, formed by a set of peer-agents. The evolution of the process is tied to message passing and state updates among nodes. Most of those examples need periodically updated information on a time scale compatible with the application requirements.

A key point in the exchange of cooperative messages, as well as in general in the broadcast of update messages to neighbors in any network of distributed agents communicating through a shared channel, is to control the congestion level so as not to impair the regular and timely update of time-critical information. Reducing the sending rate of updating messages is the obvious control variable to avoid congestion, yet it leads to a smaller refresh rate of information. The trade-off between congestion of the communication channel and refresh rate of the information carried by the messages has brought to the definition of a specific metric, the AoI. Given a table at node i whose entries are updates from another node j, the AoI at time t of those updates is measured by $A_{ij}(t) = t - u_k(j \to i)$ for $t \in [u_k(j \to i), u_{k+1}(j \to i))$, where $u_k(j \to i)$ is the time when the k-th update from j is received by i. Whenever a new update is received, the AoI drops to zero; then it grows by 1 second per second of elapsed time until the next update. A sample path of the AoI metric is shown in Figure 3.1.

AoI has been addressed specifically in [81]–[83]. Kaul et al. [81] give a nice general model to evaluate the AoI of a population of information sources coupled to a central server by means of a limited capacity channel. The general abstract

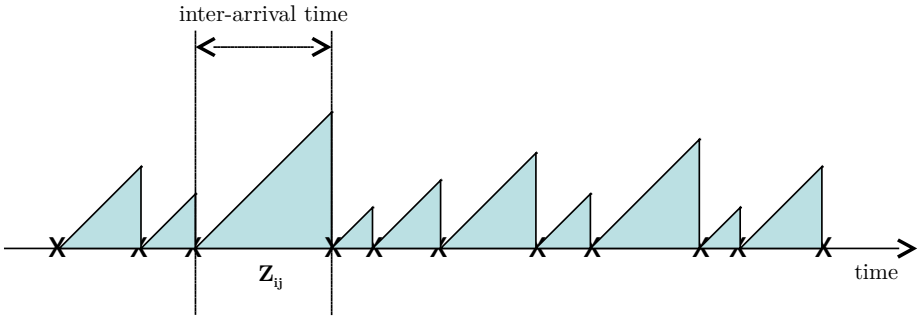


Figure 3.1 – A sample path of AoI [73].

model is able to highlight the basic trade-off involved in the AoI. Yet it refers to a single shared channel, not a distributed network of interacting agents. Kaul et al. [82] devise an adaptive beaconing strategy to minimize the age of information by balancing the load (hence the contention level) on the access wireless network and the frequency of updates. The addressed scenario is a vehicular network. Franco et al. [83] address the design of a cross-layer MAC protocol optimized to reduce the age of information in WLAN with high level of contention.

As for CSMA network modeling, since the classic model of IEEE 802.11 CSMA/CA of Bianchi [84], several extensions have been defined. Liew et al. [85] give a relatively simple model to evaluate the performance of CSMA networks with partial sensing. The model addresses saturated nodes. Besides, it implies the computation of Maximum Independent Sets (MISs) of the network graph, where an arc is introduced between nodes i and j if both i and j can hear each other. This computation becomes impractical for graphs with more than several hundred links. A general CSMA network model is presented in [86] to address both saturated and non saturated CSMA/CA networks with partial sensing. The analysis is elegant and extremely powerful. However, the continuous-time model defined in that work does not account for collisions. Moreover, it brings about the same unfeasibility problem as with [85], since it requires the computation of the MISs of the network graph.

Our contribution is twofold. On a methodological level, we aim at stating and assessing an analytical model that can capture the distributed contention in a CSMA network to calculate the AoI. On the application level, we investigate how the key parameters of the MAC layer and of the messaging protocol impact on the AoI in a VANET, by considering a realistic urban scenario describing the vehicular traffic within the city of Cologne (Germany), as well as an artificially created Manhattan Grid scenario, based on artificial traffic generation and communication.

3.2 Analytical Model for the AoI Evaluation

We focus on periodic one-hop message exchange, with nodes sending messages of given length L with period T_{msg}. New messages are accepted by the MAC layer entity of a node as long as it is idle. If the node is busy, the arriving new message is stored in a buffer. Further arriving messages are overwritten so that only the latest message is taken care of by the MAC layer, as soon as the previous message has been sent out on the channel. This setting is consistent with the periodic issuing of the beacon messages, each carrying an update of the vehicle information. Given this setting, only the latest update is worth being transmitted.

The model describes the generic node operation with a renewal process. Let us consider a tagged node A sending a message at time t_k, $k \in \mathbb{Z}$. We let $Y_k = t_k - t_{k-1}$. At equilibrium, we can assume $Y_k \sim Y$. Since the times $\{t_k\}_{k\in\mathbb{Z}}$ are regeneration points for the sending process, the sequence of intervals $\{Y_k\}$ forms a renewal process[6]. We can distinguish two cases according to whether the contention plus transmission time is less than the message inter-arrival time (top picture in Figure 3.2) or not (bottom plot on Figure 3.2).

Then

$$Y = B_2 + \max\{0, T_{\text{msg}} - B_1\} \tag{3.1}$$

where B is a random variable defined as the time elapsing from the moment when the MAC layer takes in charge a PDU until it eventually sends it out on the radio channel. B is the sum of the transmission time T (including the overhead) and of

[6] At least, this is true under the simplifying assumption of independence of the stations' states.

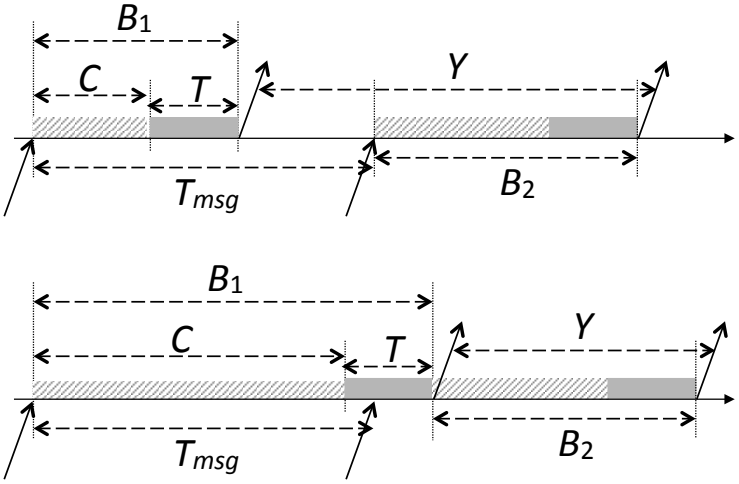


Figure 3.2 – The time interval elapsed from the moment when the message is generated by the application layer until it is sent on the radio channel [73].

the time C spent on counting until the back-off counter hits 0. If the channel is sensed busy, the counter is frozen until the channel activity terminates. Otherwise, the counter is decremented after a back-off time slot of duration δ. Let C denote the count-down time, defined as the sum of a number of "slot" times, each slot lasting either δ, the IEEE 802.11p back-off slot duration, or T, which is the time required to complete a MAC PDU transmission, including PHY and MAC overhead, and the ensuing DIFS[7]. Then

$$C = \sum_{j=1}^{N} X(j) \tag{3.2}$$

where N is a discrete random variable with uniform distribution over $[0, W_0 - 1]$, W_0 being the base contention window size of IEEE 802.11p, and $X(j)$ are i.i.d. random variables with the same distribution as X defined by

$$X = \begin{cases} \delta & w.p.\ 1-b \\ T & w.p.\ b \end{cases} \tag{3.3}$$

Here b is the probability that the tagged node senses the channel busy.

In the evaluation of the statistics of X and hence of C, we must account for the fact that only partial sensing is realized in general. In other words, while some neighbor N_1 of the tagged node A is transmitting, and hence the tagged node freezes its count-down state, some other neighbor N_2 of A could start its own transmission, in case N_2 is out of carrier sensing range of N_1, i.e., N_1 is hidden to N_2. The resulting effect as seen by the tagged node A is that its freezing time lasts *more than* T. This 'expanded' duration of the activity sensed on the channel by A depends on the maximum number of nodes that can start transmitting independently of one another, i.e., that do not sense one another. Let n_{MIS} denote the cardinality of the MIS around A. Once the transmission starts, up to $n_{\text{MIS}} - 1$ more transmissions could start with random phases. By assuming independence and uniform probability distribution of the relative phasing in $[0, T]$, it can be found easily that the time T is replaced by $T(2 - 1/n_{\text{MIS}})$. The expansion factor $\psi = 2 - 1/n_{\text{MIS}} \geq 1$ of the activity time reduces to 1 when $n_{\text{MIS}} = 1$, i.e., all neighbors of A do sense each other. A proxy of the number n_{MIS} that is easier to calculate is $\hat{n} = c + \nu(1-c)$, where ν is the number of neighbors of A and c is the clustering coefficient of A. This is simply a linear interpolation between $n_{\text{MIS}} = 1$ when $c = 1$ and $n_{\text{MIS}} = \nu$ when $c = 0$. The clustering coefficient of a graph node A is the ratio of the number of links among the ν neighbors of A divided by the maximum number of such links, i.e., $\nu(\nu-1)/2$. Given the adjacency matrix $\mathbf{A}$ of an undirected graph, the clustering coefficient of node i is $c_i = \ell_i / [\nu(\nu-1)/2]$, where ℓ_i can be found as the i-th element of the

[7] SIFS and ACK times are not included, since MAC PDUs for beaconing are sent in broadcast, hence no ACK is provided.

diagonal of the matrix $\mathbf{A}^3/2$. In the definition of X_i, the random variable X at node i, we therefore substitute T with $T\hat{\psi}_i$, where $\hat{\psi}_i = 2 - 1/\hat{n}_i = 2 - \frac{1}{c_i + \nu_i(1-c_i)}$, with $c_i = \frac{2\ell_i}{\nu_i(\nu_i-1)}$.

We have the following identities for the first two moments of $B_i = T + C_i$

$$\mathrm{E}[B_i] = T + \frac{W_0 - 1}{2}\mathrm{E}[X_i]$$
$$\sigma^2_{B_i} = \frac{W_0^2 - 1}{12}(\mathrm{E}[X_i])^2 + \frac{W_0 - 1}{2}\sigma^2_{X_i}$$

where

$$\mathrm{E}[X_i] = \delta(1 - b_i) + T\hat{\psi}_i b_i$$
$$\sigma^2_{X_i} = (T\hat{\psi}_i - \delta)^2 b_i(1 - b_i)$$

The first two moments of Y_i are found by considering all realizations β_k of the random variable B_i and the relevant probabilities, i.e., $p_i(k) \equiv \mathscr{P}(B_i = \beta_k)$. By definition:

$$\mathrm{E}[\big(\max\{0, T_{\mathrm{msg}} - B_i\}\big)^\gamma] = \sum_{k=0}^{W_0-1} p_i(k)\big(\max\{0, T_{\mathrm{msg}} - \beta_k\}\big)^\gamma$$

for $\gamma \geq 1$, and then Equation (3.1) yields

$$\mathrm{E}[Y_i] = \mathrm{E}[B_i] + \mathrm{E}[\max\{0, T_{\mathrm{msg}} - B_i\}]$$
$$Var(Y_i) = Var(B_i) + Var\left(\max\{0, T_{\mathrm{msg}} - B_i\}\right)$$

where, for $k = 0, \ldots, W_0 - 1.$, we have

$$p_i(k) = \mathscr{P}(B_i = \beta_k) = \frac{1}{W_0}\sum_{m=0}^{W_0-1-k}\binom{m+k}{k} b_i^k(1 - b_i)^m \tag{3.4}$$

and

$$\beta_k = T + W_0\delta + k(T\hat{\psi}_i - \delta) \tag{3.5}$$

The probability that the i-th node attempts a transmission on the channel is

$$\tau_i = \tau_0 \frac{\mathrm{E}[B_i]}{\mathrm{E}[Y_i]} \tag{3.6}$$

where τ_0 is the probability of attempting a transmission in a saturated CSMA/CA network, when binary exponential backoff is not used and only the basic contention window size is used. Hence, $\tau_0 = 2/(1 + W_0)$, with $W_0 = 15$, according to the IEEE 802.11p standard. Note that nodes *do not* operate necessarily in saturation,

since they are requested to send one message every T_{msg}. As long as $B_i < T_{\text{msg}}$ node i completes contention and message transmission before the next message is ready to send. This is the typical case for standard message periods (between 100 ms and 1000 ms), given that the contention time ranges between few ms and several tens of ms typically.

There remains to characterize the probability b. Let us introduce a subscript i for the tagged node. Let a_{ij} denote the entry (i, j) of the adjacency matrix $\mathbf{A}$ of the *carrier sensing* graph of the nodes. In words, $a_{ij} = 1$ if and only if node j can receive (detect) the signal emitted by node i. Since the radio channel is reciprocal, we can assume that $\mathbf{A}$ is symmetric. In this model, we assume that the carrier sensing matrix $\mathbf{A}$ is given (see Section 3.4).

As τ_j is the probability that node j is found transmitting, the probability that a neighbor node j of i is not transmitting is $1 - \tau_j a_{ji}$. We adopt the common independence assumption, whereby the states of the competing nodes in the CSMA network are assumed to be independent of one another. Then, the probability that node i senses an idle channel, i.e., that all its neighbors are silent, is[8]

$$1 - b_i = \prod_{j=1}^{n} (1 - \tau_j a_{ji}) \tag{3.7}$$

where n is the number of nodes in the network, hence the size of the adjacency matrix.

Summing up, the τ_i's can be found by solving a system of non-linear equations made up of Equations (3.3), (3.6) and (3.7). If we write $\tau \equiv [\tau_1 \ \tau_2 \ \ldots \ \tau_n]$, the equation system can be written in a compact form as $\tau = \mathbf{F}(\tau)$. The function $\mathbf{F}(\cdot)$ is continuous and maps the unit hypercube into itself. Hence, Brouwer's theorem guarantees that there exists a fixed point.

Once the transmission probabilities τ_i are computed, we can find the conditional probability of success, $P_s(i, j)$, of the event that node j receives a message from node i, given that i transmits the message. This amounts to node i transmitting and: (i) none of the neighbors of j being active at the same time; (ii) node j not transmitting as well. We can divide the neighbors of j into two sets:

$\mathscr{A}_{i,j}$ the set of neighbors of j that are also neighbors of i;

$\mathscr{B}_{i,j}$ the set of neighbors of j that are not neighbors of i.

The nodes belonging to the first set are synchronized by the activity of i, while the other nodes are not. Therefore, the transmission probability for node $k \in \mathscr{A}_{i,j}$ is τ_k. Nodes in $\mathscr{B}_{i,j}$ are outside the communication range of i, hence they are hidden with respect to i. We assume they are completely de-synchronized with i, hence

[8] Note that we define $a_{ii} = 0$.

node $k \in \mathscr{B}_{i,j}$ can start transmitting in any slot time of duration δ with probability $\delta/\mathrm{E}[Y_k]$. The vulnerability interval of the message sent by node i to node j comprises $m \equiv 2T/\delta - 1$ slot times. Therefore

$$P_s(i,j) = (1-\tau_j) \prod_{k \in \mathscr{A}_{i,j}} \left(1-\tau_k a_{kj}\right) \prod_{k \in \mathscr{B}_{i,j}} \left(1 - \frac{\delta}{\mathrm{E}[Y_k]} a_{kj}\right)^m$$

for all $j \neq i$. The time Z_{ij} to deliver a new message from i to j is given by

$$Z_{ij} = \sum_{r=1}^{N_{ij}} Y_i(r) \tag{3.8}$$

where $Y_i(r) \sim Y_i$ are the times between successive transmission attempts of node i, Y_i is given in Equation (3.1), and N_{ij} is the number of attempts required to make a successful message transfer from i to j. Assuming that successive attempt outcomes are independent of one another, N_{ij} has a geometric probability distribution, i.e.

$$P(N_{ij} = h) = P_s(i,j)[1-P_s(i,j)]^{h-1} \tag{3.9}$$

for $h \geq 1$. The AoI at node j for messages coming from i equals $t - t_{ij}(k)$ for $t \in [t_{ij}(k), t_{ij}(k) + Z_{ij})$, where $t_{ij}(k)$ is the time of arrival of the k-th message from i to j.

The mean value of the AoI from i to j, H_{ij}, is akin to the mean remaining service time in a queue, i.e.

$$\mathrm{E}[H_{ij}] = \frac{\mathrm{E}[Z_{ij}^2]}{2\mathrm{E}[Z_{ij}]} \tag{3.10}$$

It is

$$\begin{aligned}
\mathrm{E}[Z_{ij}^2] &= \mathrm{E}[N_{ij}(N_{ij}-1)](\mathrm{E}[Y_i])^2 + \mathrm{E}[N_{ij}]\mathrm{E}[Y_i^2] \\
&= \frac{2[1-P_s(i,j)]}{P_s(i,j)^2}(\mathrm{E}[Y_i])^2 + \frac{1}{P_s(i,j)}\mathrm{E}[Y_i^2] \\
&= \frac{2-P_s(i,j)}{P_s(i,j)^2}(\mathrm{E}[Y_i])^2 + \frac{1}{P_s(i,j)}\sigma_{Y_i}^2 \\
\mathrm{E}[Z_{ij}] &= \frac{1}{P_s(i,j)}\mathrm{E}[Y_i]
\end{aligned}$$

The expressions above allow to compute the mean AoI of messages flowing from i to j. The AoI at j can be obtained by averaging over all neighbor nodes of j, if any. If j is isolated, it receives no message actually, so AoI is meaningless. Besides this

marginal case, we can define

$$\mathrm{E}[H_j] = \frac{\sum_{i=1}^{n} a_{ij}\mathrm{E}[H_{ij}]}{\sum_{i=1}^{n} a_{ij}} \tag{3.11}$$

provide that $\nu_j \equiv \sum_{i=1}^{n} a_{ij} > 0$. The overall average AoI of the entire network can be summarized by the following definition:

$$\mathrm{E}[H] = \sum_{j=1}^{n} \frac{\nu_j}{\nu}\mathrm{E}[H_j] = \frac{1}{\nu}\sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}\mathrm{E}[H_{ij}] \tag{3.12}$$

where $\nu = \nu_1 + \ldots + \nu_n = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}$.

3.3 Simulation Model for the AoI Evaluation

To validate the proposed analytical model, we compare its performance with a realistic simulation model of a VANET. The main roles of the simulations set is to test the key simplifying assumptions of the analytical model, namely:

- independence of the node states, used to derive the message delivery success probabilities;
- vehicle mobility, not accounted for in the analytical model;
- details of the MAC protocol in a partial sensing environment.

In particular, we consider a set of two simulation scenarios. The first one is an artificially generated Manhattan Grid scenario, created using realistic road lengths and building dimensions taken from downtown Manhattan. This scenario consists of vertical roads representing main avenues, each road having a total of 4 lanes (2 lanes per direction), and of horizontal roads representing secondary streets, each street having a total of 2 lanes (one lane per direction). The distance between 2 junctions on the horizontal and vertical roads is of 275 m and 80 m respectively. Parallel roads are separated by buildings obstructing the inter-vehicle DSRC-based communication. Both vehicular mobility and networking are simulated over a larger area, but the observed region is smaller in order to avoid border effects. In particular, the target area is enclosed in a 620 m × 530 m region situated in the center of our simulated scenario. This area contains 3 vertical and 7 horizontal roads as described above.

The second simulation scenario that we consider is based on the TapasCologne[9] [87] vehicular mobility dataset (see Figure 3.3), which covers a region of 400 km^2 in

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

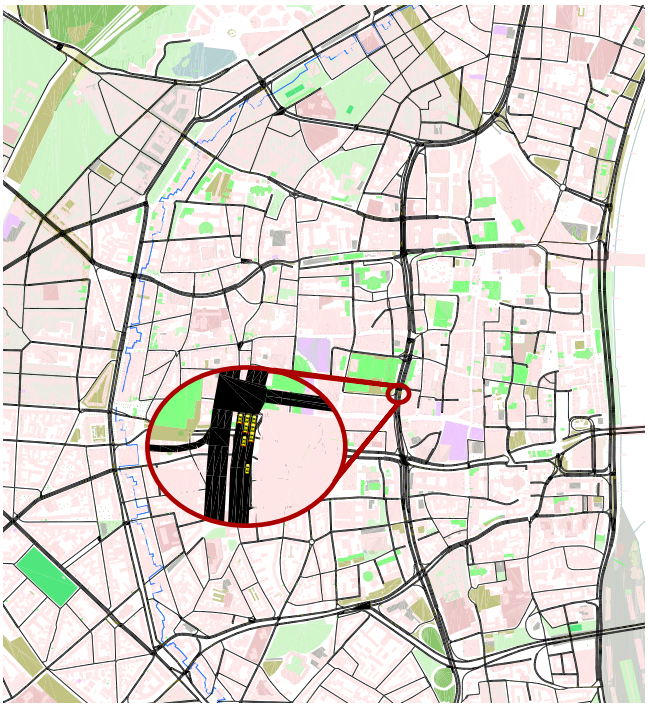


Figure 3.3 – Cologne simulation scenario [73].

the city of Cologne, Germany, and reproduces with a high level of realism the vehicular traffic for a period of 24 h. For our simulations, we delimited a 2.16 km × 2.43 km target area situated in the center of Cologne city, represented in Figure 3.3.

In our study we use the simulation framework described in Section 2.4 composed of SUMO [54], OMNeT++ [53] and Veins [55]. Vehicles are following the Krauss vehicular mobility model and the random trips traffic flow origin-destination model. The mobility is generated with a fringe factor equal to 10, meaning that it is 10 times more likely that the trips will start/end at the fringe of the simulated scenario. In this way, we model our vehicular traffic to start and end outside of the target area.

We assume that all vehicles have DSRC technology on-board. IEEE 802.11p parameters are considered for MAC and PHY. Two attenuation models are used: the free-space path loss with $\alpha = 2$, and the simple obstacle shadowing [88] to model the impact of buildings on signal propagation. The main simulation parameters can be found in Table 3.1.

At the application level, a simple beacon exchange mechanism is implemented, where every vehicle, periodically with period T_{msg} and independently, broadcasts a beacon message containing basic information, like identification, position, velocity, timestamp, etc. T_{msg} is a global parameter known to all vehicles. Also, every vehicle maintains a Local Data Base (LDB) (i.e., simulating an LDM) where the information

Parameter	Value
Manhattan Grid target area	620 m × 530 m
Cologne target area	2.16 km × 2.43 km
Manhattan Grid density (veh/km/lane)	33
Cologne density (veh/km^2)	95
IVC technology	IEEE 802.11p
IVC maximal transmit power (mW)	20
DSRC beacon frequency (ms)	100,200,300,500,1000
DSRC bitrate (Mbit/s)	3
Payload length (B)	1000

Table 3.1 – Simulation parameters

from the incoming beacon messages is stored. In a separate data structure, the last arrival times for every neighboring vehicle are saved, so as to be able to compute the time between two consecutive receptions of a beacon message from the same neighbor.

In the simulation vehicles enter the ROI in the considered urban map, roam in the ROI, then eventually they leave it. We focus on an observation time interval $\mathscr{I} = [t_0 - \Delta/2, t_0 + \Delta/2]$, where t_0 is a generic time of the statistical equilibrium regime of the simulation. Each vehicle collects messages coming from its neighbors. Let K_{ij} be the set of indices of messages originated by node i and received by node j during the interval I. Let further $\Delta t_{ij}(r)$ be the time interval spanning between the reception of the r and the $(r+1)$-th message from i at j, for $r \in K_{ij}$. Then, the estimate of the average AoI of messages from j at node i is given by (see Figure 3.1)

$$\hat{H}_{ij} = \frac{\frac{1}{2}\sum_{r\in K_{ij}} \Delta t_{ij}^2(r)}{\sum_{r\in K_{ij}} \Delta t_{ij}(r)} \tag{3.13}$$

Let $\Delta_{ij} \equiv \sum_{r\in K_{ij}} \Delta t_{ij}(r)$. The overall average of the AoI can be obtained by averaging the $\hat{H}_{ij}$'s, weighted by the fraction of the observation time when the messages have been collected, namely

$$\hat{H} = \sum_{i=1}^{N(I)} \sum_{j=1}^{N(I)} \frac{\Delta_{ij}}{\Delta} \hat{H}_{ij} \tag{3.14}$$

where $N(I)$ is the number of vehicle seen roaming in the ROI during the a time interval I in the statistical equilibrium regime. Putting together Equations (3.13) and (3.14), we can write

$$\hat{H} = \frac{1}{2\Delta} \sum_{k\in M(I)} \Delta t_k^2 \tag{3.15}$$

where $M(I)$ is the set of all messages received by some vehicle during the interval I.

3.4 Numerical Results

For the validation against simulations, the analytical model takes in input the network connectivity graph based on the carrier sensing, which we obtain from simulations by taking snapshots over time of the vehicular network. In particular, we obtain the adjacency matrix **A** from a short simulation of the beacon exchange process, where every vehicle is sending in broadcast a beacon every T_{msg}. This allows every vehicle to build its own LDB, which we then use to build the network connectivity graph (i.e., the adjacency matrix **A**). By doing so, we make sure that **A** is obtained by accounting for the radio channel model built into the simulator and described in Section 3.3. The values of AoI obtained from simulations are computed according to the Equation (3.15).

In Figure 3.4 we show the average AoI when varying the sending period for the Cologne and Manhattan Grid scenarios respectively. We can see that the results

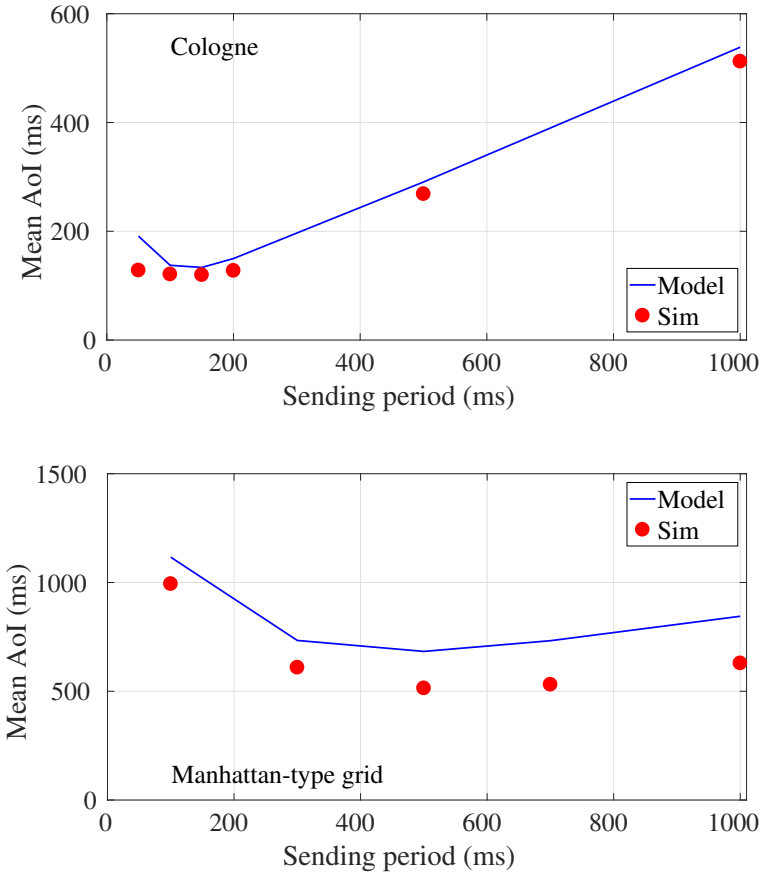


Figure 3.4 – Mean Age-of-Information with respect to sending period – analytical model (Model) vs simulation (Sim) [73].

obtained with the analytical model are close to the simulation results, meaning that the proposed model is able to capture and approximate quite well how the average AoI changes with the sending period. The model yields a less accurate upper bound for the Manhattan Grid scenario and for the lowest value of the sending period in case of the Cologne scenario. In both cases, the model provides an upper bound of the actual performance anyway. An important observation is that the model captures the optimal level of the message sending interval, which is around 150 ms for Cologne and 500 ms for Manhattan Grid. When we depart from these optimal levels, the AoI starts increasing. Moreover, even at the optimal sending interval, the AoI is quite higher than the ideal level $T_{\text{msg}}/2$.

This phenomenon can be explained by the fact that for such low intervals the channel is highly congested, as can be seen from Figure 3.5, which leads to a higher

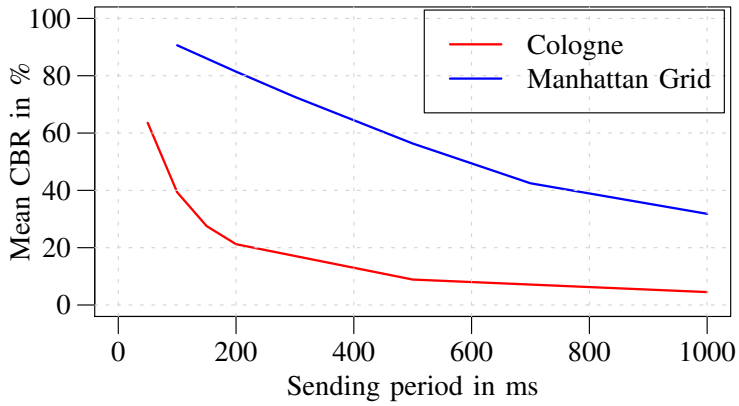


Figure 3.5 – Mean channel busy ratio with respect to sending period [73].

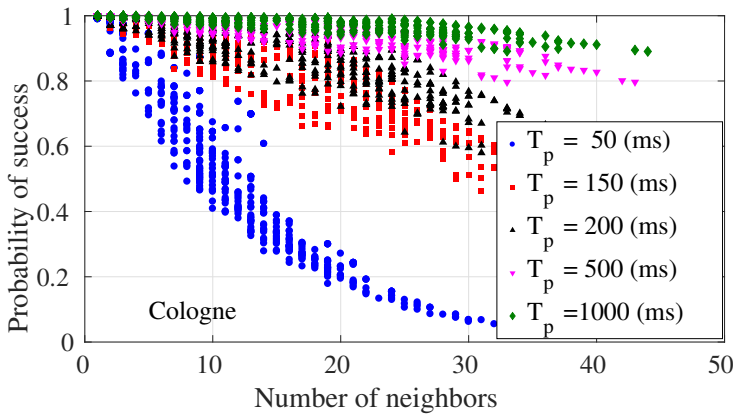


Figure 3.6 – Probability of successful message delivery with respect to the number of DSRC neighbors for different sending periods [73].

message loss ratio and to the fact that a backlogged node has to wait for a much longer time to sense the channel idle. Message loss impacts strongly the AoI, leading to variable and stochastically high gaps between received updates. On the other side, when the sending period grows, the contention on the wireless channel is relieved, but then the AoI starts increasing again because of the large time between the sending of successive updates. Figure 3.5 also shows that the Manhattan Grid scenario is much more congested than Cologne.

Figure 3.6 is a scatter plot of the average probability of successful message delivery for node i, namely $\overline{P}_s(i) = \sum_j P_s(i,j)a_{ij}/\nu_i$, as a function of the number of neighbors ν_i. It is apparent that the average success probability depends on how crowded the node neighborhood is, which is directly related to the air interface congestion. For longer message sending periods (e.g., $T_{\text{msg}} = 1000\,\text{ms}$) the dependence of $\overline{P}_s(i)$ versus ν_i is weak, whereas a wide range of levels of $\overline{P}_s(i)$ can be observed for shorter periods. This points out that the local effectiveness of message refreshing becomes critically dependent on the local vehicle density as the message sending period is decreased, i.e., locally different performance can be experienced by vehicles. This finding is consistent with the observations in [89] and [90], where the authors propose adaptive beaconing solutions to cope with variable vehicle density.

Figure 3.7 shows the (net) average throughput of a node, that is to say the average amount of data delivered by a node to its neighboring nodes successfully. This is simply $\overline{P}_s(i)/\text{E}[Y_i]$. It is apparent that the more the neighbors of a node, the less the amount of throughput that the node can sustain. This is consistent with the intuition that in crowded network spots the high level of contention hinders the possibility of delivering update messages to neighboring nodes. This is strongly amplified according to the level of T_{msg}. As the message sending frequency $1/T_{\text{msg}}$

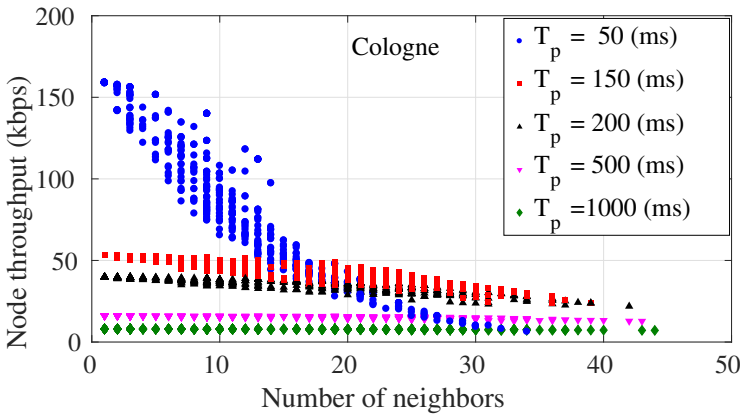


Figure 3.7 – Average net throughput of a node with respect to the number of DSRC neighbors for different sending periods [73].

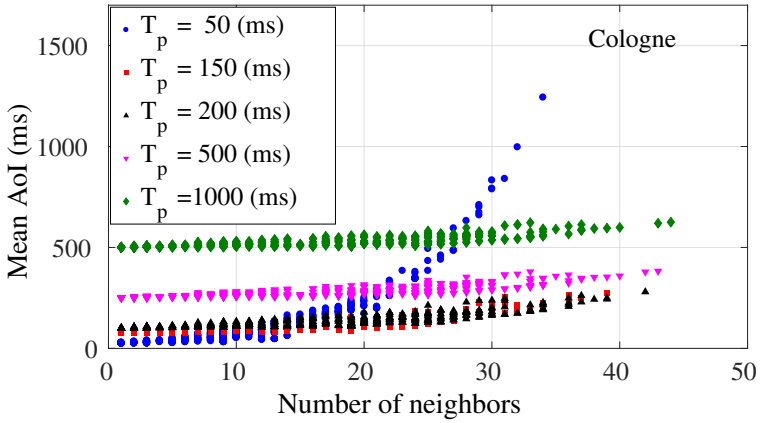


Figure 3.8 – Average AoI of nodes with respect to the number of DSRC neighbors for different sending periods [73].

grows, the throughput levels change from essentially constant with the number of neighbors to extremely sensitive to the number of neighbors (e.g., see $T_{\text{msg}} = 50\,\text{ms}$, where the throughput drops by more than one order of magnitude from low to high number of neighbors). Increasing the message sending frequency triggers an increasing level of unfairness in the message throughput performance experienced by nodes as a function of the local vehicle density, The best comprise is achieved around a message sending period between 150 ms and 200 ms.

Finally, Figure 3.8 shows the average AoI per node $\text{E}[H_i]$ against the number of neighboring nodes. In general, $\text{E}[H_i]$ grows with the number of neighbors. While the dependence is weak for large levels of T_{msg}, it gets dramatic for the lowest considered level of the message sending period, where $\text{E}[H_i]$ spans two orders of magnitude. Thus, maintaining balanced performance among different nodes requires avoiding too small levels of T_{msg}. However, large levels of T_{msg} entail a large AoI anyway, due to the sporadic refresh of information. The best compromise between 'stable' performance of nodes, irrespective of the local vehicle density and small AoI levels, is achieved for T_{msg} ranging between 150 ms and 200 ms.

A spatial representation of the per-node AoI $\text{E}[H_i]$ is shown in Figure 3.9. Dots correspond to vehicles and are scattered according to their registered position at the observation time in the simulation. The average per-node AoI calculated for each node according to the model is shown by using a heat colormap labelled with AoI levels in milliseconds. From the layout of the nodes, it is apparent that hot spots (i.e., zones where the AoI level experienced by nodes is high) are located mainly at road crossings.

The model can be used to gain insight into the effect of T_{msg} as well as other system parameters. Figure 3.10 shows the mean Age-of-Information with respect to

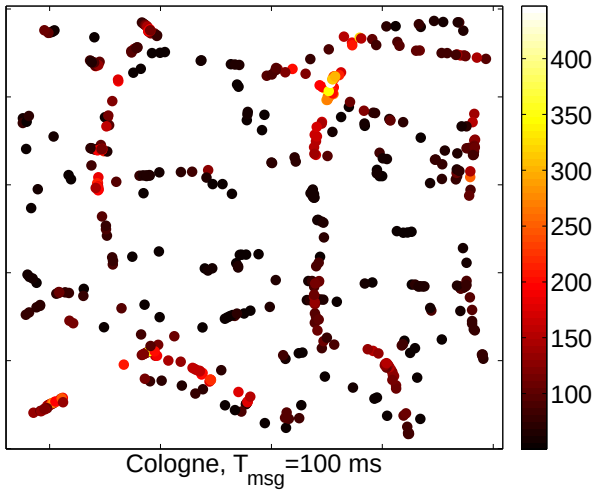


Figure 3.9 – Average AoI of nodes represented with a heat colormap (scale in milliseconds) [73].

sending period for different packet lengths L. The best operating point, i.e., the one minimizing the AoI, grows substantially as the packet length is increased. It is about 20 ms for $L = 100$ B, while for $L = 2000$ B the minimum AoI is attained around $T_{\text{msg}} \sim 500$ ms. As the message sending frequency increases, the AoI performance tends to saturate. This corresponds to the fact that the DSRC air interface gets saturated and the MAC level performance becomes essentially independent of the message sending period (i.e., nodes always have a new message to send).

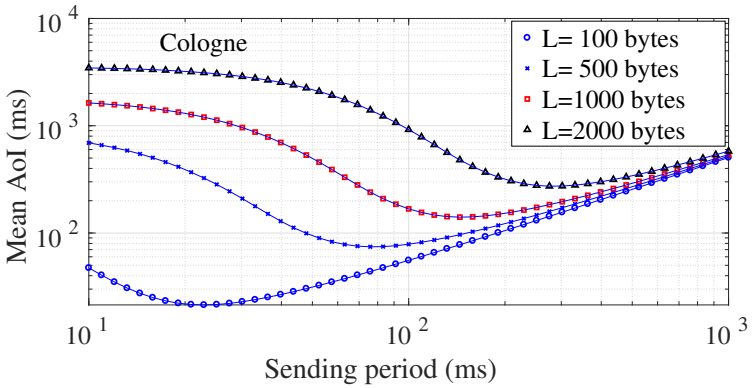


Figure 3.10 – Mean Age-of-Information with respect to sending period for different packet lengths[73].

3.5 Conclusion

In this chapter we proposed an analytical model to evaluate the Age-of-Information of a VANET, defined as the mean age of the latest update received by neighboring nodes, given the network connectivity graph of the vehicles. We validated the proposed solution by comparing it with realistic simulations of an urban area representing an area of the Cologne city and a generic Manhattan Grid scenario. Our results show that the analytical model is able to capture and approximate quite well how the average AoI changes with respect to the beacon sending frequency. It points out that an unfairness problem exists, i.e., AoI is directly related to vehicle density. Increasing the message sending period is beneficial for those vehicles that move in sparser zones, whereas it can degrade AoI strongly if used for vehicles in crowded areas. The analytical model can be used to set the sending period so as to minimize the AoI metric. It could also be used to guide the development of adaptive beaconing algorithms.

Further work could address the limitations of the analytical model, i.e., its inaccuracy as the mean vehicle density grows. The key point that calls for more investigation is capturing the effect of partial sensing, which is done in a simple, yet coarse way in our model by exploiting the clustering coefficient. A second point is to understand whether the relevant node connectivity information for the evaluation of the AoI can be reduced to some global property of the carrier sensing graph, rather than to the detailed adjacency matrix.

Chapter 4

DSRC-based FCD Collection in Vehicular Networks

Two key enablers of ITS services that can be supported by a VANET are the data dissemination and the data collection from vehicles on the road. These two features are essential to fully enable the advent of ITS and autonomous vehicles. Data dissemination can be achieved through the use of V2V DSRC-based multi-hop communications, enabling the extension of the road span covered by the RSUs or On-Board Units (OBUs) generating the data. This dissemination function is of interest for both safety and infotainment applications [91]. Another interesting function is the collection of data from vehicles, through the VANET. Vehicles can be used as sensors that monitor traffic, roads, the environment and send their data to a collection center. In opposition to the dissemination, data collection aims at gathering data, relevant to safety, traffic information, infotainment, over a given area of interest.

This chapter is based on our articles published in the *Proceedings of the 12th ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, & Ubiquitous Networks (PE-WASUN '15)* [92] and in *Elsevier Ad Hoc Networks* [93], where we propose an FCD dissemination and collection protocol for VANETs in complex urban scenarios, entirely based on the DSRC technology. We also propose here an enhanced version of the original solution described in [92] and [93] by adding a backup mechanism which allows the collection of high amounts of FCD information in particularly high density scenarios. Since FCD collection in real-time from vehicles in high density urban scenarios is a qualifying new attribute of our proposal, we focus on this aspect in this chapter. Nevertheless, the protocol we propose merges the dissemination and collection functionalities in a modular way, so that the amount of information to be disseminated and/or collected and the repetition rate of the procedure can be tailored to any vehicular application environment. The proposed protocol design induces a self-organized VANET backbone structure, with no prior

knowledge of the road map or of the intersection positions. The VANET backbone is composed on relay nodes that are elected by means of DSRC-based V2V communication following the rules of a slightly modified version of the CBF algorithm described in Section 2.3. This backbone network can be flexibly used to both disseminate and collect FCD from roaming vehicles in the target area. The flexible composition of the dissemination and collection functions in a single protocol is a distinctive feature of our proposed solution.

4.1 DISCOVER Algorithm

The considered scenario comprises a RSU and a population of V vehicles moving in a given area around the RSU. We assume CAMs [36] are periodically exchanged among single hop neighboring vehicles, with a generation interval T_{gen}. The data extracted from the CAMs is stored in an LDM [35] maintained by each vehicle OBU. The LDM is updated every time a new CAM is received or when local information lifetime expires. FCD collection consists of periodic delivery of vehicle data to a traffic monitoring server via the RSU. Messages are generated by each vehicle with its respective FCD and sent to the RSU, via multi-hop communications through the DSRC-based VANET.

Let us freeze the picture of the system at the time the RSU starts a data collection process. Let L be the length of the FCD message. The net amount of data the RSU should receive is VL at most. At the time the collection is done only $N \leq V$ out of V vehicles are connected, i.e., they belong to a connected VANET graph that includes the RSU as root. Hence, the minimum amount of data that must be received by the RSU does not exceed NL bytes. The actual amount of bytes transmitted on the air through the VANET to deliver the FCDs of the N vehicles to the RSU is bigger than NL, because of a number of reasons: (i) static overhead of the VANET protocol stack (including PHY, MAC, LLC, network and transport layers), denoted as H (i.e. a physical block of data carrying L bytes of data from the facility layer of the VANET has length $H + L$); (ii) multiple transmissions of a same FCD message due to the multi-hop networking; (iii) re-transmission of messages on each link, if ARQ mechanisms are provided; (iv) signaling messages required by the data collection protocol, besides data messages devoted to FCD transport.

Let us consider the graph G formed by the RSU and the N vehicles that are connected among themselves and to the RSU at time t. The main idea of our protocol, named DISCOVER, is to select a sub-set of vehicles to act as relay nodes, thus creating a temporary backbone network that will be used for data dissemination and collection. The backbone of relay nodes should ideally form a minimum covering node set of the graph G.

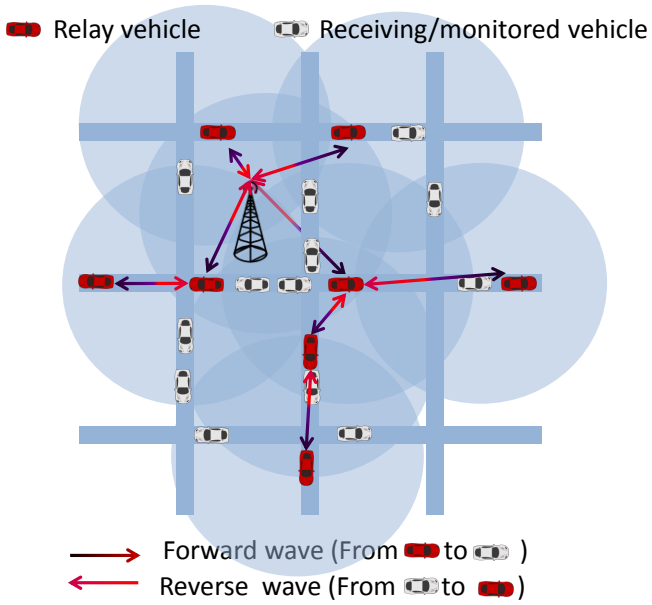


Figure 4.1 – Example of forward and reverse waves [93].

The operation of DISCOVER is organized in two phases (see Figure 4.1): a so called *forward wave*, characterized by the outbound propagation of a triggering message, named *Request*, issued by the RSU. The forwarding process is used to elect the relay nodes, while the message dissemination expands over the target ROI centered at the RSU (hence the name of forward wave). The *Request* message broadcasted by the RSU is received by vehicles traveling in the RSU's coverage area and then forwarded across the network in a multi-hop fashion. The message dissemination phase (forward wave) goes on up to a given number of hops, defined according to the desired ROI. When the forwarding has reached the target number of hops, a *reverse wave* begins. From peripheral nodes of the ROI *Reply* messages crawl back inbound towards the RSU, hopping through the designated relay nodes, this time following a backward path to the RSU. The *Reply* messages carry the FCDs in their payloads.

Both the forward (dissemination) and reverse (collection) waves are carried out by following local, autonomous rules. Each vehicle node exploits local positioning information, as provided by its local sensors, the information stored in its LDM, and the information received from the messages exchanged in the DISCOVER protocol operation. Three types of messages are defined:

- *Request*: message originated periodically by the RSU and sent during the forward wave; these messages create the backbone network by triggering the election of the relay nodes.

- *Reply*: message sent by the relay nodes back to the RSU; these messages contain the FCD collected over the ROI spanned during the forward wave.
- *Backup*: message identical to *Reply*; it is needed by the backup mechanism in the collection phase.

A *Request* message is represented by the tuple $\langle TM, ID, POS, D_{\max}, H_{\max}, H \rangle$, where:

TM Type of message: discriminates between *Request* and *Reply* messages;

ID Identification: is a unique message identifier (e.g., a counter);

POS Position: gives the geographical coordinates of the transmitting vehicle node;

$D_{\mathbf{max}}$ Maximum distance: a parameter set by the RSU to identify the maximum distance from a transmitting vehicle at which a receiving vehicle should participate in the relay node election;

$H_{\mathbf{max}}$ Hop limit: a positive integer set by the RSU according to the extension of the desired ROI;

H Hop count: non negative integer field, initialized to 0 by the RSU and incremented by each relay node; it is used to count the hops traveled by the *Request* message.

A *Reply* message is represented by the tuple $\langle TM, ID, H, S_{\text{fcd}} \rangle$, where H is the vehicle hop count from the RSU set up by the forward wave, and S_{fcd} is a data structure containing the aggregated FCDs; it is obtained by merging vehicle's own FCD with FCDs extracted from other received *Reply* messages.

4.1.1 Dissemination Phase

The dissemination phase pseudo-code is described in Algorithm 4.1 and is based on the CBF algorithm proposed by ETSI and described in Section 2.3. This algorithm runs in the application layer of the RSU, as well as in every OBU equipped vehicle. The RSU triggers the data dissemination and collection process with a frequency of T_{col}, according to the required collection frequency defined by a central collection unit (e.g., traffic monitoring server), by sending in broadcast a *Request* message. Notice that the RSU includes in every *Request* the parameters $D_{\max}$, $H_{\max}$, the message identification, RSU's geographical position, and the current hop count H initialized to 0.

Every vehicle roaming inside the target area, upon receiving the *Request*, computes its own ranking from the RSU by incrementing the hop count from the current *Request* message by 1. Let's focus on a generic receiving vehicle V_{rx}. The first condition V_{rx} has to check is if its distance from the sending vehicle V_{tx} is within $D_{\max}$.

```
1: uniqueID: a unique message identification
2: myPosition: vehicle's current geographical position
3: myHopCount: vehicle's current hop count from the RSU defined during the
   forward wave
4: relayNode: a boolean indicating whether the vehicle was selected as a relay node
   or not
5: replied: a boolean showing if the vehicle sent its own Reply or not in the current
   collection phase
6: T_curr: the current time instance

7: upon event TriggerRequest do
8:     if Self==RSU then
9:         Request.setID(uniqueID)
10:        Request.setPosition(myPosition)
11:        Request.setDmax(D_max)
12:        Request.setHmax(H_max)
13:        Request.setHopCount(0)
14:        broadcastMessage(Request)
15:        scheduleEvent(TriggerRequest, T_col)
16:    end if
17: upon event Request received do
18:    myHopCount = Request.getHopCount()+1
19:    myDistance = distance(myPosition, Request.getPosition())
20:    if myDistance < Request.getDmax() then
21:        if firstCopy(Request.getID()) then
22:            relayNode = FALSE
23:            replied = FALSE
24:            if myHopCount < Request.getHmax() then
25:                scheduleEvent(ForwardRequest, T_curr + T_req)
26:            end if
27:        else
28:            RequestCopy = getExistingCopy(Request.getID())
29:            if Request.getHopCount() > RequestCopy.getHopCount() then
30:                cancelEvent(ForwardRequest)
31:            end if
32:        end if
33:    end if
34: upon event ForwardRequest do
35:    relayNode = TRUE
36:    Request.setHopCount(myHopCount)
37:    Request.setPosition(myPosition)
38:    broadcastMessage(Request)
39:    scheduleEvent(ForwardReply, T_curr + T_rep)
```

Algorithm 4.1 – DISCOVER operation: dissemination phase

If this is true, then V_{rx} checks if this is the first copy of the just received *Request* message and if its ranking from the RSU does not exceed H_{max}. If this is also verified, then V_{rx} triggers the relay node selection process, by locally computing a *Request* forwarding timeout, T_{req}, according to the Equation (4.1).

$$T_{req} = \begin{cases} T_{req}^{max}\left[\alpha\left(1-\frac{D}{D_{max}}\right)+\beta\,\mathscr{U}(0,1)\right] & \text{if } D \leq D_{max} \\ \infty & \text{if } D > D_{max} \end{cases} \tag{4.1}$$

where T_{req}^{max} is a parameter defining the maximum timeout value, D is the current distance from the sending vehicle, and $\mathscr{U}(0,1)$ is a uniformly distributed random value needed to avoid simultaneous retransmissions. Also, α and β are weighting parameters, defining the influence of the distance and randomness factors correspondingly. These parameters can take real values between 0 and 1, while their sum must be equal to 1 in order to preserve the T_{req}^{max} constraint. According to the Equation (4.1) a vehicle having a distance closer to D_{max} with respect to the sending vehicle is more likely to become a relay node. Notice that Equation (4.1) is a slightly modified version of Equation (2.3), where we set $T_{min} = 0$, add some randomness to the algorithm to avoid simultaneous retransmissions, as well as some flexibility by making the involved parameters tunable. Moreover, since the goal is to minimize the number of relay nodes in order to have less broadcasts in the same area, vehicles outside of D_{max} simply ignore the *Request* message.

If the *Request* message is not the first copy that V_{rx} received, it means that there is another vehicle in the V_{rx}'s communication range that forwarded the same *Request* before its timeout expired. In this case, if the hop count of the received *Request* message is greater than the existing local copy, V_{rx} gets inhibited and cancels the timeout T_{req}.

If V_{rx}'s timeout expires without being inhibited, then it becomes a relay node, meaning that it will also participate in the collection phase. At this point, V_{rx} has two tasks: (i) it updates the hop count, the position, and re-broadcasts the *Request* message; (ii) it schedules a reply timeout for the collection phase, T_{rep}, according to the Equation (4.2).

$$T_{rep} = T_{rep}^{max}\left(1-\frac{H}{H_{max}}\right)+\frac{T_{rep}^{max}}{H_{max}}\mathscr{U}(0,1) \tag{4.2}$$

where T_{rep}^{max} is a time bound for the hop count factor. Notice that the maximum value T_{rep} can assume is $T_{rep}^{max}+\frac{T_{rep}^{max}}{H_{max}}$..

4.1.2 Collection Phase

The collection phase is somehow initialized during the dissemination phase, when every selected relay node sets up its own local reply timeout T_{rep}. According to the Equation (4.2) relay nodes with a smaller value of H (i.e., closer to the RSU) have higher timeout values with respect to relay nodes having greater H values (i.e., further from the RSU). This timeout setting ensures the fact that inner relay nodes hold back enough time to receive the *Reply* messages from outer relay nodes and are thus able to merge the received FCDs before replying. The pseudo-code of the collection phase operation is given in Algorithm 4.2.

When a vehicle's reply timeout expires, it attaches the FCD set S_{fcd} to the *Reply* message and sends it in broadcast on the DSRC channel. S_{fcd} set is obtained by merging the local FCDs present in the vehicle's LDM with other potential FCDs received from other relay nodes that sent their *Reply* messages earlier. In fact, each relaying vehicle has a local FCD set where the received information from other relay nodes is merged and stored. Notice that only the relay nodes that are selected during the dissemination phase are allowed to send back their *Reply* messages, hence participate in the collection phase. Differently from T_{req}, T_{rep} is never canceled before the expiration, meaning that each relay node will eventually send its *Reply* message. An illustrative representation of a complete collection phase is shown in Figure 4.2.

Despite the algorithm's effort to minimize the number of vehicles participating in the FCD collection process, the collection itself is still challenging. The main issue comes from the fact that in the collection phase we have a problem similar to the one of many source nodes sending data to one sink. The problem is more challenging as the information approaches the RSU, since more data is being merged and sent, meaning that the size of the *Reply* messages to be sent is higher at each step. Moreover, because of the increasing amount of merged and collected information, the size of the *Reply* messages can easily exceed the IEEE 802.11p Maximum Transmission Unit (MTU), which means that additional packet fragmentation has to be performed. Although there is a uniformly distributed random factor in the reply timeout computation, this does not guarantee the lack of collisions.

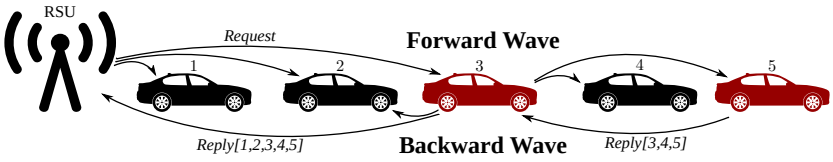


Figure 4.2 – Example of a collection instance performed by DISCOVER.

MAX_BACKUP_RTX: a parameter defining the maximum number of allowed *Backup* retransmissions
backupRTX: a variable showing the current left retransmissions for a given *Backup* message
receivedFCDSet: a local data structure containing the FCDs received from other neighboring vehicles and extracted from the corresponding *Reply* messages
localFCDSet: a local data structure containing the FCD records extracted from the vehicle's LDM

upon event *ForwardReply* **do**
    replied = TRUE
    backupRTX = MAX_BACKUP_RTX
    S_{fcd} = merge(*receivedFCDSet*, *localFCDSet*)
    Reply.setID(*uniqueID*)
    Reply.setHopCount(*myHopCount*)
    Reply.setFCDs(S_{fcd})
    Backup = *Reply*
    broadcastMessage(*Reply*)
    scheduleEvent(*ForwardBackup*, $T_{\text{curr}} + T_{\text{backup}}$)
upon event *Reply* received **do**
    if *relayNode* == TRUE **then**
        if *replied* == FALSE **then**
            merge(*receivedFCDSet*, *Reply*.getFCDs())
        else
            if $S_{\text{fcd}} \subset$ *Reply*.getFCDs() **then**
                cancelEvent(*ForwardBackup*)
            end if
        end if
    end if
upon event *ForwardBackup* **do**
    if *backupRTX* > 0 **then**
        broadcastMessage(*Backup*)
        backupRTX = *backupRTX* -1
        scheduleEvent(*ForwardBackup*, $T_{\text{curr}} + T_{\text{backup}}$)
    end if
upon event *Backup* received **do**
    if *replied* == FALSE **then**
        merge(*receivedFCDSet*, *Backup*.getFCDs())
    else
        if *myHopCount* < *Backup*.getHopCount() **then**
            broadcastMessage(*Backup*)
            backupRTX = MAX_BACKUP_RTX
            scheduleEvent(*ForwardBackup*, $T_{\text{curr}} + T_{\text{backup}}$)
        else
            cancelEvent(*ForwardBackup*)
        end if
    end if

Algorithm 4.2 – DISCOVER operation: collection phase

To cope with this issue, we enhance the original DISCOVER protocol [92], [93] by adding a backup mechanism in the collection phase. The backup mechanism is based on overhearing and allows a relaying vehicle to retransmit its *Reply* up to a predefined number of times, in case its message has not reached the destination. Basically, whenever the reply timeout expires, meaning that the vehicle has to broadcast its *Reply* message, it creates a local copy of this message, named *Backup*, and schedules a backup timer, T_{backup}, according to the Equation (4.3):

$$T_{\text{backup}} = \frac{2T_{\text{rep}}^{\max}}{H_{\max}} + T_{\text{b}}\mathscr{U}(0,1) \tag{4.3}$$

where T_{b} is a time bound for the randomness factor. If T_{backup} expires, the vehicle checks if it does not exceeded the maximum allowed *Backup* retransmissions, hence broadcasts the *Backup* and re-schedules T_{backup}. However, if while a vehicle waiting for its T_{backup} to expire overhears a *Reply* or *Backup* message containing its S_{fcd}, and that has been sent by another relaying vehicle having a smaller hop count (i.e., a vehicle closer to the RSU), then this vehicle cancels its T_{backup} to avoid unnecessary retransmissions.

4.2 Simulation Setup

To validate our proposed solution, we consider the Luxembourg SUMO Traffic (LuST) scenario [94], [95], a realistic vehicular traffic scenario that was specifically built and tailored to support the evaluation of vehicular networking protocols and applications. This scenario represents the Luxembourg City, a typical mid-size European city with common characteristics in terms of road topology and mobility patterns. In particular, LuST covers 932 km of roads and an area of 156 km^2, containing 38 different bus routes with 563 bus stops. Road topology and segments, buildings geometry, points of interest, traffic lights, and other environment information have all been extracted from OpenStreetMap[10].

The vehicular traffic model in LuST is based on a realistic mobility study that describes the traffic characteristics of Luxembourg City over recent years. LuST models the traffic pattern over a 24 h time period. This information was used in SUMO [54] micro-mobility simulator to create and simulate the vehicular traffic mobility. SUMO is coupled with OMNeT++ [53] simulation tool and Veins [55], which are all described in Section 2.4, and that are used to simulate the communication process, including the operations of the PHY, MAC, and network layers, as well as our protocol implementation. To model the impact of buildings and other obstacles to signal propagation, we have used jointly two attenuation models: the Two-Ray

[10] `www.openstreetmap.org`

Interference model [96], [97] with $\epsilon_r = 1.02$, and the Obstacle Shadowing model [88], which reproduces in Veins the shadowing effect of a real urban environment by describing the attenuation as a function of the depth of the buildings crossed by radio links.

Although the vehicular mobility is simulated over the entire LuST scenario, the network simulation is concentrated in a 2 km × 2 km central area, identified as our ROI to be monitored (see Figure 4.3), while the statistics are collected over a smaller area, 1.4 km × 1.4 km (the area delimited by the green vehicles in Figure 4.3), to avoid border effects. The buildings and the points of interest are represented by red polygons. An RSU placed at the center of this area, represented by the blue circle in Figure 4.3, periodically triggers the FCD collection process by sending in broadcast *Request* messages every $T_{col} = 5$ s. A simulation run lasts 100 s and every run is repeated 15 times for statistical confidence (95 % confidence intervals are also computed). In our study, the 100 s simulation time is chosen in three different points in time from the entire 24 h range, so as to cover three different vehicular densities. In particular, we identify a high vehicular density scenario at approximately 8:00 o'clock in the morning, a medium density scenario at 13:00 o'clock, and a low density scenario at 11:00 o'clock. The main simulation parameters are displayed in Table 4.1.

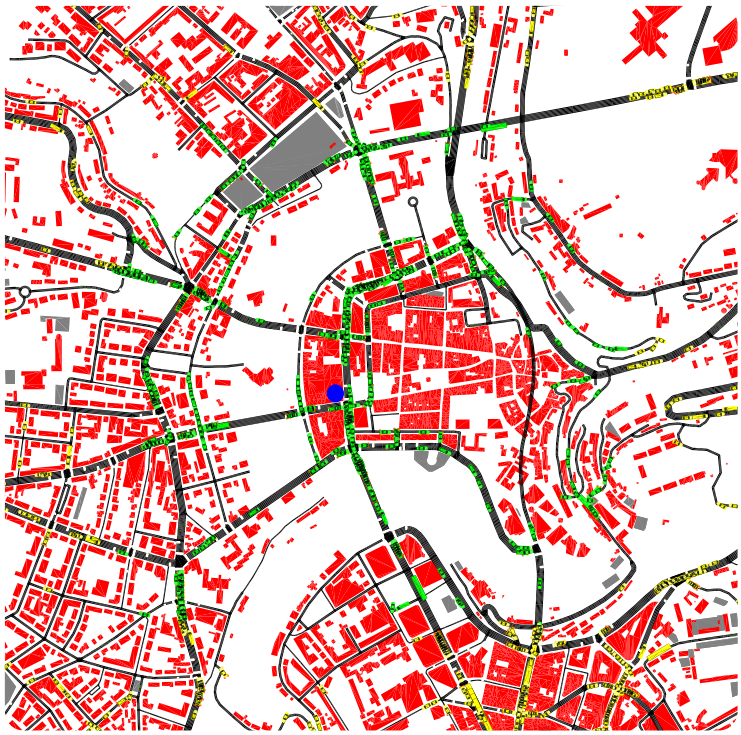


Figure 4.3 – The simulated scenario.

Parameter	Value
Luxembourg ROI	2 km × 2 km
Avg. number of vehicles in ROI	816 (high), 506 (medium), 267 (low)
Simulation duration	100 s
IVC technology	IEEE 802.11p
IVC maximal transmit power	100 mW
DSRC beaconing frequency T_{gen}	1 Hz
DSRC bitrate	6 Mbit/s
Carrier frequency	5.89 GHz
FCD size	40 B
Beacon size	40 B
T_{col}	5 s
D_{max}	100 m, 300 m, 500 m, 700 m and 900 m
H_{max}	4, 10 and 16
α, β	0.8 and 0.2
$T_{\text{req}}^{\text{max}}$	0.1 s
$T_{\text{rep}}^{\text{max}}$	0.4 s, 0.7 s and 1.0 s
T_{b}	0.1 s
MAX_BACKUP_RTX	0, 1, 2, 3 and 4

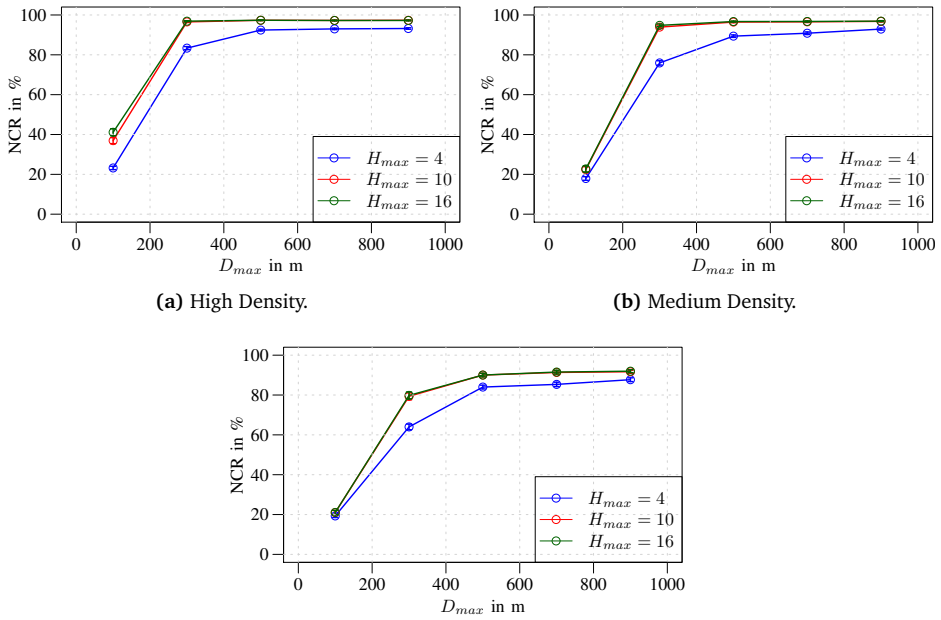
Table 4.1 – Main simulation parameters

4.3 Dissemination Phase Evaluation

The purpose of the data dissemination phase is for the *Request* message sent by the RSU to reach as many vehicles as possible within the target ROI. To measure this, we define the Node Coverage Ratio (NCR) as the ratio of the number of vehicles that received the *Request* message over the total number of vehicles roaming inside ROI at the time instance when the *Request* message was issued by the RSU. From the protocol description we can see that D_{max} and H_{max} are the main parameters that can affect the NCR metric, which is why in our analysis we vary these two parameters.

In Figure 4.4 we show the performance of DISCOVER in terms of NCR when varying D_{max} and H_{max}. The first thing that we can notice is that for low values of D_{max} the protocol is less effective, especially for the low vehicular density scenario. The main reason is that the algorithm is designed to only allow vehicles within D_{max} from the sending vehicle to participate in the relay node selection process. This means that for low values of D_{max} it is more likely for the dissemination process to be interrupted due to the lack of vehicles roaming within that distance. In fact, for $D_{\text{max}} \geq 300$ m in case of high and medium densities, and $D_{\text{max}} \geq 500$ m in case of low density, DISCOVER is able to reach on average almost 100 % of vehicle roaming inside ROI, if H_{max} is high enough. This result suggests also the fact that D_{max} should be tailored to the vehicular density. Notice that with $H_{\text{max}} = 4$ the mean NCR does not overcome 80 % simply because the limited number of hops imposed by the RSU does not allow the *Request* message to propagate further.

Another metric of interest that measures the dissemination performance is the Relay Nodes Ratio (RNR), defined as the ratio of the number of relaying vehicles



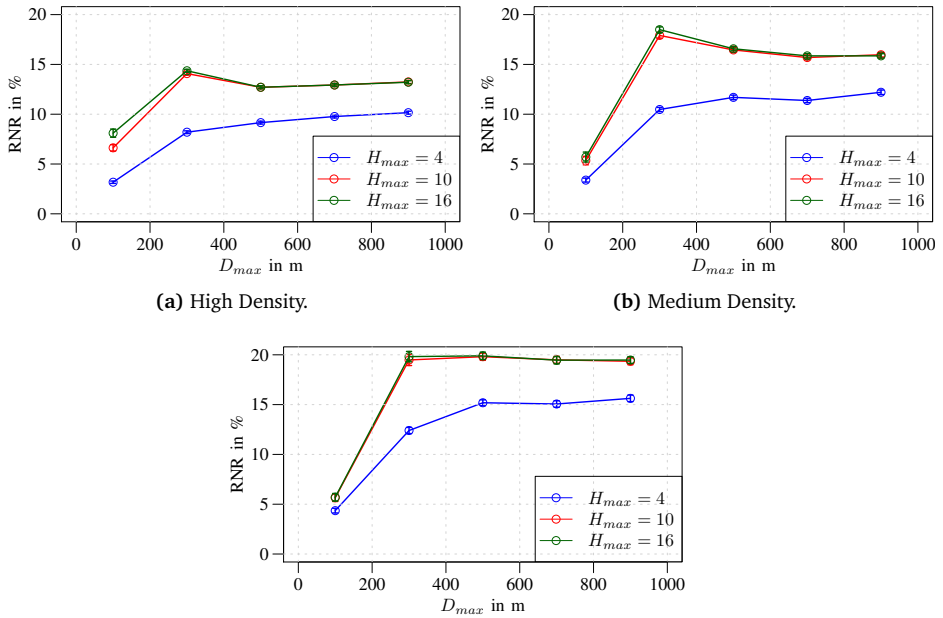
(a) High Density.

(b) Medium Density.

(c) Low Density.

Figure 4.4 – Mean NCR for different $D_{\max}$ and $H_{\max}$ values, and for different vehicular density scenarios.

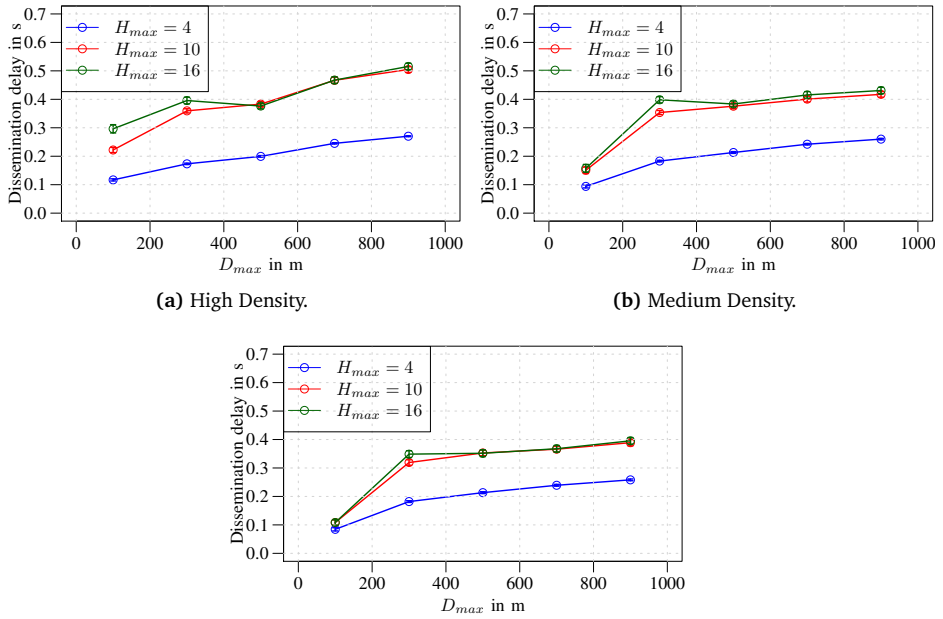
over the total number of vehicles roaming inside ROI. In particular, this metric measures how efficient is the CBF algorithm in minimizing the number of relaying vehicles. Figure 4.5 illustrates the RNR metric for different values of $D_{\max}$ and $H_{\max}$, as well as for the three considered vehicular densities. We can notice that DISCOVER is quite efficient in terms of RNR metric. As expected, the RNR values are smaller for high vehicular density scenarios, and larger for low density scenario. This is simply because CBF tends to elect relaying vehicles that are geographically separated by a distance $D_{\max}$, independently on how many vehicles are roaming in the area. This means that even if the total number of vehicles in the scenario is growing, the number of selected relay nodes remains more or less the same. As a consequence, DISCOVER is actually more efficient in terms of RNR for high density scenarios, with RNR $\approx 14\,\%$, while for low density scenario RNR $\approx 20\,\%$. Notice that the very low RNR values when $D_{\max} = 100\,\text{m}$ and $H_{\max} = 4$ are due to the fact that the dissemination process is interrupted prematurely, as can be seen from Figure 4.4. The spike at $D_{\max} = 300\,\text{m}$ can be explained by the fact that with this value DISCOVER is able to cover most of the vehicles in ROI with a higher number of hops with respect to $D_{\max} > 300$. Moreover, increasing $D_{\max}$ over a certain value does not decrease significantly the RNR, suggesting the fact that in order to cover a certain area we need a minimum number of relaying vehicles.



(a) High Density. **(b)** Medium Density.

(c) Low Density.

Figure 4.5 – Mean RNR for different $D_{\max}$ and $H_{\max}$ values, and for different vehicular density scenarios.



(a) High Density. **(b)** Medium Density.

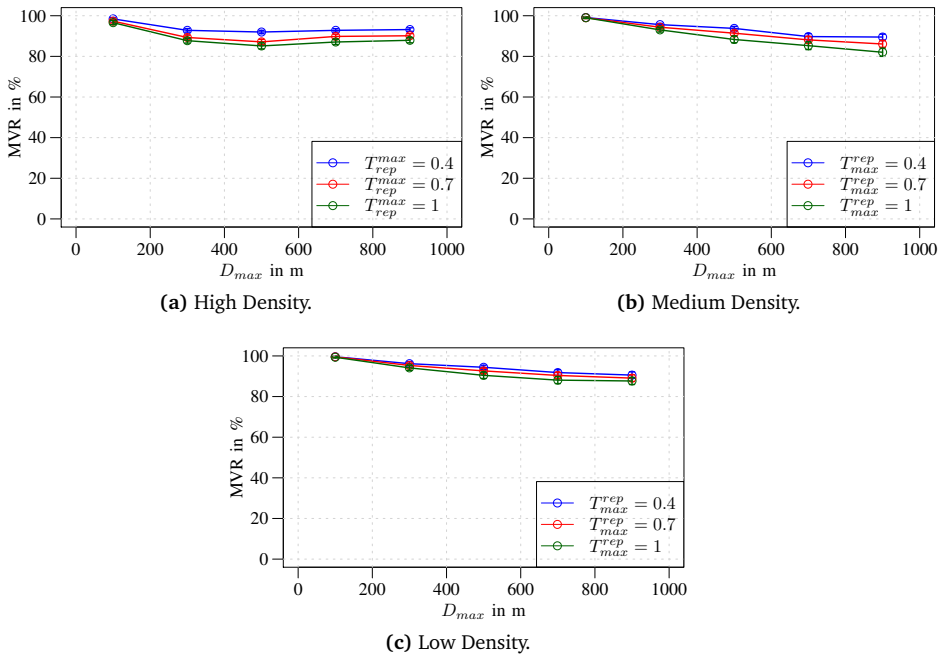
(c) Low Density.

Figure 4.6 – Mean dissemination delay for different $D_{\max}$ and $H_{\max}$ values, and for different vehicular density scenarios.

Finally, Figure 4.6 shows the dissemination delay, measured as the time interval between the moment when the RSU issues the *Request* message and the time instance when the last vehicle inside our monitored area receives it. The first thing to notice is that in the worst case scenario it takes roughly 0.5 s for the dissemination phase to reach all possible vehicles inside our area of interest. Another observation is that there is a sort of optimal delay value for $D_{\text{max}} = 300\,\text{m}$ and $D_{\text{max}} = 500\,\text{m}$, when the *Request* message reaches most of the vehicles in approximately 0.4 s. The low dissemination delays for $H_{\text{max}} = 4$ are related to the poor reachability and are consistent with the results from Figures 4.4 and 4.5.

4.4 Collection Phase Evaluation

The main purpose of DISCOVER is to collect FCD information from all vehicles roaming inside a ROI. To measure this, we define a new metric, named Monitored Vehicles Ratio (MVR), which is computed as the ratio of the number of vehicles whose FCDs arrived to the RSU at the end of a collection phase (i.e., vehicles monitored by the RSU) over the total number of vehicles that received the *Request* message. Figure 4.7 plots the MVR metric with respect to D_{max} and $T_{\text{rep}}^{\text{max}}$, and for the three considered vehicular density scenarios. Overall, DISCOVER performs quite



(a) High Density.

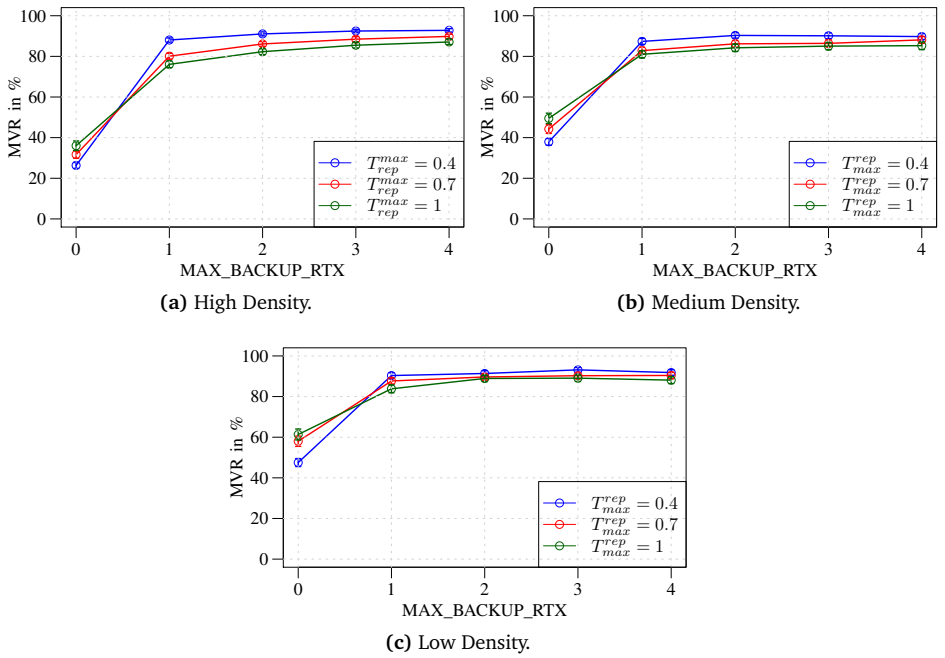
(b) Medium Density.

(c) Low Density.

Figure 4.7 – Mean MVR for different D_{max} and $T_{\text{rep}}^{\text{max}}$ values, and for different vehicular density scenarios.

well and is able to collect roughly 90 % of the vehicles that received the *Request* message. Notice that for the backup mechanism in this set of simulations we fixed MAX_BACKUP_RTX = 4. This result is not affected so much by the $D_{\max}$ or $T_{\text{rep}}^{\max}$ parameters. Moreover, DISCOVER is robust independently from the considered vehicular density scenario. We can notice that for $D_{\max} = 100\,\text{m}$ we are able to monitor almost 100 % of the vehicles, but this is simply due to the fact that the dissemination phase reaches very few vehicles. Also, it seems that $T_{\text{rep}}^{\max} = 0.4\,\text{s}$ is a little bit better than $T_{\text{rep}}^{\max} = 1\,\text{s}$, meaning that the stability of the backbone network formed by the selected relay nodes is slightly affected by the mobility of the vehicles.

But we are also interested in evaluating the impact of the backup mechanism on all the metrics that we will consider for the collection phase. In Figure 4.8 we fix $D_{\max} = 700\,\text{m}$ and show how the MVR metric is affected by the MAX_BACKUP_RTX parameter. We can see that with no backup mechanism the MVR values drop to approximately 30 %, 40 % and 50 % for the high, medium and low density scenarios respectively. Such poor performance comes from the high number of *Reply* packet collisions. Remember that the size of the *Reply* messages increases when they are getting closer to the RSU, since more information is being merged together. Hence, without a backup mechanism, the collisions have a significant impact on the protocol performance. However, we can see that already with MAX_BACKUP_RTX = 1 the



(a) High Density.

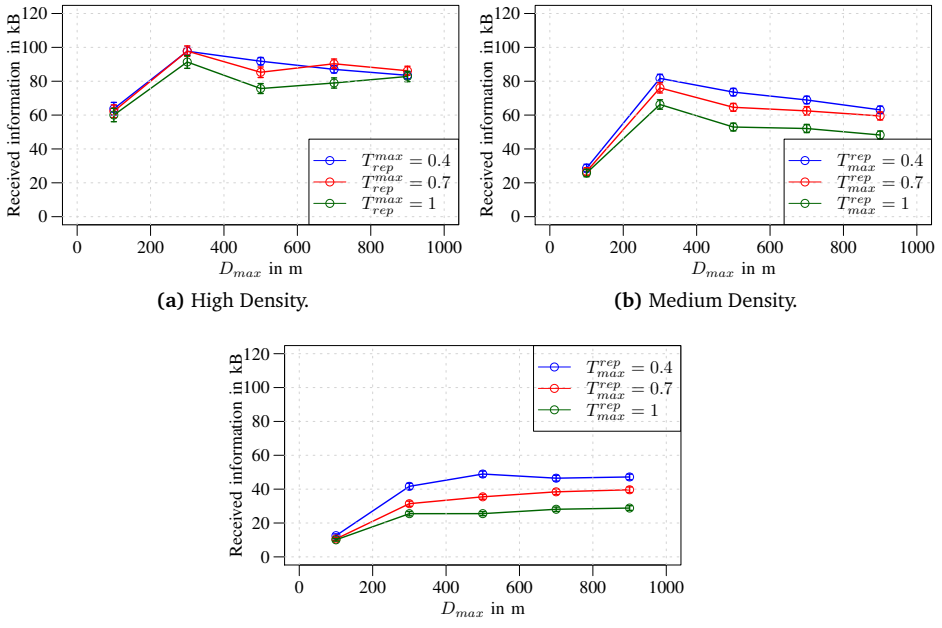
(b) Medium Density.

(c) Low Density.

Figure 4.8 – Mean MVR for different MAX_BACKUP_RTX and $T_{\text{rep}}^{\max}$ values, and for different vehicular density scenarios.

MVR values go up to 90 %, while higher values of MAX_BACKUP_RTX have a limited impact.

Figure 4.9 gives a quantitative assessment of the average amount of information collected by the RSU during one FCD collection cycle. Of course, the amount of the collected information depends on the number of vehicles roaming inside the ROI. In the high vehicular density scenario the amount of collected information ranges between 80 kB and 100 kB, when considering MAX_BACKUP_RTX = 4 and reasonable values of D_{max} (i.e., the ones with high NCR and MVR values), while for the low density scenario the RSU receives between 20 kB and 50 kB of data. Since with $D_{\text{max}} = 100$ m there are only few vehicle reached by the *Request*, the amount of FCDs is also small. A slight performance difference can also be noticed for different values of $T_{\text{rep}}^{\text{max}}$, with more information being received for low values of $T_{\text{rep}}^{\text{max}}$. This suggests the fact that for higher values of $T_{\text{rep}}^{\text{max}}$ the aggregation mechanism works better, hence the backup mechanism is invoked fewer times. This last result is confirmed also by the Figure 4.10, where we show the amount of information received by the RSU when varying the MAX_BACKUP_RTX parameter with a fixed $D_{\text{max}} = 700$ m. We can notice the amount of received information increases for higher values of MAX_BACKUP_RTX and lower values of $T_{\text{rep}}^{\text{max}}$.

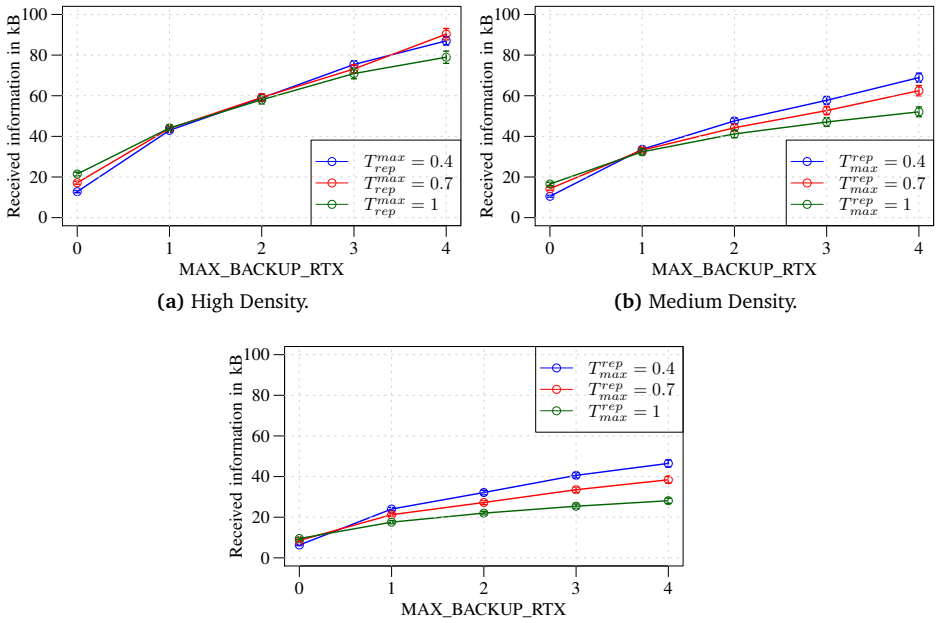


(a) High Density.

(b) Medium Density.

(c) Low Density.

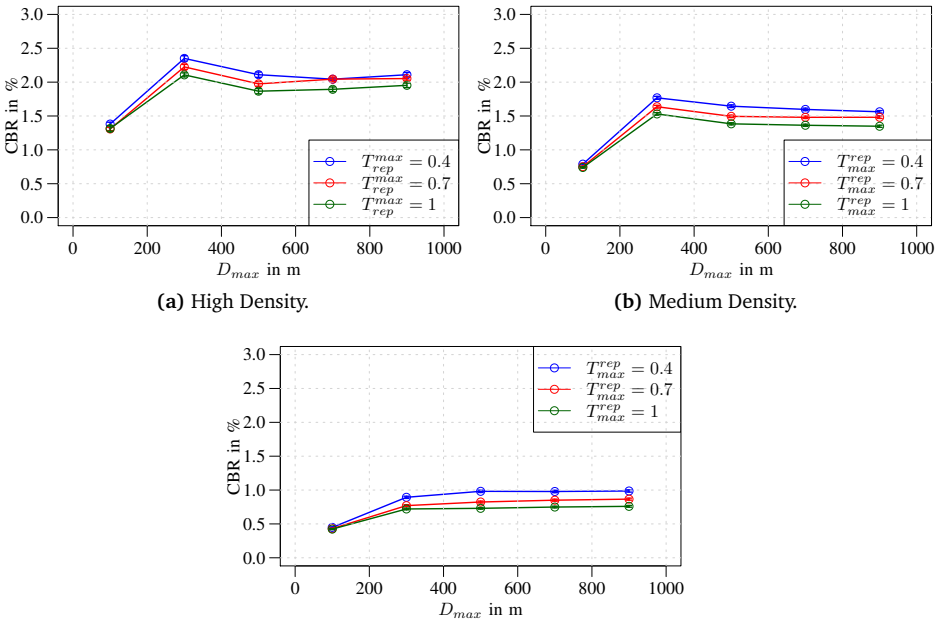
Figure 4.9 – Amount of information received by the RSU during one collection cycle for different D_{max} and $T_{\text{rep}}^{\text{max}}$ values, and for different vehicular density scenarios.



(a) High Density. **(b)** Medium Density.

(c) Low Density.

Figure 4.10 – Amount of information received by the RSU during one collection cycle for different MAX_BACKUP_RTX and $T_{\text{rep}}^{\text{max}}$ values, and for different vehicular density scenarios.



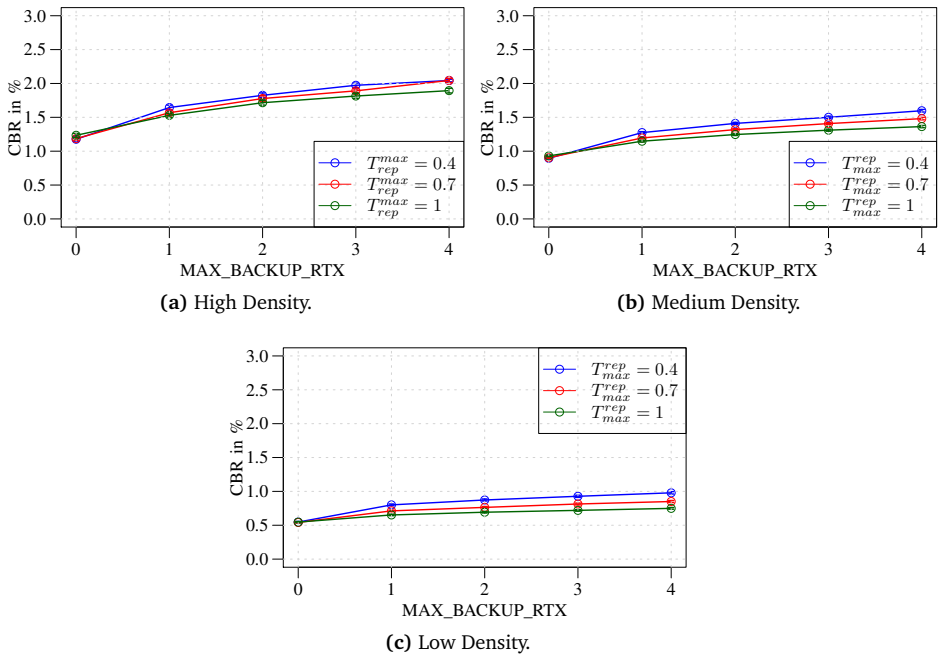
(a) High Density. **(b)** Medium Density.

(c) Low Density.

Figure 4.11 – Mean CBR for different D_{max} and $T_{\text{rep}}^{\text{max}}$ values, and for different vehicular density scenarios.

The amount of the collected information has a direct impact on the DSRC channels, as can be seen from Figure 4.11, where we highlight the mean Channel Busy Ratio (CBR) for different values of $D_{\max}$ and $T_{\text{rep}}^{\max}$, as well as for different vehicular density scenarios. The CBR metric is computed as the ration of the amount of time a vehicle senses the DSRC channel busy to the total simulation time related to that vehicle. This metric is also an indication of how congested is the DSRC channel. Notice that in Figure 4.11 we plot the average values over the total number of vehicles and over the total simulation time. Although the mean absolute CBR values are low, it is worth mentioning that these values are not uniformly distributed in time. In particular, the time series of the CBR values (which we do not illustrate here) show concentrated spikes right after the moment when the RSU starts the collection process, which are related to the duration of the collection process itself.

The obvious thing is that the CBR metric is directly affected by the number of DSRC-equipped vehicles roaming inside the ROI, as can be seen from Figure 4.11. Numerically, CBR $\approx 2\,\%$ for the high density scenario, CBR $\approx 1.5\,\%$ for the medium one, while it drops down to less than $1\,\%$ for the low vehicular density scenario. We also show the CBR metric against MAX_BACKUP_RTX in Figure 4.12. Again, we have a confirmation of the fact that the backup mechanism affects the DSRC channels. In



(a) High Density.

(b) Medium Density.

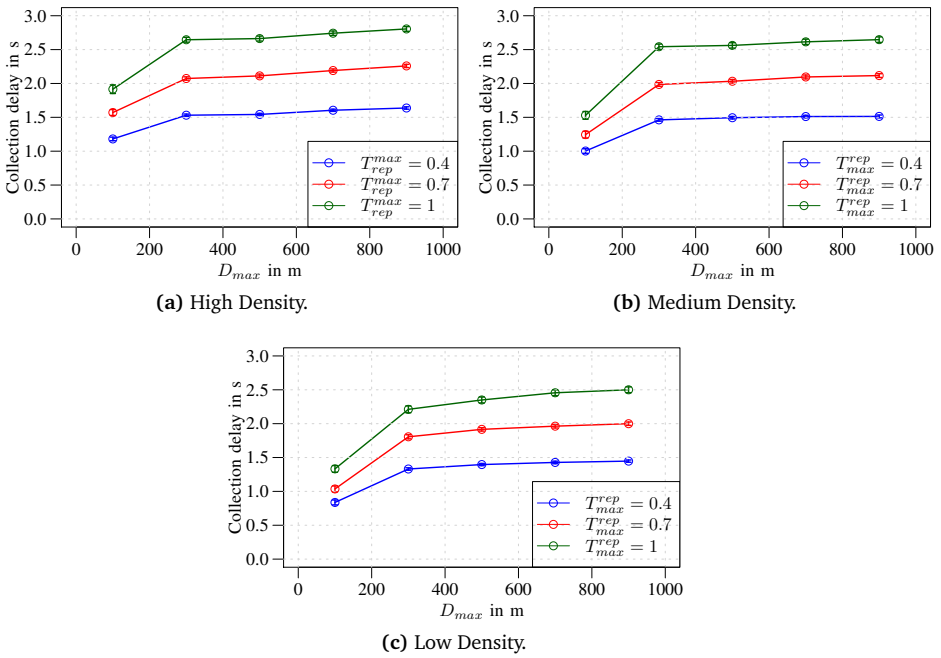
(c) Low Density.

Figure 4.12 – Mean CBR for different MAX_BACKUP_RTX and $T_{\text{rep}}^{\max}$ values, and for different vehicular density scenarios.

particular, increasing the values of the MAX_BACKUP_RTX on one hand improves the MVR metric, but on the other hand puts more load on the communication channels.

Finally, in Figure 4.13, we illustrate the average collection delay with respect to $D_{\max}$ and $T_{\text{rep}}^{\max}$ and for the three considered vehicular densities. The collection delay, that is the time needed by the algorithm to complete a single collection cycle, is measured as the time difference between the moment when the RSU starts the collection process by issuing in broadcast a *Request* and the time instance the RSU receives the last FCD message in the current collection interval. As we can see from Figure 4.13, the collection delay is not affected so much by the $D_{\max}$ parameter. The only exception is for $D_{\max} = 100\,\text{m}$, where the collection time seems to be lower. The reason is that the dissemination process is interrupted very close to the RSU, meaning that the overall amount of collected information is small. This leads to fewer collisions of *Reply* messages and, as a consequence, the backup mechanism, which is time consuming, is invoked fewer times. The same explanation is valid when analyzing the impact of the vehicular density on the collection delay, even though the influence is quite limited. With fewer vehicles, less is information is collected, hence, less *Reply* messages loss.

These results are confirmed also in Figure 4.14, where we highlight the collection delay when varying the MAX_BACKUP_RTX parameter. We can see that the average

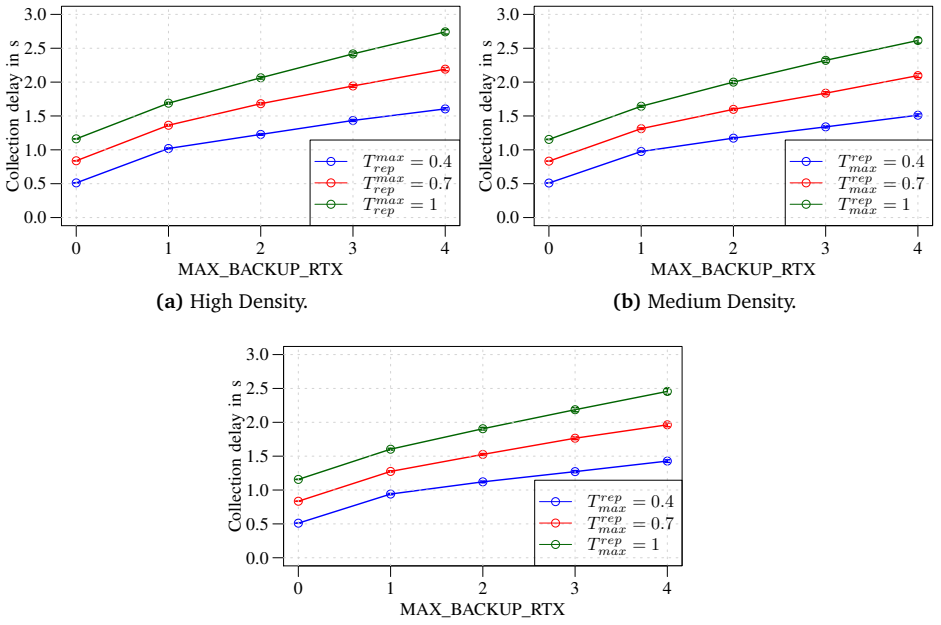


(a) High Density.

(b) Medium Density.

(c) Low Density.

Figure 4.13 – Mean collection delay for different $D_{\max}$ and $T_{\text{rep}}^{\max}$ values, and for different vehicular density scenarios.

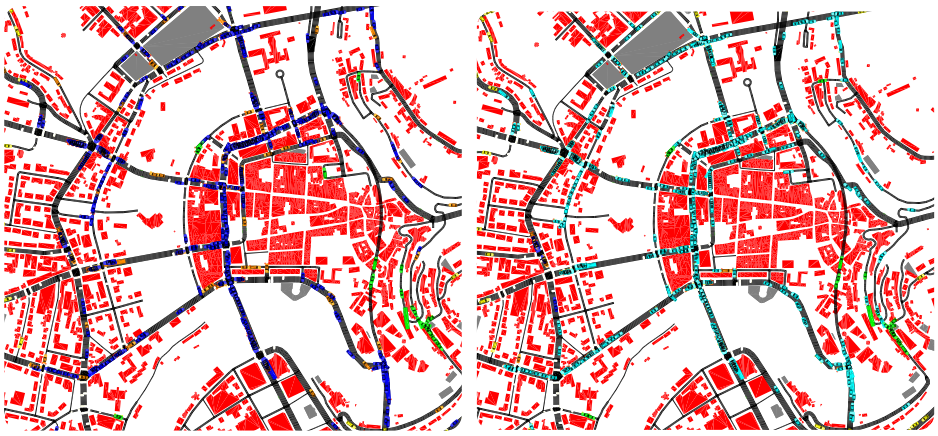


(a) High Density.

(b) Medium Density.

(c) Low Density.

Figure 4.14 – Mean collection delay for different MAX_BACKUP_RTX and $T_{\text{rep}}^{\text{max}}$ values, and for different vehicular density scenarios.



(a) Dissemination phase.

(b) Collection phase.

Figure 4.15 – An illustrative example of a collection cycle using DISCOVER.

collection delay increases with MAX_BACKUP_RTX. The parameter that seems to have a greater impact on the collection delay is $T_{\text{rep}}^{\max}$. In fact, we can notice that high values of $T_{\text{rep}}^{\max}$ generate high collection delays, while small values of $T_{\text{rep}}^{\max}$ help collecting data faster. Overall, DISCOVER is able to complete a collection cycle in less than 3 s in the considered ROI.

To have an idea of what a traffic monitoring application is able to see after one collection cycle, in Figure 4.15 we show the result of one simulated collection instance divided in the two phases: dissemination (Figure 4.15a) and collection (Figure 4.15b). In Figure 4.15a vehicles colored in blue are the ones that received the *Request* issued by the RSU, the green vehicles are the ones that are roaming inside ROI but did not received the message for different reasons (e.g., collision, disconnected from the graph, etc), while the brown vehicles are the relay nodes selected by the CBF algorithm and that will be the ones participating in the collection phase. In Figure 4.15b cyan colored vehicles are the ones whose FCDs arrived at the RSU at the end of the collection phase. As we can notice, DISCOVER was able to collect the FCDs from most of the vehicles that received the *Request* message.

4.5 Conclusion

In this chapter we studied the feasibility of the DSRC technology to support periodic collection of FCD information needed by non-safety applications. We proved that the DSRC technology is perfectly suitable for real-time collection of vehicular information, at least for reasonable FCD message sizes, i.e., if including only basic information in the FCD message, enough for a traffic efficiency application.

To do this, we proposed an integrated data collection protocol, named DISCOVER, that exploits the vehicular backbone network made of relay vehicles elected by means of a standardized data dissemination protocol, named CBF, to periodically collect FCD messages from all vehicles roaming inside a large urban area. We evaluated the proposed solution by means of simulations and illustrated a possible application outcome. The performance evaluation shows that DISCOVER is able to collect on average 90 % of vehicle roaming inside a 1.4 km × 1.4 km typical European urban area in less than 3 s. We simulated realistic packet sizes and measured the amount of information moved by the VANET towards the RSU. Also, we evaluated the impact of the data collection protocol on the DSRC channels.

As a future work, the protocol could be tested in larger city areas and with greater amount of information to be collected. In fact, an interesting study would be to measure the performance of the protocol, but also of the DSRC technology itself, when collecting massive amounts of information.

Chapter 5

Heterogeneous LTE-DSRC FCD Collection in Vehicular Networks

Most of IVC research proposes DSRC as the main technology to be used for vehicular *safety* applications. One of the main motivations is the very low transmission delay (in the order of ms) required by these applications and satisfied by DSRC. Another reason is that DSRC operates on a dedicated spectrum (75 MHz in the United States and 50 MHz in Europe at 5.9 GHz frequency), which is specifically assigned for ITS. On the downside DSRC suffers from scalability issues. Also, in order to support centralized services and applications, additional gateways and hardware is needed, like Road Side Units. The deployment of such infrastructure is expensive [98]. Moreover, the technology itself is not yet widely available.

LTE has been identified as a good candidate technology for supporting *non-safety* applications [62], like urban sensing and traffic efficiency. These applications are generally delay tolerant and aim at improving the vehicle traffic flow, traffic coordination and assistance, as well as providing up-to-date locally relevant information bounded in space and/or time. The applications usually require intermittent collection of data from every vehicle roaming inside a target area. The collected information can contain kinematic data for traffic monitoring (e.g., vehicles' position, speed, direction of travel, time), technical and service data for vehicle monitoring, or environmental data for urban sensing. This information, known in the literature as FCD, needs to be periodically reported to a remote central server for processing. Of course, the granularity of the collected data and the reporting frequency depends on the target application type. In this context LTE offers high throughput, promises high penetration rate, and has the advantage of being already widely deployed. However, LTE has several drawbacks. First of all, it operates in a licensed spectrum, meaning that its performance and availability is highly dependent on the mobile and network operators. Also, in high density urban scenarios the periodic data transmissions from

many vehicles can use a significant part of the LTE channels, possibly degrading the normal operation of traditional applications. In order to support the increasing amount of data traffic, LTE needs further upgrades, like decreasing the cell sizes, or adding more spectrum. All these upgrades are not for free, requiring additional investments from the network operators.

Optimizing the utilization of the LTE resources when periodically collecting information in vehicular networks is a challenging task. A typical approach aiming to solve this issue is the adoption of clustering mechanisms in multi-technology heterogeneous vehicular networks [23]–[25], [66]–[69]. The main motivation for this is to use other technologies to offload the traffic from the cellular network. Generally, these clustering algorithms consist in selecting a subset of vehicles, named Cluster Head vehicles, to act as local *aggregators* and *forwarders* towards the cellular network. The forwarder election itself can be done either in a centralized [68] or distributed [23]–[25], [66] fashion, while the aggregation inside each cluster is performed through IVC.

In the following sections, we present and discuss different FCD collection schemes for heterogeneous vehicular networks, that exploit the DSRC technology to significantly decrease the LTE channel utilization.

5.1 VANET-Driven LTE Off-loading

In this section, considering two realistic urban scenarios, we present a hybrid networking mechanism under which a VANET-based V2V networking protocol is employed for the purpose of supporting LTE-based FCD collection operation. The work presented here is based on our articles published in the *Proceedings of 12th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2016)* [99] and *Elsevier Vehicular Communications* [100].

The aim is to substantially reduce the number of concurrently active LTE channels and the information message load carried across the LTE cellular network for the same accuracy of the vehicular traffic description obtainable when FCD are collected via LTE from each individual vehicle. We define a distributed procedure that exploits the "horizontal" capability of vehicles to communicate among themselves via the VANET, to elect representative nodes. The election process exploits the logic of the so called dissemination protocols. The representative nodes are responsible for communicating aggregated FCD via the LTE infrastructure. The performance gains achieved through the use of the proposed approach rapidly increase as the vehicular density increases. Under such high density conditions, the traffic load of the LTE cellular network can become critically high, while VANET networking connectivity improves. Under low vehicular density levels, our procedure falls back onto the use

of a plain LTE-based FCD collection scheme. The employed operation and protocols rely on the use of geographical information known individually by each vehicle (e.g., via GPS), not requiring the use of external databases (such as those that make use of urban city maps and junction proximity sensors).

5.1.1 Typical FCD Collection Solutions

Most of the current state-of-the-art proposals based on a hybrid LTE plus DSRC networking technology follow a common paradigm, in which two main algorithmic phases can be recognized: i) SETUP; ii) COLLECTION (see Figure 5.1) [68], [71], [72]. The SETUP phase exploits the LTE cellular network, while the COLLECTION phase is based on both local communication in the VANET among neighboring vehicles and transmission of the collected data over LTE channels. If communication in the VANET is not possible, meaning that only the LTE technology is available, then SETUP and COLLECTION collapse in a unique phase. In this case FCD messages are collected individually by each vehicle (e.g., LTE box in Figure 5.1).

The SETUP phase aims at gathering status information involving vehicles that roam in the target area, and making this information available to the remote server. Proposed techniques published to-date envision an operation during the SETUP phase under which each vehicle communicates the relevant data individually to the server via LTE connections. This information is then used to set up, in an optimal fashion, the process governing the mode of operation to be used during the ensuing FCD COLLECTION phase.

During the COLLECTION phase, the vehicular population is split into clusters. Cluster head vehicles are elected based on the information collected during the SETUP phase. The choice of cluster head can be the outcome of an optimization problem that takes into account: i) information on vehicles' positions, velocities and directions; ii) VANET connectivity information (neighbors of each vehicle, according to the received CAMs); iii) information on the CQI of the LTE channels measured by

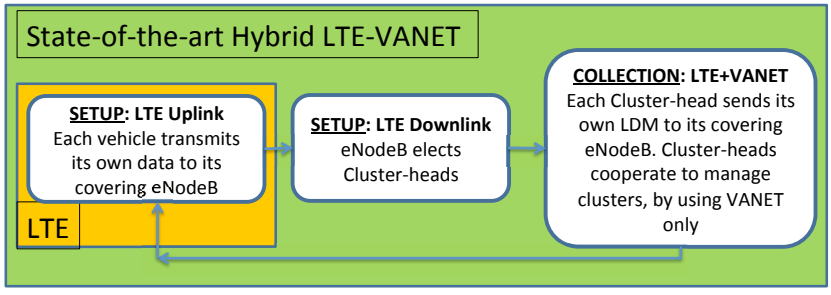


Figure 5.1 – LTE and Hybrid LTE-DSRC FCD collection schemes in the existing proposals [100].

the vehicle on board units. A centralized optimization approach, run in the remote server, can be used to identify the best candidate vehicles for the role of cluster head nodes. The cluster heads are then designated at the end of SETUP phase, before the start of the ensuing COLLECTION phase, by sending control messages on the LTE downlink channels that cover the target area. A cluster head is responsible for collecting FCDs from its 1-hop neighboring vehicles via VANET wireless links. The cluster head then forwards the collected data to the remote central server using its LTE connection. In this manner, only cluster heads (rather then each vehicle) use LTE channels.

Let N denote the number of vehicles[11] in the target monitored area A. According to the above described operational paradigm, N independent LTE channels are established and activated during the SETUP phase, while only $M \ll N$ LTE channels are used during the COLLECTION phase, where M is the number of cluster heads. The average vehicular density is equal to $\rho = N/A$. If R denotes the radio coverage range realized by a single nodal VANET transmitter, one can estimate the M level to satisfy: $M \sim A/(\pi R^2) = N/(\rho \pi R^2)$. For a vehicle density of $\rho = 100\,\text{veh/km}^2$ and $R = 300\,\text{m}$, we have $M/N \approx 0.0354$.

In general $K \geq 1$ FCD collections are performed during a COLLECTION phase. The COLLECTION phase continues in an uninterrupted manner until it is determined that the current cluster layout deviates beyond a margin level from a currently calculated optimal configuration. A new SETUP phase is then triggered. The topological layout of cluster heads and their election operations are thus adapted to new system conditions, refreshing the information required to optimally synthesize the layout and operations governing the ensuing COLLECTION phase. The duration of the COLLECTION phase is therefore tied to the scope and features of the monitored area and to the dynamics of the vehicular traffic roaming the area. Summing up, we envisage a time period T_{cycle} to refresh the SETUP of the collection network. Within the time frame of duration T_{cycle}, one SETUP phase is carried out, with duration T_{S}, as well as K COLLECTION phases, each of duration T_{C}. Then, it is $T_{\text{cycle}} = T_{\text{S}} + K T_{\text{C}}$.

To summarize, there are different state-of-the-art approaches for both SETUP and COLLECTION phases. These can be depicted by a sort of flow chart (see Figure 5.2). We can list the following different operations:

- $\text{LTE}_{\text{SETUP\_UP}}$: each vehicle transmits in uplink its own data to its covering eNodeB;
- $\text{LTE}_{\text{SETUP\_DW}}$: the eNodeB elects cluster head vehicles by sending election messages in the downlink;
- $\text{VANET}_{\text{SETUP}}$: election of cluster heads using VANET only; each cluster head transmits its own partial LDM to its covering eNodeB;

[11] N is assumed to stay constant over one SETUP+COLLECTION cycle

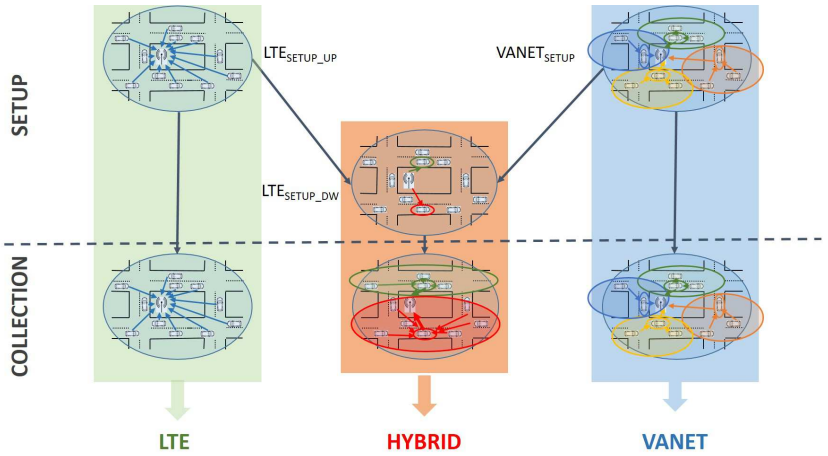


Figure 5.2 – Summary of the three different schemes for SETUP and COLLECTION [100].

- COLLECTION: each node (in the *LTE* case) or cluster head (in the *HYBRID* or *VANET* cases) sends its own FCD and those of the vehicle nodes it is responsible for (if any) to its covering eNodeB.

5.1.2 Proposed Solution

The lesson learned from the studies cited above is that the number of used LTE channels can be reduced by using only specifically designated nodes to send collected data through the LTE access network to the remote server. Each such designated node would aggregate and forward data that represents the status of vehicles in its immediate neighborhood. This status data stored in the LDM is available at each vehicle, as each one continuously collects such data through the maintenance of a background CAM exchange process.

The key idea of our proposal is that such designated nodes can be identified by executing an election process across the VANET. The distributed protocol for the designated node election can be derived from the logic of dissemination protocols. A dissemination logic provides for the multi-hop transport of messages across the vehicular network through the election of certain vehicle nodes to act in forwarding a received message to other vehicles. By definition, the dissemination logic implies the designation of special nodes, that make up a connected set of nodes, covering the area spanned by the VANET. The designated nodes are employed as local data aggregation points that are used for collecting and sending FCD information obtained from neighboring vehicles to the server via LTE. The effectiveness of the

dissemination procedure increases as the vehicular density increases. This is just the scenario where offloading for LTE access network is most critical.

A distinguishing feature of our proposal is that it is fully seamless for the LTE network. Differently from most previous approaches, no modification or new logic is required in the LTE cellular network. The designation of representative nodes and the local collection of FCD is carried out by a "horizontal" process that makes use of the VANET only. Elected representative nodes upload aggregated FCD to the remote server via LTE channels, without any further intervention from the LTE network, e.g., to orchestrate or manage vehicles clusters. This approach achieves a useful decoupling between vehicle specific functions (FCD aggregation and maintenance of up-to-date LDMs) and generic communication functions (uploading of FCD to the server via the LTE network). The proposed approach aims at leveraging the strong points of either technology: the VANET for its ease of direct communication among neighboring vehicles, and the LTE access network for its potentially high capacity and pervasive availability. As a consequence the proposed approach adapts automatically to any given penetration rate of DSRC equipment.

Summing up, the key features of our proposed approach are as follows:

- We take advantage of utilizing the dedicated spectrum bands assigned for VANET services to reduce the traffic loads imposed on the LTE wireless access network.

- Our proposed mechanism can be realized in a manner that is fully compliant with current technology and standards, e.g., by using the CBF algorithm of the GeoNetworking protocol [27] as the dissemination logic. Alternatively, it could be programmed as an application level function sitting on top of standard PHY, MAC and network protocols.

- No special new function is required of the LTE cellular network, i.e., the proposed approach for FCD collection is fully seamless to the LTE network control plane.

- The LTE network, as well as other future cellular networks, can offer message transport at much higher communications rates. Cell sizes are becoming smaller and high inter-cell interference effects become dominant. The latter limits the attainable system throughput efficiency level. It is consequently more effective to employ a lower number of nodes for the forwarding of larger amounts of data aggregates, instead of a large number of sources of relatively small amounts of data.

Election of Representative Nodes: Connected VANET Case

We define a REQUEST message that is originated by a *trigger node*, starting the dissemination-like process. The trigger node can be a RSU located in a central position of the target area, or it can be a designated OBU. The REQUEST message is disseminated according to the rules used by the GeoNetworking CBF protocol outlined in Section 2.3. The nodes that are elected as forwarders of the REQUEST message during this dissemination phase, are identified as *representative nodes*. They are in charge of reporting the status data of their neighboring vehicles to the remote server via LTE connections.

Let A denote a generic node that sends the REQUEST message (hence A is the trigger node or any of the elected representative nodes). The message sent by A contains: i) an identifier; ii) the geographical position of A; iii) a count-down hop-count field, initialized by the trigger node to the maximum number of hops $H_{\max}$ that the REQUEST message is allowed to travel and decremented by each re-broadcasting node; iv) a list L_{A} of vehicle node IDs that A commits to report to the server on.

By re-broadcasting the REQUEST message, a node A recognizes to have been designated to act as a representative node for the vehicles listed in L_{A}. Hence, the node A constructs a reduced neighbor vehicle database R_{LDM} by omitting from its full LDM those nodes whose IDs are listed in the REQUEST message that A has received. The list of IDs contained in the R_{LDM} is inserted in the copy of the REQUEST message that A sends out. A will report FCD relative to only those vehicles that appear in its R_{LDM}. Since a single representative node is elected for each VANET radio neighborhood (the maximum IEEE 802.11p vehicular radio transmission range $D_{\max}$ being in the order of several hundred meters), the number of LTE channels that are effectively used for the transmission of messages is drastically reduced.

Election of Representative Nodes: Multiple Connected Components Case

Let T_{cycle} define the duration of the SETUP plus COLLECTION phases. The trigger node starts a new time period by issuing a new REQUEST message every T_{cycle} seconds. This time period can be broken up into a SETUP phase of duration T_{S}, when representative nodes are elected, and the ensuing COLLECTION phase, when a new set of FCD is sent by current representative nodes every T_{C} seconds, until the COLLECTION phase is terminated and a new set of representative nodes is to be elected. If the collection phase is repeated K times, then $T_{\mathrm{cycle}} = T_{\mathrm{S}} + KT_{\mathrm{C}}$.

Given the maximum number of hops $H_{\max}$ that the REQUEST message is allowed to traverse (which is related to the ratio between the radius of the target areas and $D_{\max}$; typically $H_{\max}$ is limited up to few tens), the REQUEST message dissemination delay over the connected component of the VANET that the trigger node belongs to assume a value that lies between $H_{\max}T_{\min}$ and $H_{\max}T_{\max}$. Practical values of $T_{\max}$

are in the order of 100 ms. Then, the maximum message dissemination delay is typically below few seconds.

While the trigger node role can be played by suitably scattered RSUs, a simple distributed, OBU-based protocol can be defined to trigger the representative node election. The only requisite is that a vehicle that subscribes the service knows the collection time schedule (i.e., T_S and T_C) and realizes it is inside the area where the collection service is active. This requisite is easily met by using predefined data stored into the information collection application and the GPS on board the vehicle. At the beginning of every cycle, each vehicle inside the service area sets a timer by choosing a value uniformly at random in the interval $[0, T_{\text{trig}}]$ and schedules a REQUEST message. If the timer expires and the node has not received any other REQUEST message, it sends out its own REQUEST message and elects itself as a representative node. If the node receives such a message before the trigger timer expires, it cancels its scheduled REQUEST message and schedules the forwarding of the message it has just received, by selecting a timer value according to the rule of Section 2.3. Then, the protocol proceeds as detailed in Section 2.3 (forwarding and inhibition rules). This fully distributed protocol finds a suboptimal coverage of the vehicle nodes, i.e., the elected representative nodes are in general more than required by a minimum covering set. On the other hand, the mechanism described above is adaptive to the penetration rate of DSRC equipment. It falls back automatically to the case where each single vehicle reports directly its own data via LTE (LTE only approach) as little or no DSRC equipped vehicles are around. On the opposite, as the vehicle density grows, which is the critical case for the cellular network loading, the connectivity of the VANET graph and the distributed procedure outlined above ensure that only a fraction of the vehicle nodes gets elected, as shown in the performance evaluation.

5.1.3 Simulation Models and Scenarios

We consider an urban area scenario covered by one or more LTE macro-cells. FCD updates originated by vehicles moving in the underlying coverage area are collected continuously over time and fed to a number of ITS related applications. Conceptually, we think of the collected FCD as processed by a remote backhaul server. The placement of the server is immaterial to the ensuing discussion. The relevant point is that FCD collected from the monitored area, often encompassing more than a single LTE macro-cell, are processed together, thus exploiting jointly the information collected over the entire monitored area.

Vehicles are assumed to be equipped with OBUs supporting LTE and DSRC communication technologies, plus a GPS device. Vehicles generate, send and receive CAMs periodically, as described by the ETSI standard [36]. The CAM exchange is conducted through the DSRC VANET operation over its dedicated bandwidth. By

receiving CAMs, each vehicle creates its own LDM. In this manner, it is aware of the states of other vehicles in its neighborhood area, including their time-stamped positions, velocities, moving directions, vehicle attributes. We do not have an RSU but we assume that the scenario is based only on OBUs, delegating the role of the trigger node to one or more of these ones as described in Section 5.1.2.

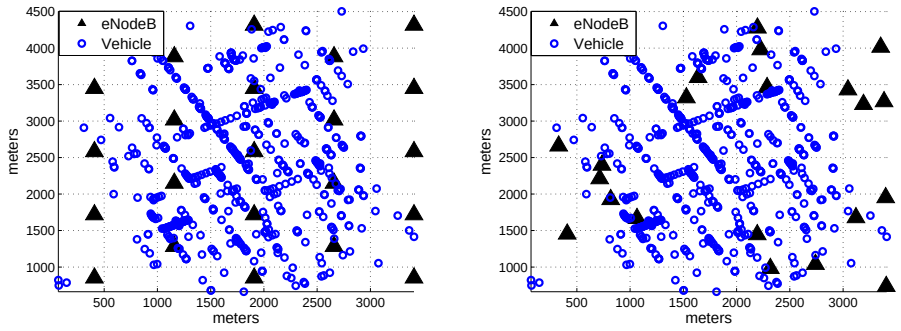
We evaluate the performance of our proposed mechanisms by using the multi-layer simulation framework described in Section 2.4 and composed of SUMO [54], OMNeT++ [53] and Veins [55]. To define two urban scenarios, we consider actual urban maps of the city centers of Rome and New York (Figure 5.3), obtained from OpenStreetMap[12]. The first is a part of the district of Manhattan in the city of New York (see Figure 5.3a). This map is mainly characterized by a regular grid of avenues and streets that create a considerable number of junctions. The second considered scenario covers the neighborhood of Termini Central Station in the city of Rome (see Figure 5.3b). In contrast with the first scenario, this one is characterized by a high level of road layout irregularity and a higher measure of stochastic street orientations. Both considered maps extend over an area of about 12 km^2. Figure 5.4 shows a snapshot of the vehicle positions in the New York map (circle markers) with the superimposed layout of LTE cellular eNodeBs (triangle markers). We consider two different cases for the placement of LTE eNodeBs: a) eNodeBs located according to a regular hexagonal grid of radius R_{eNodeB} (Figure 5.4a) b) eNodeBs scattered in accordance to a random process, Figure 5.4b.

Mobility of vehicles is generated by the micro-mobility simulator SUMO, according to the Krauss vehicular mobility model and the so called "random trips" model. The movement of the vehicles is governed by the car-following model with a target

[12] `https://www.openstreetmap.org/`

(a) Manhattan District. **(b)** Rome.

Figure 5.3 – Considered urban scenario maps [100].



(a) hexgonal grid placement of eNodeBs **(b)** Randomly scattered eNodeBs

Figure 5.4 – Monitored urban area covered by LTE macro-cells (Manhattan, NY) [100].

speed of 50 km/h. According to vehicular traffic features (vehicle density in each road lane, velocity limits, traffic lights) the actual realized velocity can be lower than the target one.

The OMNeT++ tool is used to simulate the behavior of the communication process, including the operations of the Physical, MAC and network layers. The MAC and PHY parameters are set equal to those specified by the IEEE 802.11p standard (see Table 5.1). We invoke the packet broadcasting operations mode, under which no ACK frames are produced at the MAC layer, as conducted under the IEEE 802.11p MAC specification. We have embedded the implementation of the representative node election logic described in Section 5.1.2 in the network layer.

As for the VANET, we have jointly used two attenuation models: the Two-Ray Ground [101] and the Simple Obstacle Shadowing Model (SOSM) [88]. Two-Ray

Parameter	**Value**
Vehicle density λ(veh/km^2)	$50-110$
Vehicle target speed (km/h)	50
R_{eNodeB} (m)	$500 \div 3500$
$D_{\max}$ (LOS) (m)	827
$T_{\min}$ (ms)	0
$T_{\max}$ (ms)	100
$H_{\max}$	20
Path loss model for IEEE 802.11p	Two-Ray Ground + SOSM
VANET MAC, PHY parameters	IEEE 802.11p
IEEE 802.11p Link Rate (Mbit/s)	6
IEEE 802.11p tx power (dBm)	27
Carrier frequency IEEE 802.11p (GHz)	5.9
LTE UE tx power (dBm)	27
Carrier frequency LTE (GHz)	0.8

Table 5.1 – Notations and simulation parameter values

Ground models the distance dependent component of the power loss: it assumes that the attenuation is $A(d) = \kappa d^{\alpha_1}$, for distances d up to a break point value d_{bp}. For $d > d_{bp}$, it is $A(d) = \kappa d_{bp}^{\alpha_1 - \alpha_2} d^{\alpha_2}$. Typical values of the path loss parameters are $d_{bp} = 120\,\text{m}$, $\alpha_1 = 2$, and $\alpha_2 = 4$. SOSM reproduces in Veins the shadowing effect of a real urban environment: it describes the attenuation as a function of the depth of the obstacles (e.g., buildings) crossed by radio links. The description of the obstacles in the considered map layouts is taken from the metadata provided by the OpenStreetMap repository.

The COST-Hata model [102] of path loss for urban areas has been used to evaluate the vehicle node CQI and the LTE cell that each vehicle node is associated with (the one with the best detected CQI). The Modulation and Coding Scheme (MCS) is set by each vehicle node transmitting over the LTE channel according to its observed CQI, unless stated otherwise. In our model we did not include the data overhead and the association time delay [18] for the connection establishment. This means that for quite large cells (the ones considered in our study) it is assumed that this overhead is negligible since for the time duration of a SETUP and a COLLECTION a vehicle associates only to one eNodeB. We avoid the use of femto and small cells on purpose, to represent the best case in the use of LTE and to measure in this way only the amount of data exchanged for the FCD without the overhead for the cell association. Notice that in case of small and femto cells, due to the vehicle mobility (13.8 m/s), a quite large overhead should be instead considered for the LTE association.

Numerical values used for simulation parameters are listed in Table 5.1. Every considered scenario, over a zonal scope of about $12\,\text{km}^2$, has been analyzed under different vehicular densities λ, as reported in Table 5.1. The baseline solution, taken as a benchmark in the performance comparison, sets a configuration under which each vehicle sends its own FCD directly to the eNodeB using the LTE access network. This solution represents the performance obtained when vehicular data are gathered by using only the LTE network [103][62][104]. Also, it represents the performance behavior of all proposed Hybrid LTE-DSRC mechanism during the SETUP phase.

5.1.4 Performance Metrics and Evaluation

We employ the following performance metrics:

F_{RV} fraction of all vehicles roaming in the target area that are reached by the REQUEST message propagated in the VANET according to the representative node election logic;

F_{RN} fraction of all vehicles roaming in the target area that are elected as representative nodes (vehicles that forward the REQUEST message) in the VANET;

$F_{\mathbf{MV}}$ fraction of all vehicles roaming in the target area whose data are reported to the remote server via LTE connections established by the representative nodes;

$D_{\mathbf{RQ}}$ REQUEST message delay: time needed to complete the propagation of the REQUEST message, measured from the instant that this REQUEST message is issued by the trigger node to the time that it has completed its dissemination over the graph component to which the trigger node belongs to;

$M_{\mathbf{CH}}$ number of LTE PUSCHs [105] that must be established in a cell to make nodes report their FCD data to the server via the LTE network.

$M_{\mathbf{RB}}$ average number of LTE Resource Blocks [105] per LTE cell, required by vehicles for communicating over the LTE system;

As for M_{CH}, it is assumed that each node reporting data to the remote server uses a single PUSCH in each COLLECTION instance. Note that a reporting node can aggregate data from other vehicles through the VANET, or it can just report its own data, in case it has no DSRC neighbors.

The number M_{RB} is calculated as follows. All reporting nodes are considered. Let L_{k} be the amount of data that the k-th representative node must report. The spectral efficiency of the k-th node is obtained from its CQI level. Let it be r_{k} bit/RB. Then, the number m_{k} of RBs required by the k-th reporting node is $m_{\mathrm{k}} = \lceil L_{\mathrm{k}}/r_{\mathrm{k}} \rceil$. Let N_{eNodeB} be the number of LTE eNodeBs in the scenario and let R_{j} denote the set of representative nodes under the coverage of LTE cell j, $j = 1, \ldots, N_{\mathrm{eNodeB}}$. Then $M_{\mathrm{RB}} = \frac{1}{N_{\mathrm{eNodeB}}} \sum_{j=1}^{N_{\mathrm{eNodeB}}} \sum_{k \in R_{\mathrm{j}}} m_{\mathrm{k}}$.

The performance analysis that we carry out accounts for the conduct of the two operations: dissemination of the REQUEST message over the VANET system and vehicle data reporting by the elected representative nodes through the LTE system.

Evaluation of Representative Nodes Election

In the simulation experiments, the trigger node is a randomly selected vehicle. The trigger vehicle is chosen with uniform probability among those roaming in the central part of the considered map. This corresponds to studying the capability of the considered FCD collection protocols in the area surrounding the trigger node.

Performance behavior is assessed by means of evaluation of the metrics F_{RV}, F_{RN}, F_{MV} and D_{RQ} in the two urban scenarios described in Section 5.1.3. Results are presented in Table 5.2. When considering the New York scenario, we can notice that F_{RV} is almost insensitive to the vehicle density level and it equals about 90 %. The observed values of F_{RN} range between 0.24 for $\lambda = 70\,\mathrm{veh/km^2}$ down to 0.19 for $\lambda = 110\,\mathrm{veh/km^2}$. The fraction of vehicles that serve as representative nodes is thus noted to reduce as the vehicular density λ grows, i.e., the efficiency of the

Scenario	λ (veh/km^2)	F_{RV}	F_{RN}	F_{MV}	D_{RQ} (s)
New York	70	0.89	0.24	0.95	0.31
	110	0.90	0.19	0.98	0.40
Rome	70	0.93	0.27	0.94	0.35
	87	0.93	0.22	0.94	0.48

Table 5.2 – Performance metrics for the dissemination of the REQUEST message in the New York and Rome scenarios.

aggregation operated by the representative nodes improves with growing levels of λ.

As for the Rome map, F_{RV} is again stable with different vehicle density levels. It settles to slightly higher values than with the New York scenario ($F_{RV} \simeq 0.94$ for Rome). Also in this case F_{RN} decreases with the vehicle density, consistently taking higher values than in the New York case, namely $F_{RN} \simeq 0.27$ for $\lambda = 70\,\text{veh/km}^2$, $F_{RN} \simeq 0.22$ for $\lambda = 87\,\text{veh/km}^2$.

For both New York and Rome scenarios the fraction of monitored vehicles F_{MV} is close to 1 and insensitive to the vehicle density level. In other words, the designated representative nodes do actually represent (cover) essentially all vehicles roaming in the target area.

The dissemination time D_{RQ} is dependent on the vehicular density level λ. In the New York scenario, for the lower λ, it took approximately 310 ms for the message to reach 89 % of the vehicles. The message dissemination delay increases to 400 ms for the higher λ level. The corresponding values for the Rome scenario are between 350 ms and 480 ms. The higher levels of delay and F_{RN} observed in the Rome map are due to the irregularity of the street layout that is noted to have lower vehicular communications connectivity, so that a larger number of hops are needed to reach out vehicles distant from the trigger node point.

Evaluation of LTE Cellular System Load: SETUP Phase

Once the representative nodes are elected, they proceed to report the FCD of the vehicles roaming in the ROI. We investigate the case where the reported FCD data contains the vehicles' geographical positions. Under our approach, each representative node sends a REPORT message with its own FCD and the positions of the vehicles whose IDs are listed in the R_{LDM} built during the REQUEST dissemination phase. The REPORT message sent by each representative node consists of:

- network plus transport headers (IPv6+UDP) of 48 B [103];
- an application level header of 48 B, that contains the representative node ID, its position and the same data as envisaged in the Vehicle High Frequency

Container of the CAMs[13]; moreover, it contains also the number $n \geq 0$ of ensuing records, relevant to neighbor vehicles' data;

- a list of records, each record having a length of 32 B, and being composed of: i) a 1 B sequence number; ii) a 17 B encoding of the 17 characters US National Highway Traffic Safety Administration (NHTSA) standard Vehicle Identification Number; iii) the position of the reported vehicle, encoded with 14 B.

Overall, a REPORT message containing data from n neighborhood vehicles has a length of $96+32n$ B. We investigate the performance behavior of the urban scenarios by varying the value of the LTE eNodeB distance R_{eNodeB} and by considering different vehicular densities. The crucial points are: i) the overhead implied by setting up and maintaining an active LTE connection, hence the number of used LTE channels per cell; ii) the load seen by an LTE eNodeB due to the overall number of vehicle nodes under its coverage that require an LTE channel.

The impact of the vehicle data transfer through the LTE access network is highlighted by the results in Figures 5.5 and 5.6. The metrics M_{CH} and M_{RB} are plotted as a function of the inter-eNodeB distance, R_{eNodeB}, for New York (Figure 5.5) and Rome (Figure 5.6) scenarios. In these figures, we compare two approaches: i) each vehicle sends its own data individually, by using its own dedicated LTE connection (curves labelled with *LTE*); ii) our proposed protocol is used, representative nodes are elected and only those nodes report data about themselves and about their respective neighbors via their LTE connections, as described in Section 5.1.2 (curves labelled with *VANET*).

We can notice that, under the *LTE* approach (curves with the square marker), the number of LTE channels M_{CH} assigned by an eNodeB to report FCD to the remote server is equal to the total number of vehicles under the coverage area of the eNodeB. This number grows quickly as the area covered by a single eNodeB expands. In comparison with the *LTE* approach, we note that the *VANET* scheme (curves with circle markers) is able to reduce the number of nodes elected to report vehicles' data via the LTE access network, leading to a substantial reduction of the number of required LTE channels per cell. The presented curves flatten for growing values of R_{eNodeB}, since eventually only a single LTE eNodeB covers most of the considered map area. When a single cell covers most of vehicles, further increments of R_{eNodeB} do not change the load of the single LTE cell in the scenario.

To expand the performance evaluation as the vehicle density and the eNodeBs layout are varied, in the case of the New York map we have set up specific models.

[13]Our setting is consistent with [103], where it is mentioned that the maximum length of a CAM containing only the mandatory fields, including the Basic Container and the Vehicle HF Container, is 50 bytes

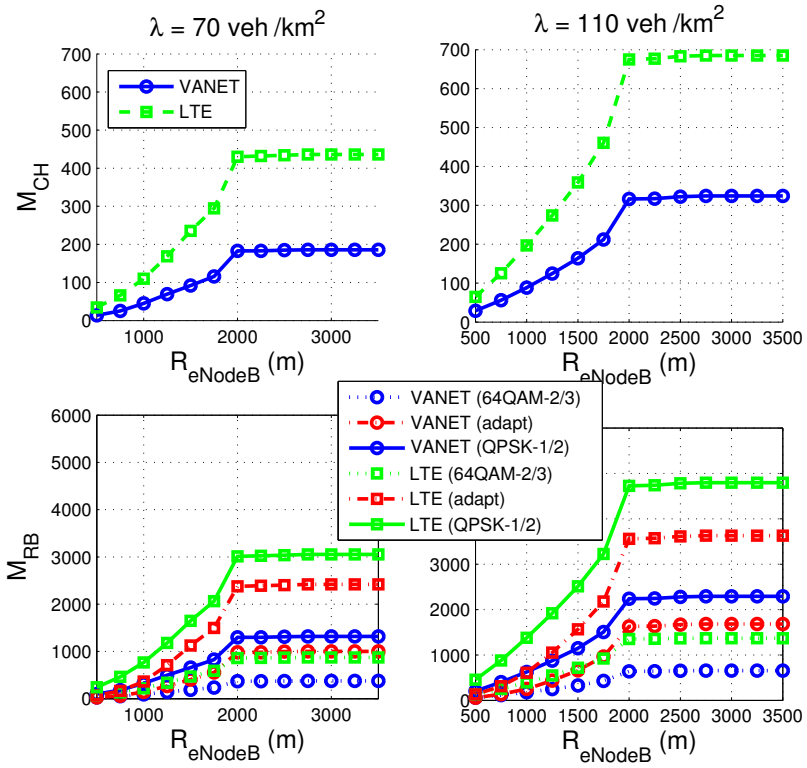


Figure 5.5 – New York scenario: average number of LTE uplink radio channels used per cell, M_{CH}, (top graphs) and average number of uplink RBs used per cell, M_{RB}, (bottom graphs) vs. the eNodeB transmission range R_{eNodeB} for two different vehicular density levels [100].

The positions of the eNodeBs have been generated according to a hard-core spatial random process, namely Matern type II [106]. It is constructed starting from a uniform spatial Poisson Point Process with mean density μ_b. Then points are assigned with random marks drawn form a uniform probability distribution over $[0,1]$. Points having a neighbor within distance d with a mark level less than their own are labelled with a '0'. After removing all points labelled with a '0', the residual points cannot be closer than the chosen distance d. The relevant mean density μ is

$$\mu = \frac{1 - e^{-\mu_b \pi d^2}}{\pi d^2} \tag{5.1}$$

If the eNodeBs were laid out according to a regular hexagonal grid of radius R_{eNodeB}, the resulting density would be $\mu = \frac{2}{3\sqrt{3}R_{\mathrm{eNodeB}}^2}$. This value can be plugged into Equation (5.1), hence the value of μ_b can be found, given d. In our simulations we set $d = 100\,\mathrm{m}$ and let R_{eNodeB} vary from 200 m up to 3000 m. Vehicle positions

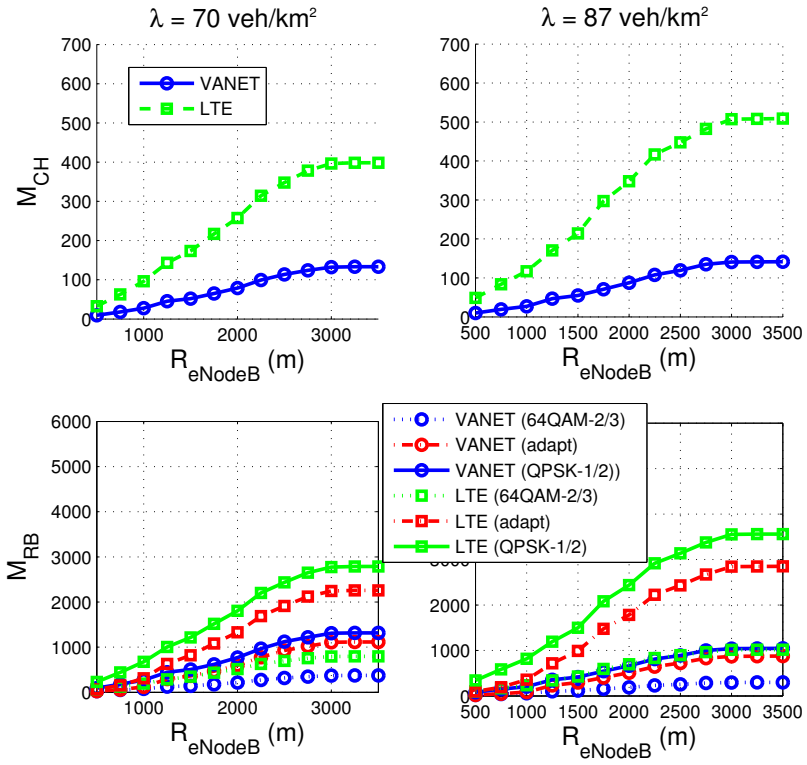


Figure 5.6 – Rome scenario: average number of LTE uplink radio channels used per cell, M_{CH}, (top graphs) and average number of uplink RBs used per cell, M_{RB}, (bottom graphs) vs. the eNodeB transmission range R_{eNodeB} for two different vehicular density levels [100].

are obtained from the SUMO simulation of the New York area, as for Figure 5.4. To let the vehicle nodes vary, we sample vehicles with probability p, i.e., we assume that only a fraction p of the vehicles moving in the considered area take part in the traffic information collection. Hence, the vehicle node density is $p\lambda$, where λ is the average density of all vehicles. Vehicle nodes are associated to the closest eNodeB, i.e., the serving eNodeB is chosen as the one having the least average path loss to the vehicle node. In these simulations we have used the fully distributed trigger node procedure outlined in Section 5.1.2 (multiple connected components case), since we vary the vehicle density and hence we consider cases where the VANET graph is sparse and disconnected.

Figure 5.7a plots M_{CH} as a function of the average cell radius for the case of randomly scattered eNodeBs for an average density of vehicle of 82.2 veh/km^2. Figure 5.7b plots M_{CH} as a function of the average vehicle node density, obtained by vehicle sampling as explained above; the average cell radius is 600 m. Blue square

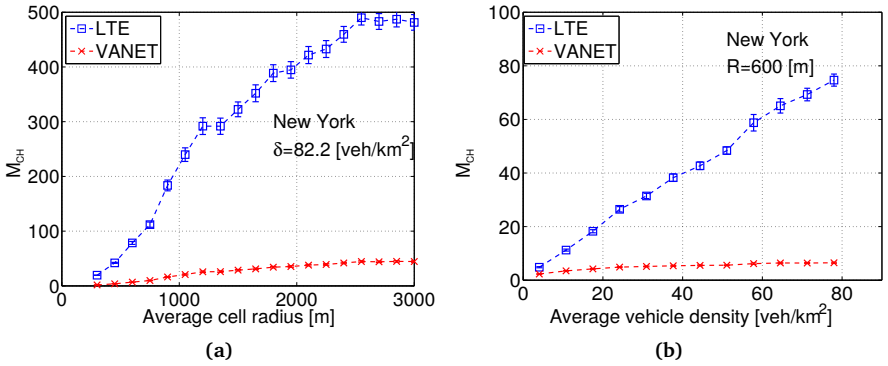


Figure 5.7 – Average number of LTE uplink radio channels used per cell, M_{CH} under randomized eNodeB positions: (a) varying average cell radius; (b) varying vehicle node density [100].

markers refer to the case where only the LTE cellular network is used for the vehicle data collection, while the red cross markers correspond to our protocol, based on VANET level election of representative nodes that are the only one using an LTE channel. 95 % level confidence intervals are shown as well.

The major offload brought about by the use of the VANET is highlighted by the performance curves shown in Figures 5.5 and 5.6 for M_{RB}, which is the average number of RBs used by representative nodes in each LTE cell, as a function of R_{eNodeB}. We identify two performance bounds: i) the best case, when each node using an LTE channel is able to use the high rate MCS, namely 64 QAM with code rate 2/3; ii) the worst case, when every node using an LTE channel must use the lowest rate MCS, namely QPSK with code rate 1/2. Besides those bounds, we also evaluate the intermediate case, where each node using an LTE channel measures its CQI and infers what is the best MCS that it can use (curves labelled with *adapt*). The worst performance exhibited under the *VANET* approach is close to the best performance obtained under the *LTE* scheme for the highest λ. The performance gap between the corresponding bounds and between the two approaches (as measured by the adaptive case) broadens as the number of eNodeBs is reduced. This is a critical issue, since low intensity data collection of FCD should be taken care of by macro-cells, rather than by hot spot micro-cells, intended to boost the capacity offered in special areas for broadband users. On the other hand, macro-cells cover urban areas that can encompass hundreds of vehicles. Hence, the *VANET* scheme proposed herein is highly effective in supporting massive FCD collection.

Another performance advantage offered by the proposed approach is appreciated by examining the results shown in Figure 5.8. The metric M_{RB} is plotted vs. R_{eNodeB} for two different vehicular density levels, in the New York and Rome scenarios. We

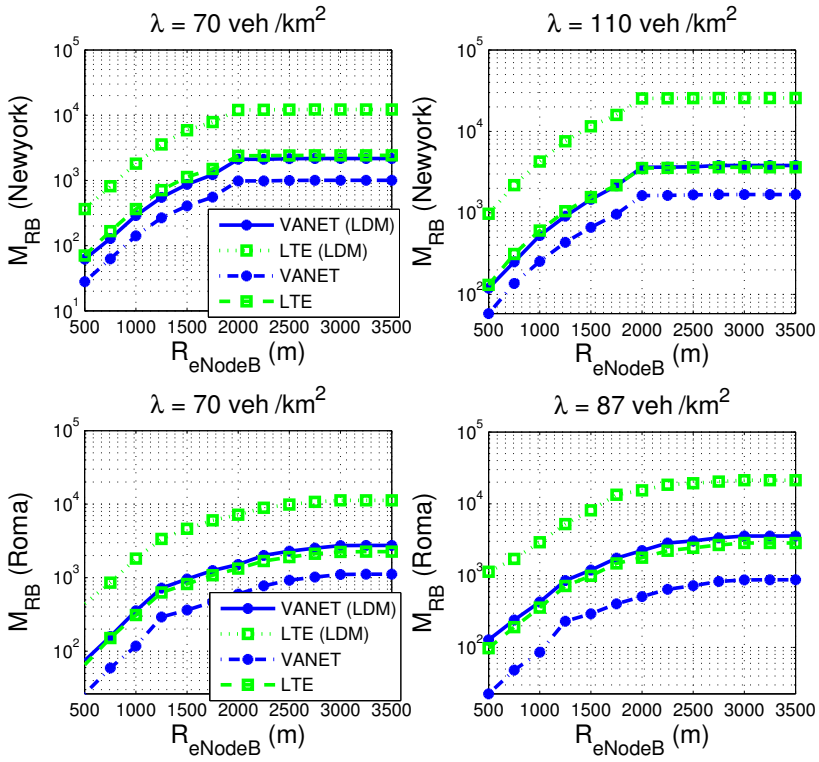


Figure 5.8 – Average M_{RB} vs. R_{eNodeB} for two density levels, New York map (top graphs) and Rome map (bottom graphs). The green curves (square markers) refer to the case where full connectivity information is transferred, in addition to node positions. The blue curves (circle markers) correspond to the case where only nodal position information is transferred [100].

compare the adaptive LTE channel performance obtained under the *VANET* and *LTE* approaches under two alternative cases: i) only vehicular positions are reported to the remote server via LTE (the same case as the one shown in Figures 5.5 and 5.6); ii) both vehicular positions *and* VANET connectivity information are reported to the server (curves denoted with *VANET (LDM)* and *LTE (LDM)*). The latter case is appealing for a centralized optimization of inter-vehicular communications and, in general, whenever the knowledge of the VANET topology can be exploited profitably.

We can notice that the advantage of our approach is enhanced when it is required to transfer information that includes nodal positions as well as their connectivity relationships within the VANET. In fact, under our *VANET* approach, this amounts to transfer the full list of neighboring nodal IDs and positions, rather than only those listed in the reduced table R_{LDM}. Hence, the difference is impacted by the number

of common neighbors of adjacent representative nodes. Under the LTE framework, the knowledge of the VANET connectivity requires each vehicular node to report information about itself plus the full list of its neighbor's IDs and positions.

Evaluation of LTE Cellular System Load: COLLECTION Phase

Another interesting study is the impact of the vehicular data gathering on the LTE access network during the COLLECTION phase. We compare three approaches (the acronyms are used as labels in the graphs):

LTE Each vehicle sends its own FCD information directly to eNodeB in a dedicated LTE channel.

VANET Representative nodes are elected by means of a VANET driven process (the REQUEST message dissemination described in Section 2.3). Then, representative nodes are responsible to send their R_{LDM} to their covering eNodeBs, thus reporting their own FCD plus those of part of their vehicular neighborhood.

HYBRID This is a state-of-the-art Hybrid approach as described at the beginning of Section 5.1. The information about all monitored vehicles in the target area, collected at the server during the SETUP phase, is used to synthesize a set of cluster-head nodes that cover all the target area. The designated cluster heads are responsible to aggregate and send the FCD of their respective vehicular neighborhoods. Note that the identification of a set of cluster head nodes covering all other vehicles requires the server to acquire the entire VANET connectivity graph.

More in depth, we have implemented the following scheme to select cluster heads according to the *HYBRID* approach. First, we order the vehicle nodes inside the coverage area of a eNodeB by decreasing levels of CQI. Second, the node with the highest CQI level is elected cluster head. This node and all its neighbors are removed from the list of the nodes under the coverage of the considered eNodeB. We iterate this selection process until the list is empty. Note that the cluster heads are elected on the basis of their LTE channel quality, maximizing the cluster head LTE radio capacity. Moreover, the election process guarantees that there is no duplication of FCD reported to the remote server, thus minimizing the overall amount of information to be sent through LTE channels. Clearly, the implementation of the cluster head election according to the Hybrid process requires that full VANET topology information be collected during the SETUP phase. MCS is set according to the transmitting node CQI.

Figures 5.9 and 5.10 compare the load induced on the LTE access network by the three data collection approaches listed above: *LTE* (green square markers); *VANET*

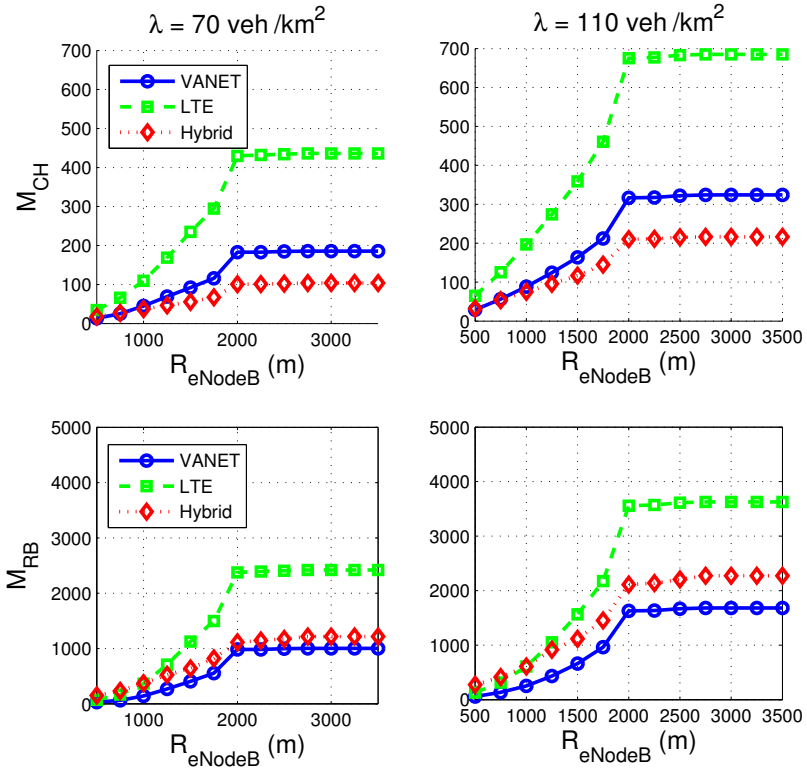


Figure 5.9 – New York scenario: average M_{CH} (top graphs) and M_{RB} (bottom graphs) vs. R_{eNodeB} for two different vehicular density levels [100].

(blue circle markers); *HYBRID* (red diamond markers). The obvious result shown in these figures is that using VANET communications (as done in the *VANET* or *HYBRID* approaches) we can drastically reduce the number of PUSCHs and RBs occupied in each LTE cell, in comparison with the *LTE* approach. The performance gap between the *LTE* and the other two approaches broadens as the number of eNodeBs is reduced. These results make a strong case for the exploitation of peer-to-peer communication networks among vehicle nodes, as allowed by the DSRC VANET, to aggregate FCD before sending them through the LTE cellular network.

The less obvious result is that *HYBRID* turns out to use a smaller number of LTE channels with respect to *VANET*, whereas the latter consumes a smaller amount of RBs to carry the FCD with respect to *HYBRID*. This apparent contradiction is explained as follows. According to *HYBRID* algorithm, representative nodes are chosen so as to obtain a sparse, yet full coverage of the vehicle nodes in the target area. There is no requirement that representative nodes form a connected network, only that each given vehicle node can communicate with one representative node.

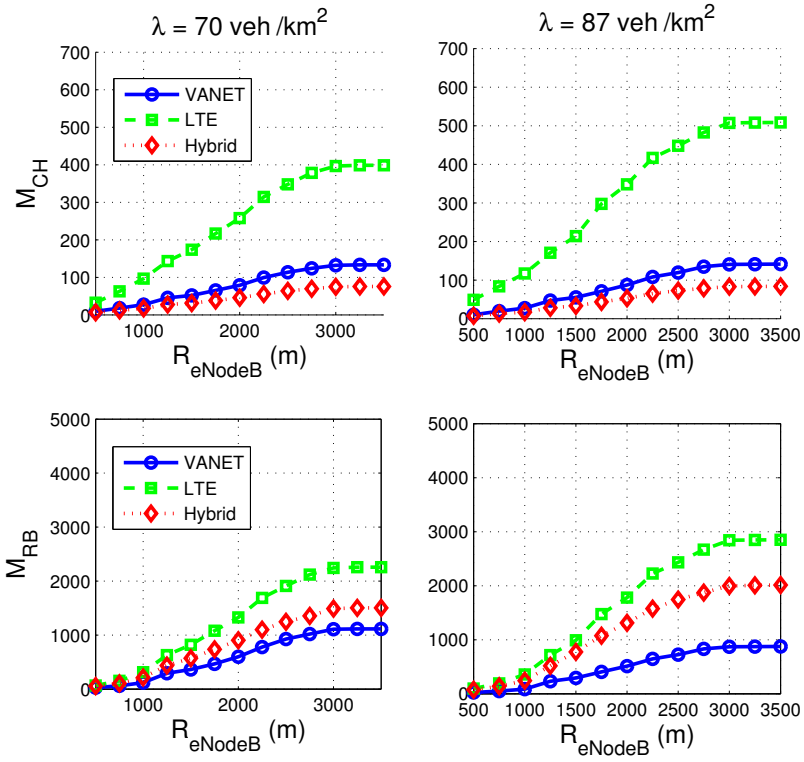


Figure 5.10 – Rome scenario: average M_{CH} (top graphs) and M_{RB} (bottom graphs) vs. R_{eNodeB} for two different vehicular density levels [100].

On the other hand, with *VANET* the identification of the representative nodes is driven by the dissemination logic, hence they form a connected set. As a matter of example, in a span of road of length L, the number of representative nodes is in the order of $L/(2d_{\max})$ with *HYBRID*, while it is in the order of $L/d_{\max}$ in case of *VANET*. Conversely, the number of vehicle nodes that each representative node has to report on is smaller with *VANET* than with *HYBRID*, hence less RBs are enough to carry the FCD in case of *VANET* with respect to *HYBRID*.

Evaluation of LTE Cellular System Load: Multiple Originators

Here we study the impact of our *VANET* solution when more than one vehicle generate the REQUEST message. This is the case with the election protocol in the multiple connected components case. Note that REQUEST messages are instrumental to identify the elected representative nodes, hence the originating trigger node is irrelevant. This implies that a REQUEST message originated from source node X can

inhibit a vehicle node that has received a REQUEST message originated by another node Y and is currently running its timer.

We consider the same simulation scenarios as in the collection phase, for New York and Rome and for two different vehicular density levels. In each scenario we select randomly n_t trigger vehicle nodes that are responsible to generate a REQUEST message. The trigger vehicle nodes are chosen randomly in different areas of the map, e.g., in case of 4 trigger nodes the map is divided into four quarters and one vehicle node is elected randomly within each map quarter. We have run simulations for $t = 1 \ldots 5$.

Figure 5.11 shows the performance of our proposed approach (*VANET*), compared with the *LTE* and *HYBRID* solutions, during the COLLECTION phase in New York. The plotted metrics are the number of LTE channels and RBs used by nodes sending FCD via the LTE access network, according to the three approaches. Figure 5.12 shows the same performance metrics in case of the Rome map. The number of trigger vehicle nodes originating REQUEST messages is annotated in the graph and ranges between 1 and 5. We can notice that the approaches making use of the VANET

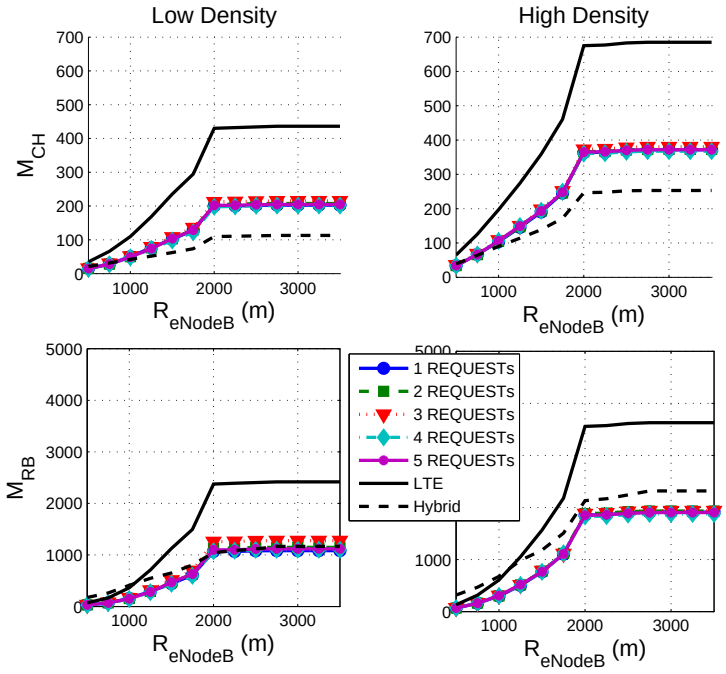


Figure 5.11 – New York: average M_{CH} (top graphs) and M_{RB} (bottom graphs) vs. R_{eNodeB} for two different vehicular density levels during the COLLECTION phase [100].

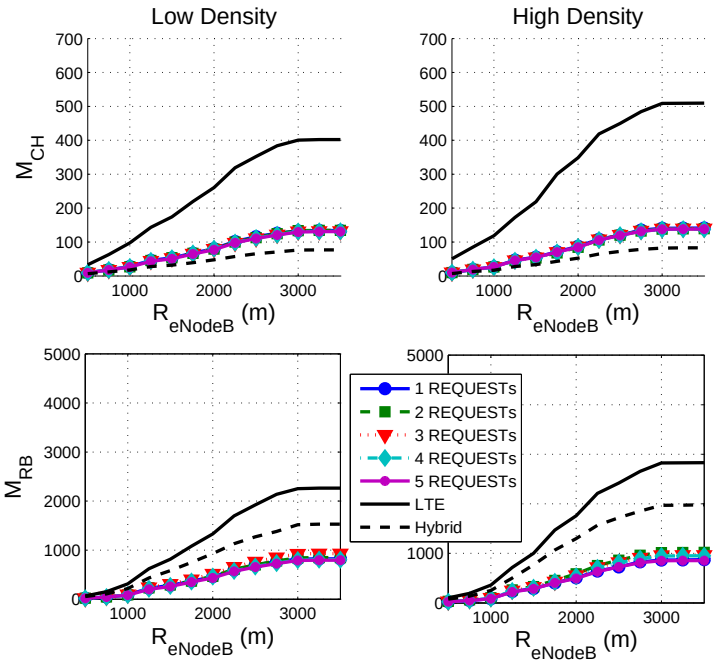


Figure 5.12 – Rome: average M_{CH} (top graphs) and M_{RB} (bottom graphs) vs. R_{eNodeB} for two different vehicular density levels during the COLLECTION phase [100].

communication links, namely *VANET* and *HYBRID*, attain performance levels close to one another. They definitely outperform the *LTE* solution.

The main result that we can deduce from this set of simulations is the low sensitivity of the considered performance metrics with respect to the number of the REQUEST originators. This is particularly important for the following considerations. In our approach, we use the VANET to disseminate the so called REQUEST message to trigger the FCD collection process and, as we have shown in our previous work [99], this mechanism works well when the REQUEST generation is centrally performed using a fixed access point called RSU. In our analysis we have considered the case when no specific additional VANET infrastructure node is present, delegating the whole work to the moving vehicles. The first set of simulations shows that our idea is also able to outperform traditional solutions, like *LTE* and *HYBRID*, to off-load the cellular network. We also show that the centralized scenario vision guaranteed by an RSU is not necessary for a good REQUEST dissemination process, or in other words, vehicles can be directly used to trigger the FCD collection process. Also, we are not bounded to use a particular stringent algorithm with a full knowledge of the network, being able to select the "best" vehicle node among a multitude. On

the contrary, the impact of having multiple trigger nodes scattered at random in the target area is marginal, i.e., the grid of representative nodes that emerges out of the REQUEST message dissemination exhibits robust performance levels with respect to the position of the initial trigger nodes.

5.2 On-the-Fly Distributed Clustering Formation

The existing clustering algorithms for heterogeneous vehicular networks rely on DSRC as the main technology for IVC and cluster creation, while LTE is used by the selected cluster heads to periodically report the aggregated information to the remote server (see Figure 5.13). In this context, the parameters used in the cluster head selection process become very important. The number of DSRC neighbors is the most used parameter when the objective is to minimize the LTE channel utilization. The reason is that the DSRC connectivity parameter helps in minimizing the number of cluster heads accessing the LTE channel, hence reducing the packet header overhead.

Although the DSRC connectivity parameter cannot be ignored, as we further show in this section, using it as the only parameter for electing cluster head vehicles turns out to be suboptimal. We identify another relevant parameter that has to be used in the selection process in order to further reduce the LTE channel utilization, namely the CQI in the LTE uplink [105]. We prove that for choosing the *right* vehicles to act as cluster heads, both DSRC connectivity and CQI parameters have to be used in the election process, as well as some jitter [107] – a randomly varying timing that aims at preventing vehicles from simultaneous transmissions. To this purpose, we propose a distributed clustering algorithm that combines the above mentioned parameters. The proposed solution does not require any a priory knowledge (e.g., road intersection coordinates) or dedicated infrastructure.

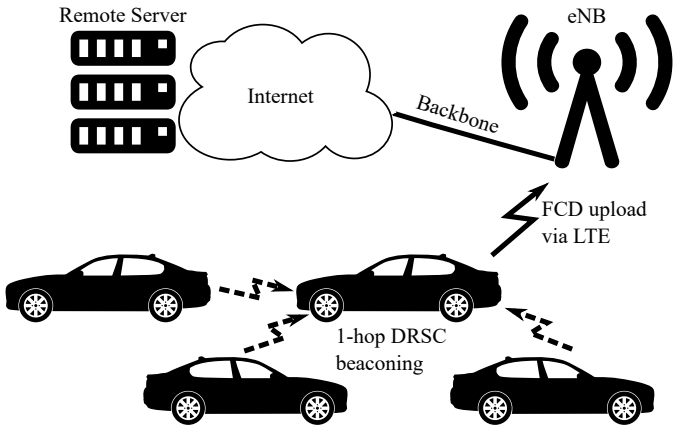


Figure 5.13 – FCD collection scenario.

The work presented in this section is based on our article published in the *Proceedings of 8th Vehicular Networking Conference (VNC 2016)* [26]. We first present a sample application used in our study, showing a simple LTE-based data collection algorithm. Then we describe the On-the-Fly Clustering (OFC) algorithm proposed in [26], as well as its extended version, named OFC with Duplicate Suppression (OFCDS). OFC and OFCDS use both LTE and DSRC technologies to collect data in a heterogeneous vehicular network. Finally, we show how OFC can be turned into a baseline state-of-the-art distributed clustering algorithm that is basing its cluster head selection procedure on the current number of DSRC neighbors only.

5.2.1 PureLTE Algorithm

We consider a traffic monitoring system as use case example for our study, but any other application that needs periodic exhaustive collected information is relevant. We assume that every vehicle inside the target area has LTE communication technology available on board. The application itself consists in periodically reporting FCD messages via LTE to the traffic monitoring system server. The updating frequency, which is common to all vehicles, is decided by the traffic monitoring system and is set up in the collection interval parameter (I_{col}) by every vehicle (i.e., when the application starts, it can immediately send a request to the remote server via LTE asking for the desired reporting frequency).

A simple algorithm that periodically collects FCD messages in such a scenario is presented in Figure 5.14. We will further refer to this approach as PureLTE. Basically, whenever the application starts, it periodically schedules a time-out event, named I_{out}, equal to the collection interval parameter. When the time-out expires, the application sends a *Data* message via LTE to the traffic monitoring system server containing updated information about the vehicle itself. Notice that a *Data* message can contain one or more FCD messages. In this particular case *Data* consists of only one FCD message created by the transmitting vehicle itself, since no IVC communication is

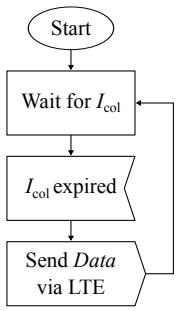


Figure 5.14 – PureLTE data collection algorithm [26].

present. The transmissions are not synchronized among different vehicles. The only common information that must be known to all vehicles is the parameter I_{col}.

Although this approach is very simple, it implies that every vehicle has to periodically report its FCD, which can introduce a high load over the LTE channels, especially in the case of urban scenarios with high vehicle density [108]. Considering that many different vehicular applications, as well as all regular LTE traffic, will have to share the same limited LTE bandwidth provided by the mobile and network operators, this issue becomes even more critical.

5.2.2 OFC Algorithm

The FCD collection application assumes each vehicle maintains an LDB where relevant information about the vehicle itself and about its current neighbors is stored. A background exchange of one hop messages on DSRC keeps the LDBs up to date. When the time comes for sending a report, the elected forwarding vehicle reads its current LDB content and sends it to the remote server. An example of such a process is already envisaged explicitly by the ETSI standards, where the CAMs exchanged among neighboring vehicles and the Local Dynamic Map [35] data base are defined to maintain vehicle awareness of the surrounding vehicular traffic environment. We do not pursue the details of the LDB maintenance further, since this has been widely investigated in the literature (e.g., see [109][89]).

The main idea behind OFC is to allow only a subset of vehicles, named *forwarders*, to report via LTE their own, as well as their one hop neighbors' FCD messages. These forwarders are dynamically selected during every collection interval. The selection process itself is based on synchronized selection phases and takes into account the current number of DSRC neighbors, the CQI in the LTE uplink information, and a uniformly distributed random jitter. OFC operation is highlighted in Figure 5.15. Unlike the PureLTE approach, where no synchronization is needed since no IVC is present, with OFC the time instance when the collection interval starts must be the same for all vehicles. Although the forwarder selection mechanism is performed locally, it has to start at the same point in time for all vehicles, since the considered parameters have to refer to the same time instance. Hence, every vehicle is periodically computing the next collection interval according to

$$T_{\text{col}} = T_{\text{cur}} - (T_{\text{cur}} \bmod I_{\text{col}}) + I_{\text{col}} \tag{5.2}$$

where T_{col} is the point in time when the collection interval starts, T_{cur} is the current time instance (i.e., we assume every vehicle has a GPS device on board which can provide the current time) and I_{col} represents the collection interval span.

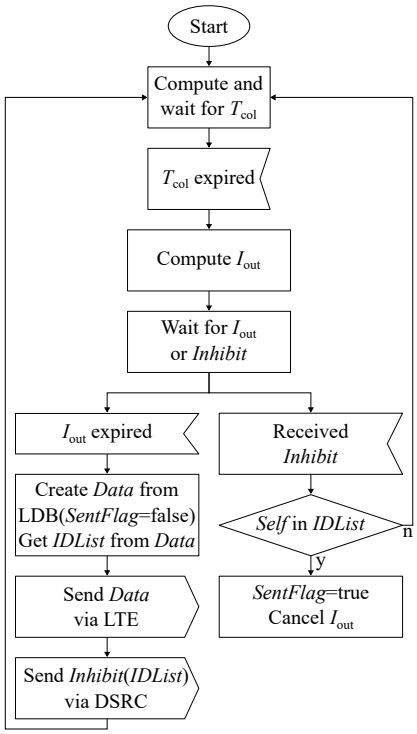


Figure 5.15 – OFC data collection algorithm.

Upon collection interval starting, every vehicle computes its own sending time T_{send} according to

$$T_{\text{send}} = T_{\text{col}} + I_{\text{out}} \tag{5.3}$$

where the time-out interval I_{out} is given by

$$I_{\text{out}} = I_{\text{col}}(\alpha X + \beta Y + \gamma Z) \tag{5.4}$$

Here α, β, and γ are non-negative weights chosen so as that $\alpha + \beta + \gamma = 1$, and $\alpha, \beta, \gamma \in [0, 1]$. X, Y, and Z represent the DSRC connectivity, the CQI in the LTE uplink, and the jitter respectively and are computed as

$$X = 1 - \frac{N_{\text{cur}}}{N_{\text{max}}} \tag{5.5}$$

$$Y = 1 - \frac{Q_{\text{cur}}}{Q_{\text{max}}} \tag{5.6}$$

$$Z = \mathscr{U}(0, 1) \tag{5.7}$$

where N_{cur} and Q_{cur} represent the current number of one hop DSRC neighbors and the current CQI in the LTE uplink of a generic vehicle (in case of subband-level CQI

reporting, the average value over all subbands is considered), while $N_{\max}$ and $Q_{\max}$ are the corresponding maximum values. Notice that $Q_{\max}$ refers to the maximum CQI index, which is globally known to all vehicles, while $N_{\max}$ is locally computed by every vehicle. In particular, N_{cur} is included in the beacon exchange process, meaning that every vehicle knows the number of neighbors for each one of its one-hop DSRC neighbors. At this point a vehicle can compute $N_{\max}$ by finding the maximum N_{cur} value among all its neighbors.

It is important to study the impact that each of the three considered factors has. This is why we introduce three weight parameters, namely α, β, and γ, which are needed for tuning the considered factors (see Section 5.2.7). According to Equation (5.4), vehicles having a higher number of DSRC neighbors and a better CQI in the LTE uplink are scheduled for transmission first. Vehicles whose time-out expire, become forwarders and prepare their *Data* message to be sent to the traffic monitoring system by reading their LDB. Immediately after sending the *Data* message via LTE, a forwarder informs its neighbors by broadcasting an *Inhibit* message over the DSRC network, containing the identifiers of all vehicles whose FCD was enclosed in *Data*. If a vehicle waiting for its time-out to expire receives an *Inhibit* message, it checks whether its identifier is present. If this is true, then it immediately cancels the time-out I_{out}, aborting its scheduled transmission. According to this approach, once a vehicle becomes a forwarder, all its 1-hop DSRC neighbors are inhibited.

Notice that an inhibited vehicle can be in the transmission range of more than one forwarder, meaning that multiple copies of the same FCD message can be sent to the server, increasing the LTE channel utilization. OFC has a duplicate suppression mechanism that takes advantage of the already existing beacon exchange process. In particular, the beacon messages sent in background are extended with a flag, named *SentFlag*. At the beginning of each collection cycle, vehicles set their *SentFlag* to FALSE. As soon as a vehicle *A* receives an inhibition message from a neighbor, announcing that the neighboring vehicle has reported *A*'s FCD message to the remote server, *A* turns its flag to TRUE. Whenever a vehicle node updates the application information by sending a message to its neighbors, it includes the current value of its *SentFlag*. As a consequence, updates of the application data sent by the inhibited vehicle *A* every I_{beacon} seconds carry the flag set to TRUE and cause the relevant information to be updated in the LDBs of *A*'s neighbor vehicle nodes. If any of those neighbors report their *Data* to the remote server, they will exclude *A*'s FCD.

5.2.3 Baseline algorithm

Current state of the art solutions consider the DSRC connectivity as the main parameter in the forwarder election mechanism [24], [25]. These are usually heuristics

for finding approximations to the Minimum Dominating Set problem, that aim at maximizing the offloading level by minimizing the number of forwarders.

OFC structure has the flexibility to be easily turned into such algorithm, that we will further refer to as Baseline, by using only the number of DSRC neighbors as the main parameter, while not using the CQI in the LTE uplink at all. This can be done by properly setting the values for the weighting factors α, β, and γ. For instance, by setting $\beta = 0$ we end up having a heuristic which is minimizing the number of forwarders by selecting those vehicles with the highest number of DSRC neighbors. Notice that we still keep $\gamma = 0.2$ (i.e., jitter) to reduce simultaneous transmissions and obtain a fair comparison with the other considered solutions.

The features of the Baseline algorithm are similar to the *RB* clustering mechanism proposed by Stanica et al. [24] and described in Section 2.5. According to this solution each vehicle transmits in a slot selected at the beginning of every collection interval, chosen among N_s available, only if no other neighboring vehicle transmitted its information first. However, since the authors do not specify the slot selection criterion, we can assume that with this heuristic a vehicle selects its own slot uniformly at random. In [26] we proved that a clustering algorithm that considers only the randomness factor in the forwarder selection process turns out to be sub-optimal. For this reason, in our implemented version of the Baseline algorithm we consider the number of DSRC neighbors by giving priority to those vehicles that have more DSRC neighbors.

5.2.4 OFC with Duplicate Suppression

The effectiveness of the *SentFlag* mechanism described in Section 5.2.2 depends on the ratio between the time interval I_{beacon} of the background application LDB periodic update and the data collection time interval I_{col}: the smaller $I_{\text{beacon}}/I_{\text{col}}$, the more effective the *SentFlag* mechanism. However, for applications that need frequent information updates from the vehicular network, this mechanism can be less effective in preventing the transmission of duplicate messages on LTE, causing a higher resources utilization. For this reason, we extend here the OFC algorithm with a new duplicate suppression scheme that does not depend on the $I_{\text{beacon}}/I_{\text{col}}$ ratio. The main idea behind this new approach, named OFCDS, is to immediately disseminate the IDs of all 1-hop DSRC neighbors whose information is being sent on LTE by an elected forwarder to all its 2-hop DSRC neighbors.

OFCDS operation is displayed in Figure 5.16. In particular, the new parts of the extended algorithm, as well as the modified parts from Figure 5.15, are represented by the gray blocks. In Algorithm 5.1 we present the pseudo-code of our proposed solution. Notice that the inhibition mechanism in OFCDS is similar to OFC. In particular, once a vehicle is elected as forwarder, it immediately broadcasts an *Inhibit*

1: v - current vehicle
2: I_{out} - the timeout for selecting forwarding vehicles
3: LDB - the local data base containing updated beacons
4: L - a list of neighboring vehicle IDs whose FCDs were sent to the server by the elected forwarding vehicle
5: $Data$ - a message containing the aggregated FCDs to be sent to the server via LTE
6: $Inhibit$ - a message to be sent on DSRC by the elected forwarder to inhibit the 1-hop neighbors; this message includes the list L and the list of selected relay vehicles
7: $Notify$ - a message to be sent on DSRC by the selected relay vehicles to notify the 2-hop neighbors about L
8: OHN_i - the list of 1-hop DSRC neighbors of vehicle i
9: THN_i - the list of 2-hop DSRC neighbors of vehicle i

10: **upon event** I_{out} expired **do**
11: $sentNeighbors =$ getSentNeighbors(LDB)
12: deleteSentNeighborsFromData($sentNeighbors$)
13: sendToServer($Data$)
14: $selectedRelays =$ selectRelays()
15: $Inhibit.L = L$
16: $Inhibit.Relays = selectedRelays$
17: broadcastOnDSRC($Inhibit$)
18: **upon event** $Inhibit$ received **do**
19: cancel(I_{out})
20: $selectedRelays = Inhibit.Relays$
21: **if** $v \in selectedRelays$ **then**
22: $Notify.L = Inhibit.L$
23: broadcastOnDSRC($Notify$)
24: **end if**
25: **upon event** $Notify$ received **do**
26: $sentNeighbors = Notify.L$
27: setSentNeighborsInLDB($sentNeighbors$)
28: **function** SELECTRELAYS()
29: **while** THN_v is not empty **do**
30: $u = \text{argmax}_{u \in \text{OHN}_v} \text{THN}_v \bigcap \text{OHN}_u$
31: $selectedRelays$.insert(u)
32: $\text{THN}_v = \text{THN}_v - (\text{THN}_v \bigcap \text{OHN}_u)$
33: **end while**
34: return $selectedRelays$
35: **end function**

Algorithm 5.1 – Duplicate suppression algorithm

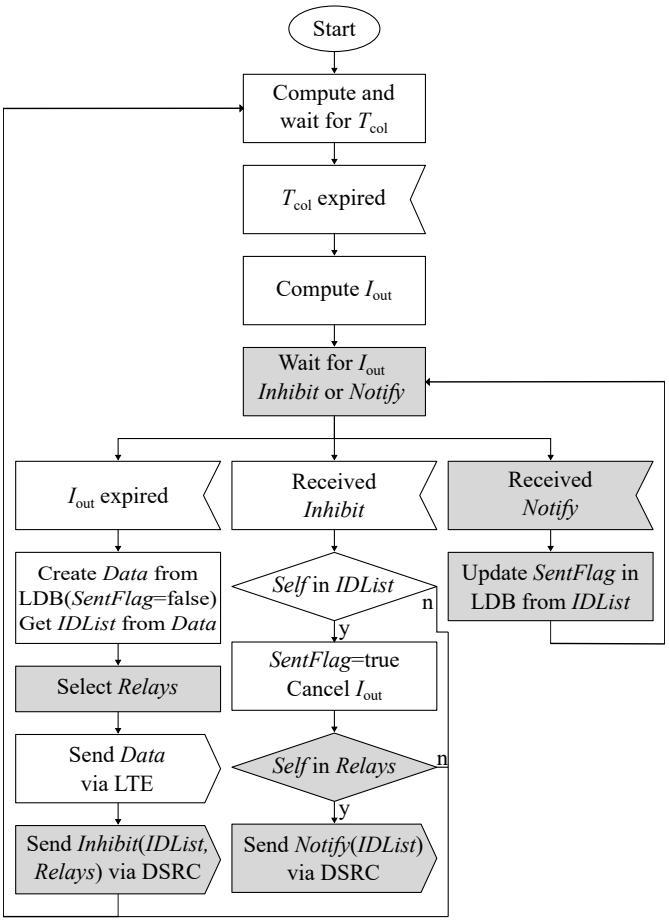


Figure 5.16 – OFCDS data collection algorithm.

message containing the ID list L of its 1-hop DSRC neighbors whose information was included in the *Data* message. The main difference is in the behavior of the vehicles that receive the *Inhibit* message. Unlike OFC, where the inhibited vehicles wait for the next beaconing opportunity to inform their neighbors about the fact that their information was already sent, in OFCDS the inhibited vehicles must disseminate as soon as possible to all their neighbors the full list L received from the elected forwarder. They do this by broadcasting a *Notify* message including the list L, with a small random delay to avoid simultaneous transmissions. Notice that *Inhibit* is telling which vehicles should cancel their I_{out} timers, while *Notify* is only informing about the inhibited vehicles, so that other potential forwarders can exclude the corresponding FCDs from their *Data* messages.

However, to avoid the congestion of the DSRC channel, not all inhibited vehicles are broadcasting the *Notify* message. The idea is that the elected forwarders select a subset of their 1-hop DSRC neighbors to be in charge of sending such message.

The relay selection procedure (*selectRelays()* function in Algorithm 5.1) consists in iteratively selecting a subset of 1-hop neighbors, such that all 2-hop neighboring vehicles are covered. This is possible only if all vehicles have 2-hop DSRC awareness. We achieve this by extending the beacon structure with the current ID list of 1-hop DSRC neighbors. Notice that while the standard beacon has a fixed constant size, the extended beacon size depends on the current number of DSRC neighbors.

A time diagram of the OFC and OFCDS algorithms is shown in Figure 5.17. Assume vehicle *A*'s timeout expired first, hence *A* is elected as a forwarder. Besides sending the *Data* message, *A* immediately broadcasts an *Inhibit* message containing the ID list of its neighbors whose information was just included in *Data*. Such behavior is common to both OFC and OFCDS. At this point, if OFC runs on top of *B*, then *B* will wait for the next beacon opportunity to sent its own *Beacon* message containing the *SentFlag* equal to TRUE. Of course, *C*'s timeout can expire before *C* receives all the updated flags from its neighbors, including *B*. This means that *B*'s FCD will be sent via LTE by both *A* and *C*. On the other hand, with OFCDS, assuming that *B* was elected by *A* as one of the relay nodes, it will immediately re-broadcast the ID list received from *A*, so that all *A*'s 2-hop neighbors, including *C*, know whose information was sent by *A* to the remote server.

5.2.5 Simulation Framework and Scenario

For evaluating the proposed algorithms we use the simulation framework described in Section 2.4 composed of SUMO [54], OMNeT++ [53] and *Veins LTE* [57], an LTE

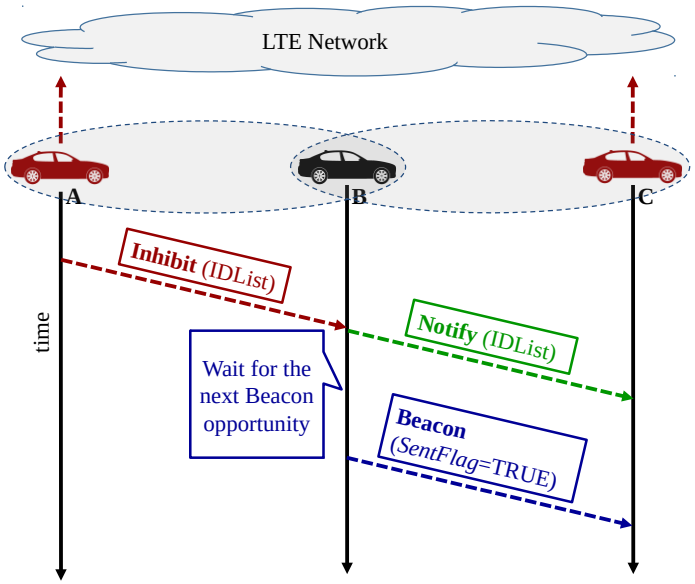


Figure 5.17 – Time diagram of the OFC and OFCDS algorithms.

extension of the well-known open source vehicular network simulator *Veins*[14] [55]. A realistic Manhattan grid scenario is considered for our simulations, created using real Manhattan downtown road and building dimensions (see Figure 5.18). Krauss vehicular mobility model is used, along with the random trips traffic flow origin-destination model. Although the vehicular mobility is simulated over a larger area, we enclosed the observed region to a smaller target area to avoid border effects. Also, we use the free-space path loss ($\alpha = 2$) with obstacle shadowing [88] models for DSRC, and Urban Macro path loss [45] with Jakes multi-path fading models for LTE.

We assume LTE coverage is available inside the target area. All vehicles are equipped with DSRC and LTE wireless network interfaces, while the decision whether to send a packet on one interface or on another is taken at the application layer. Considering that most likely the mobile operators will dedicate only a small portion of bandwidth to vehicular applications, for our analysis we assume a bandwidth of 3 MHz (15 available RBs). Since different traffic monitoring systems, but also other applications, might have particular requirements in terms of data reporting frequency, we analyze and compare the performance of the three considered solutions with respect to different collection intervals.

All simulations are run for 100 s preceded by 400 s of warmup time. Every simulation is repeated 25 times with independent random number seeds. The most relevant simulation parameters are displayed in Table 5.3.

5.2.6 Evaluation Metrics

We propose and compare two different implementations of OFCDS. The first one, named OFCDS-Ideal, is an idealistic implementation of the algorithm, where the

[14] `http://veins.car2x.org`

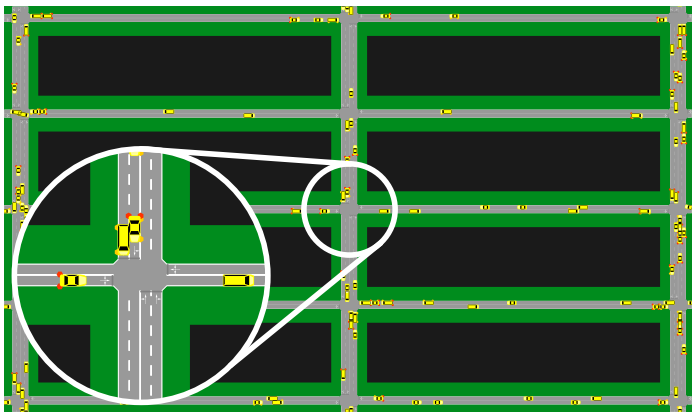


Figure 5.18 – Part of the simulation scenario.

Parameter	Value
Simulated area	620 m × 530 m
Average number of vehicles	165 and 390
Average density (veh/km/lane)	11 and 26
Simulation duration	100 s
I_{col}	1 s, 3 s, 5 s, 10 s, 15 s and 20 s
Baseline α, β, and γ	0.8, 0, 0.2
OFC and OFCDS α, β, and γ	0.3, 0.5, 0.2
IVC technology	IEEE 802.11p
IVC maximal transmit power	20 mW and 100 mW
DSRC beacon frequency	1 Hz and 10 Hz
Beacon size	400 B
Vehicle ID size S	6 B
Number of available RBs	15
LTE scheduler	MAXCI
UE transmission power	26 dBm
eNodeB transmission power	45 dBm

Table 5.3 – Network and simulation parameters.

2-hop DSRC awareness is assumed to be obtained without additional load on the DSRC communication channel. We assume here that the size of the list containing the IDs of the current neighboring vehicles, included in each beacon, is constant and has a negligible size with respect to the beacon length (e.g., 10 B for the ID list size with respect to 400 B for the beacon size). The second, named OFCDS-Real, is a realistic implementation where the size of the ID list depends on the actual number of vehicle IDs that are included in this list. For instance, let A be a generic vehicle and N_A its current number of DSRC neighbors. Then, the additional payload added to A's beacon is $N_A S$, where S is the size of a vehicle ID entry. Both OFCDS implementations are evaluated for two different vehicular densities and compared against OFC, PureLTE, and Baseline.

The aim of this evaluation is to measure the performance and the influence of the proposed solutions on both LTE and DSRC communication channels. The main evaluation metrics are defined in Table 5.4. On LTE, we are interested in measuring the RB utilization in uplink, defined as the average percentage of used RBs requested to transfer the FCDs of all vehicles roaming inside the area of interest. There are two main causes for the RB utilization: (i) the amount of transferred information on LTE, which depends on the actual payload coming from the number of FCD messages, and the overhead induced while transferring this information; (ii) the LTE channel quality of the transmitting vehicles. We tackle the second cause with our proposed solutions by considering the CQI in the LTE uplink in the forwarder election process. As for the first cause, we can act only on the generated overhead, which is coming from the network and transport layer headers, and the duplicate messages induced by the heterogeneous algorithms. We address these issues by significantly reducing the number of vehicles accessing the LTE network, that send

Metric	Definition
RB Utilization	Mean percentage of allocated RBs from the total number of available RBs per each Transmission Time Interval (TTI)
Duplicate Ratio	Mean ratio of the number of duplicate messages to the total number of received messages in each collection interval
Delay	Time interval between the moment when the FCD message is generated and the time instant when the same message arrives at the remote server
Inter-Arrival Time	Time difference between two consecutive FCD message receptions at the server belonging to the same vehicle
CBR	Mean ration of the total time a vehicle senses the DSRC channel busy to the total simulation time

Table 5.4 – Performance evaluation metrics

aggregated FCD messages, and reducing the duplicate ratio, defined as the number of duplicate messages over the number of total received messages by the remote server.

Of course, reducing the RB utilization comes with a cost, which is payed in terms of information transferring delay, defined as the time interval between the moment when the FCD message is generated and the time instant when the same message arrives at the server. Moreover, we are interested in quantifying the variability of the arrived information, which is why we measure the inter-arrival time of the reported FCDs, defined as the time difference between two consecutive FCD message receptions belonging to the same vehicle being received at the remote server. At the same time, all the heterogeneous approaches introduce some load on the DSRC channel, which we measure as the average CBR experienced by each vehicle for the entire simulation period.

All these metrics are evaluated for different parameter configurations. In particular, we consider two vehicular densities, 11 and 26 veh/km/lane, with a lighter load on the DSRC channel: each vehicle's transmission power is set to 20 mW and the beaconing frequency to 1 Hz. To evaluate the proposed algorithms under a higher DSRC load setup, we modified the low vehicular density scenario by increasing the transmission power to 100 mW and the beaconing frequency to 10 Hz.

5.2.7 Parametric Study

Our intention is to include the most relevant parameters describing the communication capabilities. We can notice that there are three different parameters that can affect the performance of our algorithm: the number of DSRC neighbors, the CQI in

the LTE uplink and the jitter. What we are interested in is to assess the influence of CQI and jitter parameters when adding them to the DSRC connectivity.

To this purpose, we investigate the performance of our heterogeneous data collection algorithm in terms of RB utilization for different values of α, β, and γ, ranging from 0 to 1 with a 0.05 step, for a fixed collection interval $I_{\text{col}} = 10\,\text{s}$. Since the above mentioned parameters are not independent, we show only the values of β and γ, while $\alpha = 1 - \beta - \gamma$.

The results of our study are shown in Figure 5.19. The x-axis represents the parameter β (i.e., the influence of CQI in the LTE uplink), while the y-axis shows the percentage of the RB utilization. Because of visibility reasons we choose to plot only three values of γ, namely $\gamma = 0.2$, $\gamma = 0.5$, and $\gamma = 1$, although we simulated the entire range from 0 to 1. However, the other curves show similar behaviors.

The first thing that can be noticed is that the LTE channel utilization is higher for low values of β, meaning that the CQI parameter must be considered with a proportion of at least 10 % when designing clustering algorithms for reducing the RB utilization. Also, a slight increase can be seen for $\beta > 0.7$, which means that increasing too much the influence of the CQI in the LTE uplink and decreasing the weight of other parameters is not the best solution. The utilization of the LTE channel is minimized for $0.1 < \beta < 0.7$. We can notice that, if choosing an influence factor for the CQI inside this range, the RB utilization can be decreased to 70 % with respect to the case when not using the CQI at all ($\beta = 0$).

When looking at the jitter influence, we can see that varying γ between 0.2 and 1 does not affect too much the performance. In fact, we can notice that the curves overlap and their shapes are similar. However, according to our results (data not shown), not using jitter at all (i.e., $\gamma = 0$) increases the RB utilization up to 400 % with respect to $\gamma = 0.2$. This confirms the need of using at least some jitter. Moreover,

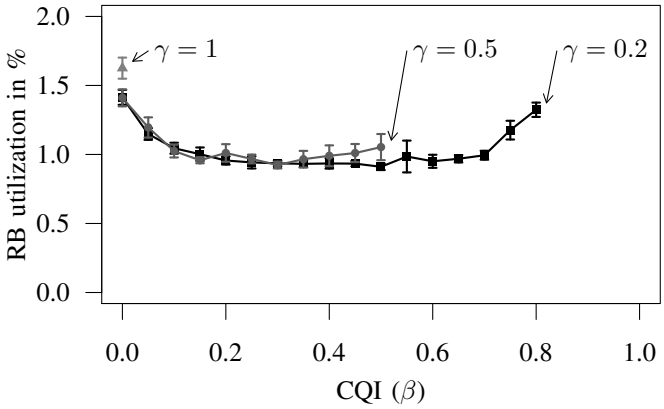


Figure 5.19 – OFC performance in terms of RB utilization as a function of β for different values of γ [26].

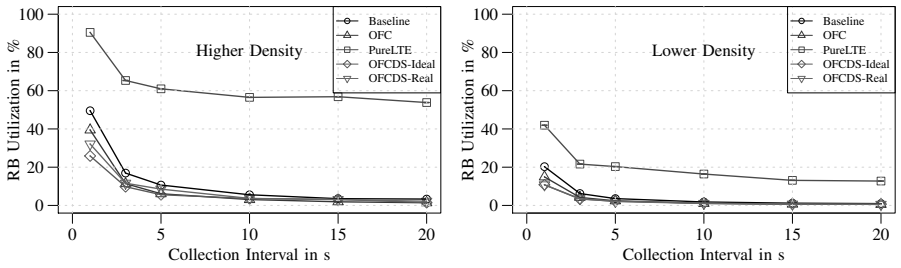
the curve shape for $\gamma = 0$ matches the other ones plotted in Figure 5.19, meaning that even in this case the CQI parameter helps in decreasing the RB utilization. According to these results, for the comparative performance evaluation in the next section we choose the following weight values: $\alpha = 0.3$, $\beta = 0.5$, and $\gamma = 0.2$.

5.2.8 Comparative Performance Evaluation

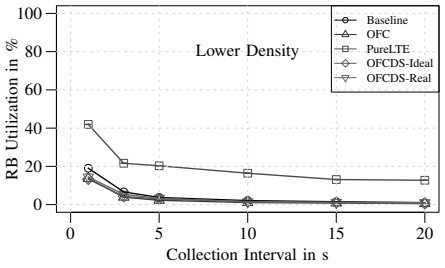
After finding the optimal parametric setup for OFC and OFCDS, we are interested in comparing our proposed solutions to other state-of-the-art protocols. To this purpose, we choose for comparison PureLTE, a simple solution described in Section 5.2.1, that collects FCD messages using the LTE technology only. Also, we compare our solution with Baseline, a state-of-the-art distributed protocol based on both DSRC and LTE communication technologies, that only considers the number of DSRC neighbors as the main parameter in the forwarder selection process.

Evaluation of Resource Blocks Utilization in LTE Uplink

The mean LTE RB utilization is depicted in Figure 5.20. We can notice that PureLTE is using the highest amount of RBs, independently from the considered collection interval and/or vehicular density, with a peak of 90 % used RBs for a collection interval of 1 s and a higher vehicular density. This is an expected result, considering



(a) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$. **(b)** $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.

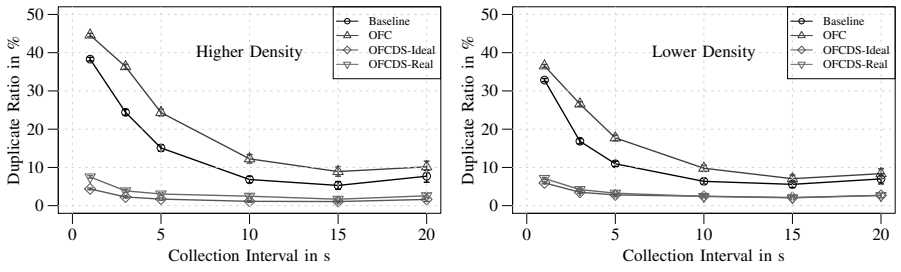


(c) $I_{\text{LDB}} = 10\,\text{Hz}$ and $P^{\text{tx}} = 100\,\text{mW}$.

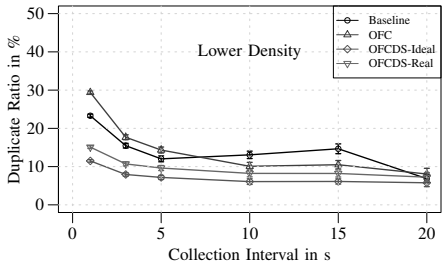
Figure 5.20 – The RB utilization as a function of the collection interval for two vehicle densities.

the fact that all vehicles roaming inside the area of interest are periodically accessing the LTE channel and requesting resources. On the other hand, for the same collection interval, all other algorithms that are exploiting the DSRC technology are significantly decreasing the LTE RB utilization, confirming the fact that the DSRC technology can help in decreasing the LTE channel utilization. Although for higher collection intervals the RB utilization is quite similar for these heterogeneous algorithms, the difference becomes noticeable when decreasing the I_{col} values. In particular, for $I_{\text{col}} = 1\,\text{s}$, Baseline drops down the RB utilization to 47 %, OFC to 37 %, OFCDS-Real to 27 %, while OFCDS-Ideal to 23 %. Notice that Baseline uses more resources than OFC and OFCDS, since it does not consider the CQI of the elected forwarders, meaning that these vehicles send more aggregated information while having a possibly very bad CQI, wasting much more resources. The same behavior can be observed for the lower vehicular density scenario, but with overall less used RBs. It is also worth pointing out that increasing the load on the DSRC channel does not affect the LTE resource utilization, as can be seen from Figures 5.20b and 5.20c.

Figure 5.21 displays the mean duplicates ratio for different collection intervals. In general, all the considered heterogeneous algorithms introduce duplicate messages. This is related to the DSRC network topology and can be explained by the *network assortativity* phenomenon [110] from complex network theory, which implies that directly connected nodes (i.e, nodes in the same neighborhood) are likely to have



(a) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$. **(b)** $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.



(c) $I_{\text{LDB}} = 10\,\text{Hz}$ and $P^{\text{tx}} = 100\,\text{mW}$.

Figure 5.21 – The duplicates ratio as a function of the collection interval.

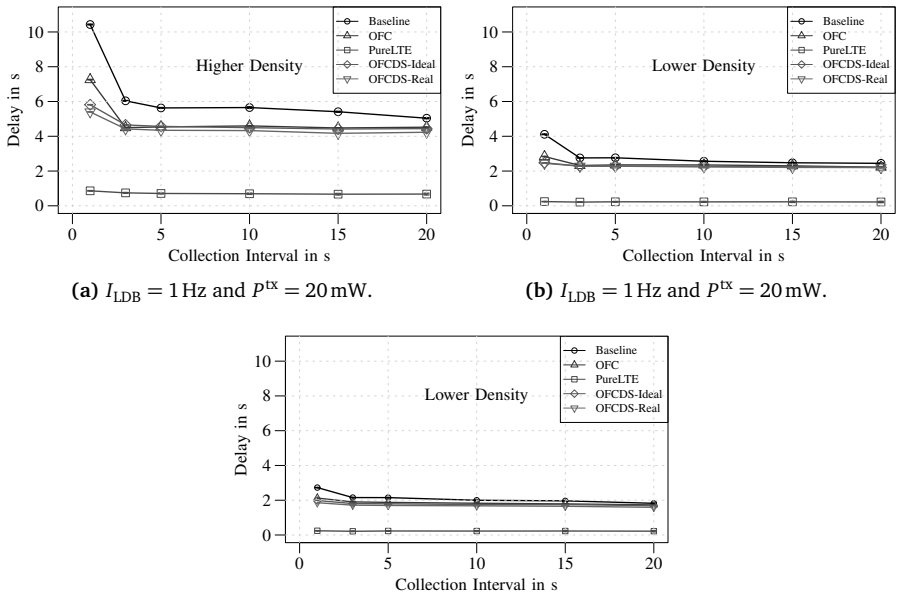
similar degree levels. On the other hand, with PureLTE each vehicle is sending its own FCD without generating duplicate messages, which is why we do not display it here.

A first observation is that OFC and Baseline increase the duplicates ratio for lower collection intervals, meaning that more information is being sent to the server. This confirms the fact that their inhibition mechanism is less efficient for greater $I_{\text{LDB}}/I_{\text{col}}$ ratios. Also, notice that Baseline induces generally less duplicates than OFC (e.g., roughly 45 % duplicates generated by OFC and 39 % by Baseline, when considering the higher density scenario and $I_{\text{col}} = 1\,\text{s}$). This is because Baseline gives priority to vehicles having more DSRC neighbors in the forwarder selection process, thus minimizing the number of forwarding vehicles. But, since OFC tends to elect as forwarders those vehicles with a better CQI in the LTE uplink, it is still able to utilize less resources with respect to Baseline, as can be seen from Figure 5.20. However, the greatest impact over the suppression of duplicates ratio is given by OFCDS. In particular, for the higher density scenario and the lowest collection interval, OFCDS-Ideal and OFCDS-Real generate only 4 % and 7 % duplicates correspondingly. The difference is less noticeable when we put more load on the DSRC channel (see Figure 5.21c), indicating the fact that the congestion slightly affects the performance of the duplicate suppression mechanism. The overall results, however, confirm the efficiency of the new proposed duplicate suppression scheme.

Delay Analysis

An important aspect to be investigated is how much time an FCD message needs to reach the server from the moment when it is generated. In Figure 5.22 we compare the considered algorithms in terms of FCD message delay for different collection intervals. We notice that the message transferring delay introduced by PureLTE is lower with respect to the considered heterogeneous approaches. This is mainly due to the fact that the size of the messages is smaller, since there is no aggregation (i.e., every vehicle is sending its own FCD message with a constant size). On the other hand, in case of a heterogeneous approach, the elected forwarders are sending aggregated messages, meaning that they send much more information, hence, needing more time to complete the transmission. Another reason is that in PureLTE the FCD message generation and the transmission starting time instances are the same, while in the heterogeneous algorithms the aggregated information to be sent via LTE has already an additional delay uniformly distributed between 0 and I_{LDB}. However, it is worth noting that we measure here only the data transfer delay, without considering the random access procedure.

Among the heterogeneous approaches, Baseline has the highest message delay. There are two main reasons for this: the first one is that Baseline does not consider



(a) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.

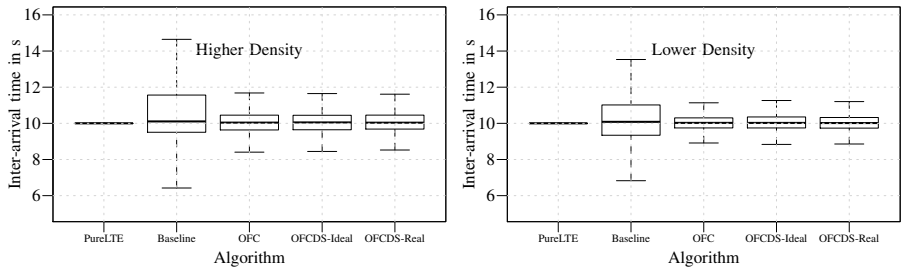
(b) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.

(c) $I_{\text{LDB}} = 10\,\text{Hz}$ and $P^{\text{tx}} = 100\,\text{mW}$.

Figure 5.22 – Information delay as a function of the collection interval.

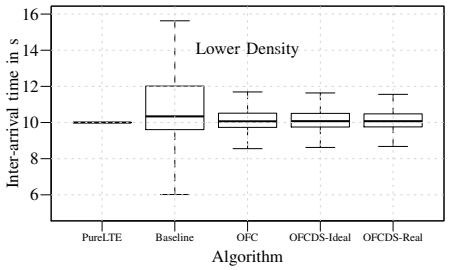
the CQI in the LTE uplink when electing forwarding vehicles, which not only leads to a higher RB utilization, as we see in Figure 5.20, but also to an increased transfer delay; the second reason is that Baseline is transferring more information due to the higher duplicates ratio (see Figure 5.21). The amount of the transferred information is precisely the reason why for the lower vehicular density scenario the average delays are generally smaller. Moreover, Figure 5.22c also suggests the fact that an increased load on the DSRC channel does not have any significant impact on the information delay. The delays induced by OFC and OFCDS are quite similar, with a slight difference for very low collection intervals, where the generated duplicates have a more negative impact on OFC with respect to OFCDS.

The FCD messages inter-arrival time is displayed in Figure 5.23 for a collection interval equal to 10 s. The values are grouped in box plots, where the box itself is representing the first and third quartiles, the median value is represented by means of a central line inside each box, while the whiskers are showing the maximum and minimum values. We can notice that all the considered algorithms have roughly the same median value, which is equal to the requested collection interval. This means that the information is arriving at the remote server with the same frequency as it is requested. What changes is the distribution of the inter-arrival times. The best result is given by the PureLTE algorithm, since every vehicle is sending its own information at the requested update frequency (i.e., collection interval). The only variability here can come from the different message transferring delays. Although



(a) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.

(b) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.



(c) $I_{\text{LDB}} = 10\,\text{Hz}$ and $P^{\text{tx}} = 100\,\text{mW}$.

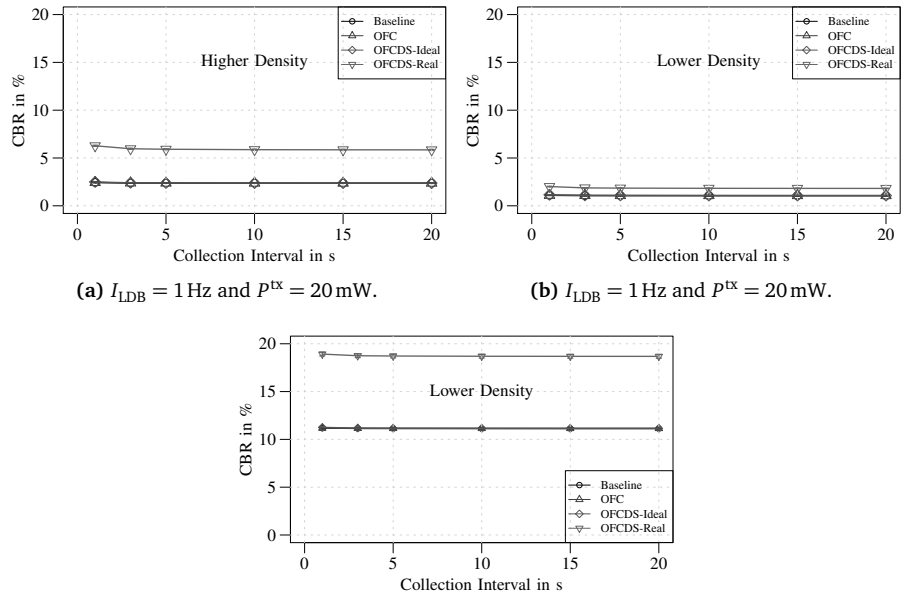
Figure 5.23 – Inter-arrival time for $I_{\text{col}} = 10\,\text{s}$.

each vehicle is sending the same amount of information, the delay depends on the quality of the LTE channel. However, since the single FCD message size is constant and relatively small, this variability is not so visible here.

The story is slightly different for the heterogeneous algorithms. Here, besides the data packets transferring delays, the variability is also caused by the fact that the forwarding vehicles are re-elected at each collection period. Since a single vehicle's FCD message may be sent by different forwarders in different collection periods, we have an additional variability already at the sender side. This is visible when looking at OFC and OFCDS, whose results are similar among them, since the forwarder election mechanism is the same. The largest distribution of the inter-arrival time values is given by Baseline, because this algorithm does not consider the CQI in the LTE uplink, which leads to a higher variability due to the data transferring delays. The DSRC channel load also affects the inter-arrival time metric, as can be seen from Figure 5.23c, suggesting the fact that an increased load on the DSRC channel leads to a higher variability of the update beacons being received by a vehicle.

Evaluation of the Impact on DSRC

In Figure 5.24 we show the impact that each of the considered heterogeneous algorithms has on the DSRC channel in terms of CBR as a function of the collection interval. An interesting observation is that, for all the considered parameter



(a) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.

(b) $I_{\text{LDB}} = 1\,\text{Hz}$ and $P^{\text{tx}} = 20\,\text{mW}$.

(c) $I_{\text{LDB}} = 10\,\text{Hz}$ and $P^{\text{tx}} = 100\,\text{mW}$.

Figure 5.24 – The DSRC Channel Busy Ratio as a function of the collection interval.

configurations, OFCDS-Real gives a higher CBR with respect to all other solutions, independently from the considered collection interval. These results are consistent with the fact that obtaining 2-hop DSRC awareness comes with a cost, especially if we attach to every beacon the raw current ID list of neighboring vehicles, as OFCDS-Real does. Notice that for the lower density scenario and beaconing frequency (Figure 5.24b) the overall CBR is smaller, since we have less vehicles periodically sending their beacons and less often. Same scenario but with 10 times the beaconing frequency (Figure 5.24c) leads to 10 times more load on the DSRC channel. OFCDS-Ideal has the same performance in terms of CBR as Baseline and OFC. This confirms the fact that the higher CBR induced by OFCDS-Real is caused only by the ID lists attached to the beacons. This also suggests that if we can come out with a good compression algorithm for the ID lists, we can do at most as well as OFCDS-Ideal.

It is worth noting that the CBR remains constant when varying the collection interval, with only a very small increase for $I_{\text{col}} = 1\,\text{s}$. This confirms the fact that beaconing is the main cause that affects the mean DSRC channel utilization, while the additional load induced by the *Inhibit* and *Notify* messages is insignificant. However, the mean CBR represented in Figure 5.24 is not uniformly distributed in time and/or space. As expected, we noticed that CBR depends on the current neighborhood vehicular density, meaning that at an individual level, each vehicle experiences a CBR that depends on the current number of DSRC neighbors. In time, we noticed

that vehicles roaming in crowded neighborhoods experience periodic spikes of CBR, which are observed when *Inhibit* and *Notify* messages are broadcasted. On the downside, these control messages are a direct consequence of vehicles being elected as forwarders, meaning that the spikes are caused by our proposed algorithm. On the upside, their impact on the overall system, and specifically on the mean CBR, is insignificant, as can be seen in Figure 5.24. Also, the algorithm itself is something that we can control, meaning that smoothing out these spikes can be the subject of a future work.

5.3 Conclusion

In this chapter we address the problem of intermittent FCD collection in an urban environment by exploiting both DSRC and LTE technologies. We show that the DSRC technology can significantly help in reducing the LTE channel utilization for non-safety applications that require intermittent collection of data. This is generally achieved through clustering mechanisms that select a subset of vehicles in charge of aggregating and sending the information to a remote server via LTE.

In Section 5.1 we presented an FCD collection protocol via LTE cellular network, where substantial off-loading is obtained by resorting to V2V direct communication links to elect representative vehicle nodes that aggregate FCD of their respective neighboring vehicle nodes before sending them through LTE channels. The identification of representative nodes is distributed, based on autonomous rules followed by each participating vehicle node, seamless to the LTE cellular network. V2V communication is assumed to take place by means of the IEEE 802.11p VANET, since this is the technology specifically designed for that purpose and that kind of transponders are expected to become part of the vehicle equipment, at least because of safety requirements. The proposed solution adapts fully to the available penetration rate of the VANET equipment; it falls back automatically to LTE only FCD collection in case VANET equipment is not available or excessively sparse. Our results confirm the significant performance gain, as expressed in terms of the saved number of used LTE PUSCH channels and RBs, that can be of an order of magnitude in case the nodes reported FCD include complete information on VANET connectivity.

In Section 5.2 we propose an on-the-fly distributed clustering algorithm with forwarder selection, named OFC, that considers the DSRC connectivity, the Channel Quality Indicator in the LTE uplink, and a randomly varying timing (i.e., jitter) as the main parameters in the forwarder selection process. Also, we propose an improved version of this algorithm, named OFCDS, which has a more efficient duplicate suppression mechanism. The main features of these proposed solutions are: (i) both OFC and OFCDS rely on a distributed procedure to periodically select

forwarding vehicles in charge of sending their own data, as well as their neighboring vehicles' information, towards a remote facility via LTE; (ii) the forwarder selection process is based on timers that depend on parameters drawn from both DSRC and LTE communication technologies; (iii) they both exploit the existing background beaconing process to populate a local data base used by the forwarding vehicles to create the aggregated information to be sent via LTE; (iv) OFCDS uses additional control messages sent on DSRC to suppress the generated duplicates.

The efficiency of the proposed algorithms is proved by means of an extensive performance analysis based on realistic simulations. In particular, we show that a proper cooperation between the VANET based on the DSRC technology and the LTE cellular network brings a significant benefit in terms of LTE radio resources utilization. The price that we have to pay for off-loading the LTE access network consists in an increase of the data transferring delay. However, this delay might be compensated by a significantly lower number of vehicles simultaneously competing during the random access procedure.

Chapter 6

Final Conclusions and Remarks

Real-time vehicular Floating Car Data (FCD) collection is one of the key enablers of a broad range of applications in the context of ITS. The research community and the automotive industry are looking at several candidate communication technologies that can support periodic and frequent collection of FCD information, such as DSRC, LTE, mmWave, VLC. Each one of the considered technologies has its own advantages and disadvantages, but there is no clear winner.

In this thesis we studied the feasibility of the DSRC technology to endorse FCD collection, both as a self contained solutions, as well as a supporting technology to offload the LTE cellular network. We proved that DSRC is a suitable technology not only for safety, but also for traffic efficiency applications. In fact, it can handle periodic collection of FCD messages and can feed the traffic monitoring systems with real-time information. Moreover, we showed that DSRC is able to significantly decrease the load on the LTE network by reducing the number of vehicles requiring access the to the LTE channels while collecting the same amount of information.

In the first part of this thesis we addressed the problem of DSRC channels congestion, which can cause information loss and increase the communication delay. The main DSRC communication paradigm is based on periodic broadcasts of FCD information, which helps each vehicle to keep track of other neighboring vehicles roaming in their vicinity by maintaining a local data base of FCD messages. Since most of the FCD collection algorithms rely on the accuracy of the information inside these local data bases, keeping them up-to-date is crucial. The primary metric that measures the freshness of the information in every vehicle's local data base is the Age-of-Information (AoI), defined as the mean age of the latest updates received from the neighboring nodes. In Chapter 3 we defined an analytical model to evaluate the AoI of a DSRC-based VANET. The model takes in input the current connectivity graph of the vehicles and gives in output the beaconing frequency that minimizes the AoI metric. We validated our model against realistic simulations of two different

urban scenarios: (i) a generic Manhattan grid scenario, and (ii) the TapasCologne scenario [87], which represents the traffic mobility in the city of Cologne, Germany. We showed that the analytical model captures quite well the variation of the mean AoI with respect to the beacon sending frequency.

We moved further and proposed in Chapter 4 an integrated DSRC-based FCD collection protocol, named DISCOVER, that exploits the beacon exchange mechanism and the GeoNetworking CBF data dissemination protocol [27] to periodically collect FCD information in an urban area. The proposed solution was validated through realistic simulations based on the LuST vehicular traffic scenario [94], [95]. The considered use case scenario comprises an RSU placed at an intersection in the central part of the Luxembourg City and a population of vehicles roaming in a $4\,\mathrm{km}^2$ area around this RSU. We performed extensive simulations and analyzed DISCOVER under different parameter configurations and vehicular densities. The obtained results showed that the proposed solution is quite robust with respect to the vehicular density and it is able to collect approximately 90 % of vehicles in less than 3 s. Moreover, we gave a quantitative and a qualitative analysis of the collected FCD information and indicated a possible direction for the protocol improvement.

Finally, in Chapter 5 we proposed and evaluated different FCD collection schemes that exploit the DSRC technology to offload the LTE network. We first showed in Section 5.1 that by simply decreasing the number of vehicles communicating the FCDs present in their local data bases via LTE can bring a significant reduction of the LTE resource utilization. The selection of this subset of vehicles is driven by the VANET and it is based again on the beaconing mechanism and on the GeoNetworking CBF algorithm described in Section 2.3. The proposed solution dynamically adapts to the DSRC equipment penetration rate. The performance gain in terms of LTE resource utilization with respect to the existing solutions that require an initial setup phase where each vehicle has to send its own FCD via LTE was proved to be significant. In Section 5.2 we took another step and proposed two fully distributed on-the-fly clustering algorithms, named OFC and OFCDS, that reduce even more the LTE resource utilization by considering the LTE CQI parameter in the clustering mechanism. We evaluated the proposed solutions for different FCD collection intervals and vehicular densities, and compared their performance with other state of the art algorithms. We proved that OFC and OFCDS outperform other solutions in terms of RB utilization thanks to the fact that they consider qualitative parameters drawn from both communication technologies in the cluster head selection process.

Of course, periodic and frequent FCD collection still remains a challenge. The increasing number of sensors installed in our cars will generate more and more information, that will be a catalyst for a huge number of new ITS applications. How the current communication technologies will handle the growing amount of information that needs be collected is still an open question. Moreover, modern

vehicles already have multiple communication technologies on-board and others are still to come. How to properly combine them could depend on the application, on the current state of the vehicle, on the type of information to be collected, or it could be the result of a common decision taken in cooperation with other neighboring vehicles.

Another issue to be considered is that of privacy. What is the impact of the periodic and frequent collection of FCD information on the people's privacy? How can we design privacy-preserving FCD collection algorithms so that to prevent tracking and identification? All these questions have to be answered in order to convince the drivers to share their data, so as to be able to build safe, efficient, and sustainable Intelligent Transportation Systems.

List of Abbreviations

3GPP	Third Generation Partnership Project
AF	Advanced Forwarding
AoI	Age-of-Information
ARI	Auto-fahrer Rundfunk Information
ARIB	Association of Radio Industries and Businesses
AS	Access Stratum
BSM	Basic Safety Message
BSS	Basic Service Set
BSSID	Basic Service Set Identifier
BTP	Basic Transport Protocol
CACS	Comprehensive Automobile Control System
CAM	Cooperative Awareness Message
CBF	Contention-Based Forwarding
CCH	Control Channel
CH	Cluster Head
CQI	Channel Quality Indicator
CSMA/CA	Carrier Sense Multiple Access / Collision Avoidance
DENM	Decentralized Environmental Notification Message
DRIVE	Dedicated Road Infrastructure for Vehicle Safety in Europe
DSRC	Dedicated Short-Range Communication
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
EDCA	Enhanced Distributed Channel Access
EPC	Evolved Packet Core
ERGS	Electronic Route Guidance System
ERTICO	European Road Transport Telematics Implementation Coordination Organization
ETSI	European Telecommunications Standards Institute
FCD	Floating Car Data
GDP	Gross Domestic Product
GPS	Global Positioning System

HARQ	Hybrid Automatic Repeat Request
ICT	Information and Communication Technologies
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
IPv6	Internet Protocol version 6
ITS	Intelligent Transportation Systems
IVC	Inter-Vehicle Communication
IVHS	Intelligent Vehicle Highway Systems
LDB	Local Data Base
LDM	Local Dynamic Map
LTE	Long Term Evolution
M2M	Machine-to-Machine
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MIS	Maximum Independent Set
mmWave	Millimeter Wave
MTU	Maximum Transmission Unit
MVR	Monitored Vehicles Ratio
NAS	Non-Access Stratum
NCR	Node Coverage Ratio
NED	Network Description Language
NHTSA	National Highway Traffic Safety Administration
OBU	On-Board Unit
OCB	Outside the Context of a BSS
OFC	On-the-Fly Clustering
OFCDS	OFC with Duplicate Suppression
OFDM	Orthogonal Frequency Division Multiplexing
OMNeT++	Objective Modular Network Testbed in C++
PDCP	Packet Data Convergence Protocol
PDU	Protocol Data Unit
PROMETHEUS	Program for a European Traffic System with Higher Efficiency and Unprecedented Safety
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Service Channel
QoS	Quality of Service
RACS	Road/Automobile Communication System
RB	Resource Block
RLC	Radio Link Control
RNR	Relay Nodes Ratio
ROI	Region of Interest

RRC	Radio Resource Control
RSU	Road Side Unit
SAE	Society of Automotive Engineers
SC-FDMA	Single Carrier Frequency Division Multiple Access
SCH	Service Channel
SDU	Service Data Unit
SINR	Signal to Interference plus Noise Ratio
SNR	Signal to Noise Ratio
SUMO	Simulation of Urban MObility
TCP	Transmission Control Protocol
TRC	Traffic Responsive Capabilities
TTI	Transmission Time Interval
UDP	User Datagram Protocol
UE	User Equipment
UMTS	Universal Mobile Telecommunication System
USDOT	U.S. Department of Transportation
UV-CAST	Urban Vehicular Broadcast
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
VANET	Vehicular Ad Hoc Network
VBN	Vehicular Backbone Network
Veins	Vehicles in Network Simulation
VLC	Visible Light Communication
WAVE	Wireless Access for Vehicular Environments
WSM	WAVE Short Message
WSMP	WAVE Short Message Protocol

List of Figures

1.1 Cars per inhabitant 1
1.2 People killed in road accidents 2
1.3 Greenhouse gas emissions in Europe and USA. 2
1.4 Traffic jam detector in Germany. 3

2.1 Relationship among protocol stack architectures 10
2.2 DSRC spectrum allocation 11
2.3 E-UTRAN protocol stack 13
2.4 CBF algorithm operation 16
2.5 Simulation framework 17

3.1 A sample path of AoI 27
3.2 Definition of the time interval Y 28
3.3 Cologne simulation scenario 34
3.4 AoI vs sending period 36
3.5 CBR vs sending period 37
3.6 Successful message delivery vs number of DSRC neighbors 37
3.7 Net throughput vs number of DSRC neighbors 38
3.8 AoI vs number of DSRC neighbors 39
3.9 AoI heat colormap 40
3.10 AoI vs sending period 40

4.1 Example of forward and reverse waves 45
4.2 Example of a collection instance performed by DISCOVER 49
4.3 The simulated scenario 52
4.4 NCR vs $D_{\max}$ and $H_{\max}$ 54
4.5 RNR vs $D_{\max}$ and $H_{\max}$ 55
4.6 Dissemination delay vs $D_{\max}$ and $H_{\max}$ 55
4.7 MVR vs $D_{\max}$ and $T_{\text{rep}}^{\max}$ 56
4.8 MVR vs MAX_BACKUP_RTX and $T_{\text{rep}}^{\max}$ 57

4.9 Collected bytes vs $D_{\max}$ and $T_{\text{rep}}^{\max}$ 58
4.10 Collected bytes vs MAX_BACKUP_RTX and $T_{\text{rep}}^{\max}$ 59
4.11 CBR vs $D_{\max}$ and $T_{\text{rep}}^{\max}$ 59
4.12 CBR vs MAX_BACKUP_RTX and $T_{\text{rep}}^{\max}$ 60
4.13 Collection delay vs $D_{\max}$ and $T_{\text{rep}}^{\max}$ 61
4.14 Collection delay vs MAX_BACKUP_RTX and $T_{\text{rep}}^{\max}$ 62
4.15 An illustrative example of a collection cycle using DISCOVER 62

5.1 Existing hybrid collection schemes 67
5.2 Summary of the existing collection schemes 69
5.3 Considered urban scenario maps 73
5.4 Monitored urban area (Manhattan, NY) 74
5.5 New York: M_{CH} and M_{RB} vs R_{eNodeB} (SETUP) 79
5.6 Rome: M_{CH} and M_{RB} vs R_{eNodeB} (SETUP) 80
5.7 M_{CH} vs R_{eNodeB} and vehicle node density 81
5.8 M_{RB} vs. R_{eNodeB} for two density levels 82
5.9 New York: M_{CH} and M_{RB} vs R_{eNodeB} (COLLECTION) 84
5.10 Rome: M_{CH} and M_{RB} vs R_{eNodeB} (COLLECTION) 85
5.11 New York: M_{CH} and M_{RB} vs R_{eNodeB} (multiple originators) 86
5.12 Rome: M_{CH} and M_{RB} vs R_{eNodeB} (multiple originators) 87
5.13 FCD collection scenario 88
5.14 PureLTE data collection algorithm 89
5.15 OFC data collection algorithm 91
5.16 OFCDS data collection algorithm 95
5.17 Time diagram of the OFC and OFCDS algorithms 96
5.18 Part of the simulation scenario 97
5.19 RB utilization as a function of α, β, and γ 100
5.20 RB utilization vs I_{col} 101
5.21 Duplicates ratio vs I_{col} 102
5.22 Delay vs I_{col} 104
5.23 Inter-arrival time for $I_{\text{col}} = 10\,\text{s}$ 105
5.24 CBR vs I_{col} 106

List of Tables

2.1 IEEE 802.11 OFDM basic PHY parameters for a 10 MHz channel . . . 11
2.2 Main LTE parameters. 14

3.1 Simulation parameters 35

4.1 Main simulation parameters 53

5.1 Notations and simulation parameter values 74
5.2 Performance metrics for the dissemination of the REQUEST message 77
5.3 Network and simulation parameters. 98
5.4 Performance evaluation metrics 99

Bibliography

[1] EUROSTAT. (2016). Dataset [Database], Product code: road_eqs_carhab, [Online]. Available: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=road_eqs_carhab&lang=en (visited on 09/06/2017).

[2] U.S. Department of Transportation, Federal Highway Administration. (2016). Highway Statistics 2014, Chart DV-1C, [Online]. Available: https://www.fhwa.dot.gov/policyinformation/statistics/2014/ (visited on 09/06/2017).

[3] U.S. Department of Transportation, Bureau of Transportation Statistics. (2016). Transportation Satellite Accounts, [Online]. Available: https://www.bts.gov/archive/publications/transportation_satellite_accounts/index (visited on 09/06/2017).

[4] EUROSTAT. (2017). Dataset [Database], Product code: tsdtr420, [Online]. Available: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tsdtr420 (visited on 09/06/2017).

[5] U.S. Department of Transportation, National Highway Traffic Safety Administration. (2017). Fatality Analysis Reporting System, [Online]. Available: https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars (visited on 09/06/2017).

[6] EUROSTAT. (2017). Dataset [Database], Product code: env_air_gge, [Online]. Available: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_air_gge&lang=en (visited on 09/07/2017).

[7] U.S. Environmental Protection Agency. (2016). Greenhouse Gas Inventory Data Explorer, [Online]. Available: https://www3.epa.gov/climatechange/ghgemissions/inventoryexplorer/ (visited on 09/07/2017).

[8] INRIX. (2016). Global Traffic Scorecard, [Online]. Available: http://inrix.com/scorecard/ (visited on 09/08/2017).

[9] O. Altintas, F. Dressler, F. Hagenauer, M. Matsumoto, M. Sepulcre, and C. Sommer, "Making Cars a Main ICT Resource in Smart Cities," in *34th IEEE Conference on Computer Communications (INFOCOM 2015), International Workshop on Smart Cities and Urban Informatics (SmartCity 2015)*, Hong Kong, China: IEEE, Apr. 2015, pp. 654–659. DOI: 10.1109/INFCOMW.2015.7179448.

[10] A. Auer, S. Feese, and S. Lockwood, "History of Intelligent Transportation Systems," U.S. Department of Transportation, Intelligent Transportation Systems, Joint Program Office, Tech. Rep. FHWA-JPO-16-329, 2016.

[11] G Nowacki, C Krysiuk, and R Kopczewski, "Development and Standardization of Intelligent Transport Systems," 2012.

[12] W. G. Najm, J. Koopmann, J. D. Smith, and J. Brewer, "Frequency of Target Crashes for IntelliDrive Safety Systems," U.S. Department of Transportation, National Highway Traffic Safety Administration, Tech. Rep. DOT HS 811 381, 2010.

[13] J. Chang, G. Hatcher, D. Hicks, J. Schneeberger, B. Staples, S. Sundarajan, M. Vasudevan, P. Wang, and K. Wunderlich, "Estimated Benefits of Connected Vehicle Applications: Dynamic Mobility Applications, AERIS, V2I Safety, and Road Weather Management," U.S. Department of Transportation, Intelligent Transportation Systems, Joint Program Office, Tech. Rep. FHWA-JPO-15-255, 2015.

[14] J. Barbaresso, G. Cordahi, D. Garcia, C. Hill, A. Jendzejec, and K. Wright, "USDOT's Intelligent Transportation Systems (ITS) ITS Strategic Plan 2015-2019," U.S. Department of Transportation, Intelligent Transportation Systems, Joint Program Office, Tech. Rep. FHWA-JPO-14-145, 2014.

[15] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021," White Paper, 2017.

[16] M. S. Ali, E. Hossain, and D. I. Kim, "LTE/LTE-A Random Access for Massive Machine-Type Communications in Smart Cities," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 76–83, Jan. 2017. DOI: 10.1109/MCOM.2017.1600215CM.

[17] E. Soltanmohammadi, K. Ghavami, and M. Naraghi-Pour, "A Survey of Traffic Issues in Machine-to-Machine Communications Over LTE," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 865–884, Dec. 2016. DOI: 10.1109/JIOT.2016.2533541.

[18] G. C. Madueño, J. J. Nielsen, D. M. Kim, N. K. Pratas, Č. Stefanović, and P. Popovski, "Assessment of LTE Wireless Access for Monitoring of Energy Distribution in the Smart Grid," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 675–688, Mar. 2016. DOI: 10.1109/JSAC.2016.2525639.

[19] T. P. C. de Andrade, C. A. Astudillo, L. R. Sekijima, and N. L. S. da Fonseca, "The Random Access Procedure in Long Term Evolution Networks for the Internet of Things," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 124–131, Mar. 2017. DOI: 10.1109/MCOM.2017.1600555CM.

[20] E. Dahlman, S. Parkvall, and J. Sköld, 4G LTE / LTE-Advanced for Mobile Broadband. Academic Press, 2014.

[21] IEEE, "IEEE Guide for Wireless Access in Vehicular Environments (WAVE) - Architecture," IEEE, Std 1609.0-2013, Mar. 2014. DOI: 10.1109/IEEESTD.2014.6755433.

[22] J. B. Kenney, "Dedicated Short-Range Communications (DSRC) Standards in the United States," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1162–1182, Jul. 2011.

[23] A. Benslimane, T. Taleb, and R. Sivaraj, "Dynamic Clustering-Based Adaptive Mobile Gateway Management in Integrated VANET - 3G Heterogeneous Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 559–570, Mar. 2011. DOI: 10.1109/JSAC.2011.110306.

[24] R. Stanica, M. Fiore, and F. Malandrino, "Offloading Floating Car Data," in *14th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2013)*, Madrid, Spain: IEEE, Jun. 2013. DOI: 10.1109/WoWMoM.2013.6583391.

[25] S. Ancona, R. Stanica, and M. Fiore, "Performance boundaries of massive Floating Car Data offloading," in *11th Annual Conference on Wireless On-demand Network Systems and Services (WONS 2014)*, Obergurgl, Austria: IEEE, Apr. 2014, pp. 89–96. DOI: 10.1109/WONS.2014.6814727.

[26] I. Turcanu, C. Sommer, A. Baiocchi, and F. Dressler, "Pick the Right Guy: CQI-Based LTE Forwarder Selection in VANETs," in *8th IEEE Vehicular Networking Conference (VNC 2016)*, Columbus, OH: IEEE, Dec. 2016, pp. 98–105. DOI: 10.1109/VNC.2016.7835938.

[27] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 4: Geographical addressing and forwarding for point-to-point and point-to-multipoint communications; Sub-part 1: Media-Independent Functionality," ETSI, EN 302 636-4-1 v1.2.1, May 2014.

[28] C. Sommer and F. Dressler, Vehicular Networking. Cambridge University Press, Nov. 2014. DOI: 10.1017/CBO9781107110649.

[29] J. Eriksson, H. Balakrishnan, and S. Madden, "Cabernet: Vehicular Content Delivery Using WiFi," in *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking (MobiCom '08)*, San Francisco, California, USA: ACM, 2008, pp. 199–210. DOI: 10.1145/1409944.1409968.

[30] D. Pesavento, G. Grassi, G. Pau, P. Bahl, and S. Fdida, "Demo: Car-Fi: Opportunistic V2I by Exploiting Dual-Access Wi-Fi Networks," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*, Paris, France: ACM, 2015, pp. 173–175. DOI: 10.1145/2789168.2789171.

[31] M. Giordani, A. Zanella, and M. Zorzi, "Millimeter wave communication in vehicular networks: Challenges and opportunities," in *6th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, 2017, pp. 1–6. DOI: 10.1109/MOCAST.2017.7937682.

[32] M. Segata, R. Lo Cigno, H.-M. Tsai, and F. Dressler, "On Platooning Control using IEEE 802.11p in Conjunction with Visible Light Communications," in *12th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2016)*, Cortina d'Ampezzo, Italy: IEEE, 2016, pp. 124–127.

[33] A. Memedi, H.-M. Tsai, and F. Dressler, "Impact of Realistic Light Radiation Pattern on Vehicular Visible Light Communication," in *IEEE Global Telecommunications Conference (GLOBECOM 2017)*, to appear, Singapore: IEEE, 2017.

[34] ETSI, "Intelligent Transport Systems (ITS); Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band," ETSI, EN 302 663 V1.2.1, Jul. 2013.

[35] ——, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Local Dynamic Map (LDM)," ETSI, EN 302 895 V1.1.1, Sep. 2014.

[36] ——, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service," ETSI, Tech. Rep. 302 637-2 V1.3.2, Nov. 2014.

[37] ——, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 3: Specification of Decentralized Environmental Notification Basic Service," ETSI, Tech. Rep. 302 637-3 V1.2.1, Sep. 2014.

[38] SAE Int., "Dedicated Short Range Communications (DSRC) Message Set Dictionary," SAE, Tech. Rep. J2735 201603, Mar. 2016.

[39] IEEE, "IEEE Standard for Wireless Access in Vehicular Environments–Security Services for Applications and Management Messages," *IEEE Std 1609.2-2016 (Revision of IEEE Std 1609.2-2013)*, pp. 1–240, Mar. 2016. DOI: 10.1109/IEEESTD.2016.7426684.

[40] ——, "IEEE Standard for Wireless Access in Vehicular Environments (WAVE) – Networking Services," *IEEE Std 1609.3-2016 (Revision of IEEE Std 1609.3-2010)*, pp. 1–160, Apr. 2016. DOI: 10.1109/IEEESTD.2016.7458115.

[41] ——, "IEEE Standard for Wireless Access in Vehicular Environments (WAVE) – Multi-Channel Operation," *IEEE Std 1609.4-2016 (Revision of IEEE Std 1609.4-2010)*, pp. 1–94, Mar. 2016. DOI: 10.1109/IEEESTD.2016.7435228.

[42] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 5: Transport Protocols; Sub-part 1: Basic Transport Protocol," ETSI, EN 302 636-5-1 v1.2.0, Oct. 2013.

[43] ——, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN);Overall description;Stage 2," ETSI, Tech. Rep. 136 300 V8.10.0, Sep. 2009.

[44] A. Virdis, G. Stea, and G. Nardini, "SimuLTE - A Modular System-level Simulator for LTE/LTE-A Networks based on OMNeT++," in *4th International Conference on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH 2014)*, Vienna, Aug. 2014.

[45] I. T. Union, "Guidelines for evaluation of radio interface technologies for IMT-Advanced," ITU-R, Report M.2135-1, Dec. 2009.

[46] M. Chaqfeh, A. Lakas, and I. Jawhar, "A survey on data dissemination in vehicular ad hoc networks," *Elsevier Vehicular Communications*, vol. 1, no. 4, pp. 214 –225, 2014. DOI: 10.1016/j.vehcom.2014.09.001.

[47] I. Rubin, A. Baiocchi, F. Cuomo, and P. Salvo, "GPS aided inter-vehicular wireless networking," in *Information Theory and Applications Workshop (ITA)*, Feb. 2013, pp. 1–9. DOI: 10.1109/ITA.2013.6502973.

[48] W. Viriyasitavat, O. K. Tonguz, and F. Bai, "UV-CAST: an urban vehicular broadcast protocol," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 116–124, Nov. 2011. DOI: 10.1109/MCOM.2011.6069718.

[49] O. K. Tonguz, N. Wisitpongphan, J. S. Parikh, F. Bai, P. Mudalige, and V. K. Sadekar, "On the Broadcast Storm Problem in Ad hoc Wireless Networks," in *3rd International Conference on Broadband Communications, Networks and Systems*, Oct. 2006, pp. 1–11. DOI: 10.1109/BROADNETS.2006.4374403.

[50] S. Kuhlmorgen, I. Llatser, A. Festag, and G. Fettweis, "Performance Evaluation of ETSI GeoNetworking for Vehicular Ad Hoc Networks," in *81st IEEE Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–6. DOI: 10.1109/VTCSpring.2015.7146003.

[51] J. Harri, F. Filali, and C. Bonnet, "Mobility models for vehicular ad hoc networks: a survey and taxonomy," *IEEE Communications Surveys Tutorials*, vol. 11, no. 4, pp. 19–41, Dec. 2009. DOI: 10.1109/SURV.2009.090403.

[52] F. J. Martinez, C. K. Toh, J.-C. Cano, C. T. Calafate, and P. Manzoni, "A Survey and Comparative Study of Simulators for Vehicular Ad Hoc Networks (VANETs)," *Wirel. Commun. Mob. Comput.*, vol. 11, no. 7, pp. 813–828, Jul. 2011. DOI: 10.1002/wcm.859.

[53] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *1st ACM/ICST International Conference on Simulation Tools and Techniques for Communications, Networks and Systems (SIMUTools 2008)*, Marseille, France: ACM, Mar. 2008.

[54] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent Development and Applications of SUMO – Simulation of Urban MObility," *International Journal On Advances in Systems and Measurements*, vol. 5, no. 3&4, pp. 128–138, 2012.

[55] C. Sommer, R. German, and F. Dressler, "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, Jan. 2011. DOI: 10.1109/TMC.2010.133.

[56] S. Krauss, "Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamic," PhD thesis, 1998.

[57] F. Hagenauer, F. Dressler, and C. Sommer, "A Simulator for Heterogeneous Vehicular Networks," in *6th IEEE Vehicular Networking Conference (VNC 2014), Poster Session*, Paderborn, Germany: IEEE, Dec. 2014, pp. 185–186. DOI: 10.1109/VNC.2014.7013339.

[58] B. Brik, N. Lagraa, H. Cherroun, and A. Lakas, "Token-based Clustered Data Gathering Protocol(TCDGP) in vehicular networks," in *9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Jul. 2013, pp. 1070–1074. DOI: 10.1109/IWCMC.2013.6583705.

[59] W. R. Chang, H. T. Lin, and B. X. Chen, "TrafficGather: An Efficient and Scalable Data Collection Protocol for Vehicular Ad Hoc Networks," in *5th IEEE Consumer Communications and Networking Conference*, Jan. 2008, pp. 365–369. DOI: 10.1109/ccnc08.2007.88.

[60] A. Soua and H. Afifi, "Adaptive data collection protocol using reinforcement learning for VANETs," in *9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Jul. 2013, pp. 1040–1045. DOI: 10.1109/IWCMC.2013.6583700.

[61] Y. Zhu, Q. Zhao, and Q. Zhang, "Delay-Constrained Data Aggregation in VANETs," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 2097–2107, May 2015. DOI: 10.1109/TVT.2014.2335232.

[62] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, "LTE for Vehicular Networking: A Survey," *IEEE Communications Magazine*, vol. 51, no. 5, pp. 148–157, May 2013. DOI: 10.1109/MCOM.2013.6515060.

[63] T. Mangel, T. Kosch, and H. Hartenstein, "A comparison of UMTS and LTE for Vehicular Safety Communication at Intersections," in *2nd IEEE Vehicular Networking Conference (VNC 2010)*, Jersey City, NJ: IEEE, Dec. 2010, pp. 293–300. DOI: 10.1109/VNC.2010.5698244.

[64] C. Ide, B. Dusza, M. Putzke, and C. Wietfeld, "Channel sensitive transmission scheme for V2I-based Floating Car Data collection via LTE," in *IEEE International Conference on Communications (ICC 2012)*, Ottawa, Canada: IEEE, Jun. 2012, pp. 7151–7156. DOI: 10.1109/ICC.2012.6364684.

[65] R. N. Clarke, "Expanding mobile wireless capacity: The challenges presented by technology and economics," *Telecommunications Policy*, vol. 38, no. 8–9, pp. 693–708, Sep. 2014. DOI: 10.1016/j.telpol.2013.11.006.

[66] C. Ide, F. Kurtz, and C. Wietfeld, "Cluster-Based Vehicular Data Collection for Efficient LTE Machine-Type Communication," in *78th IEEE Vehicular Technology Conference Fall (VTC2013-Fall)*, Las Vegas, NV: IEEE, Sep. 2013. DOI: 10.1109/VTCFall.2013.6692136.

[67] G. El Mouna Zhioua, N. Tabbane, H. Labiod, and S. Tabbane, "A Fuzzy Multi-Metric QoS-Balancing Gateway Selection Algorithm in a Clustered VANET to LTE Advanced Hybrid Cellular Network," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 2, pp. 804–817, Jun. 2015. DOI: 10.1109/TVT.2014.2323693.

[68] G. Remy, S.-M. Senouci, F. Jan, and Y. Gourhant, "LTE4V2X–Collection, dissemination and multi-hop forwarding," in *IEEE International Conference on Communications (ICC), 2012*, Nanjing, Jiangsu: IEEE, Jun. 2012, pp. 120–125. DOI: 10.1109/ICC.2012.6364412.

[69] A. Bazzi, B. M. Masini, and G. Pasolini, "V2V and V2R for cellular resources saving in vehicular applications," in *IEEE Wireless Communications and Networking Conference (WCNC 2012)*, Paris, France: IEEE, Apr. 2012, pp. 3199–3203. DOI: 10.1109/WCNC.2012.6214358.

[70] R. S. Bali, N. Kumar, and J. J. Rodrigues, "Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions," *Elsevier Vehicular Communications*, vol. 1, no. 3, pp. 134–152, Jul. 2014. DOI: 10.1016/j.vehcom.2014.05.004.

[71] S. Jia, S. Hao, X. Gu, and L. Zhang, "Analyzing and relieving the impact of FCD traffic in LTE-VANET heterogeneous network," in *21st International Conference on Telecommunications (ICT)*, May 2014, pp. 88–92. DOI: 10.1109/ICT.2014.6845086.

[72] P. M. D'Orey, N. Maslekar, I. de la Iglesia, and N. K. Zahariev, "NAVI: Neighbor-Aware Virtual Infrastructure for Information Collection and Dissemination in Vehicular Networks," in *81st IEEE Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–6. DOI: 10.1109/VTCSpring.2015.7145945.

[73] A. Baiocchi and I. Turcanu, "A Model for the Optimization of Beacon Message Age-of-Information in a VANET," in *29th International Teletraffic Congress (ITC 29)*, vol. 1, 2017, pp. 108–116. DOI: 10.23919/ITC.2017.8064345.

[74] N. M. Freris, H. Kowshik, and P. R. Kumar, "Fundamentals of Large Sensor Networks: Connectivity, Capacity, Clocks, and Computation," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1828–1846, Nov. 2010. DOI: 10.1109/JPROC.2010.2065790.

[75] H. Zhang, Z. Zhang, and H. Dai, "Gossip-Based Information Spreading in Mobile Networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5918–5928, Nov. 2013. DOI: 10.1109/TWC.2013.100113.130619.

[76] J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu, "Deterministic Gossiping," *Proceedings of the IEEE*, vol. 99, no. 9, pp. 1505–1524, Sep. 2011. DOI: 10.1109/JPROC.2011.2159689.

[77] M. Saeednia and M. Menendez, "A Consensus-Based Algorithm for Truck Platooning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 2, pp. 404–415, Feb. 2017. DOI: 10.1109/TITS.2016.2579260.

[78] J. He, P. Cheng, L. Shi, J. Chen, and Y. Sun, "Time Synchronization in WSNs: A Maximum-Value-Based Consensus Approach," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 660–675, Mar. 2014. DOI: 10.1109/TAC.2013.2286893.

[79] N. Kumar, S. Misra, M. S. Obaidat, J. J.P. C. Rodrigues, and B. Pati, "Networks of learning automata for the vehicular environment: a performance analysis study," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 41–47, Dec. 2014. DOI: 10.1109/MWC.2014.7000970.

[80] G. Scutari, S. Barbarossa, and L. Pescosolido, "Distributed Decision Through Self-Synchronizing Sensor Networks in the Presence of Propagation Delays and Asymmetric Channels," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1667–1684, Apr. 2008. DOI: 10.1109/TSP.2007.909377.

[81] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" In *Proceedings IEEE INFOCOM 2012*, Mar. 2012, pp. 2731–2735. DOI: 10.1109/INFCOM.2012.6195689.

[82] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, Jun. 2011, pp. 350–358. DOI: 10.1109/SAHCN.2011.5984917.

[83] A. Franco, E. Fitzgerald, B. Landfeldt, N. Pappas, and V. Angelakis, "LUPMAC: A cross-layer MAC technique to improve the age of information over dense WLANs," in *23rd International Conference on Telecommunications (ICT)*, May 2016, pp. 1–6. DOI: 10.1109/ICT.2016.7500469.

[84] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, Mar. 2000. DOI: 10.1109/49.840210.

[85] S. C. Liew, C. H. Kai, H. C. Leung, and P. Wong, "Back-of-the-Envelope Computation of Throughput Distributions in CSMA Wireless Networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 9, pp. 1319–1331, Sep. 2010. DOI: 10.1109/TMC.2010.89.

[86] R. Laufer and L. Kleinrock, "The Capacity of Wireless CSMA/CA Networks," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1518–1532, Jun. 2016. DOI: 10.1109/TNET.2015.2415465.

[87] S. Uppoor, O. Trullols-Cruces, M. Fiore, and J. M. Barcelo-Ordinas, "Generation and Analysis of a Large-Scale Urban Vehicular Mobility Dataset," *IEEE Transactions on Mobile Computing*, vol. 13, no. 5, pp. 1061–1075, 2014. DOI: 10.1109/TMC.2013.27.

[88] C. Sommer, D. Eckhoff, R. German, and F. Dressler, "A Computationally Inexpensive Empirical Model of IEEE 802.11p Radio Shadowing in Urban Environments," in *8th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2011)*, Bardonecchia, Italy: IEEE, 2011, pp. 84–90. DOI: 10.1109/WONS.2011.5720204.

[89] C. Sommer, O. K. Tonguz, and F. Dressler, "Adaptive Beaconing for Delay-Sensitive and Congestion-Aware Traffic Information Systems," in *2nd IEEE Vehicular Networking Conference (VNC 2010)*, Jersey City, NJ: IEEE, Dec. 2010, pp. 1–8. DOI: 10.1109/VNC.2010.5698242.

[90] C. Sommer, S. Joerer, M. Segata, O. K. Tonguz, R. Lo Cigno, and F. Dressler, "How Shadowing Hurts Vehicular Communications and How Dynamic Beaconing Can Help," *IEEE Transactions on Mobile Computing*, vol. 14, no. 7, pp. 1411–1421, 2015. DOI: 10.1109/TMC.2014.2362752.

[91] H. T. Cheng, H. Shan, and W. Zhuang, "Infotainment and road safety service support in vehicular networking: From a communication perspective," *Mechanical Systems and Signal Processing*, vol. 25, no. 6, pp. 2020–2038, 2011, Interdisciplinary Aspects of Vehicle Dynamics. DOI: 10.1016/j.ymssp.2010.11.009.

[92] I. Turcanu, P. Salvo, A. Baiocchi, and F. Cuomo, "DISCOVER: A Unified Protocol for Data Dissemination and Collection in VANETs," in *Proceedings of the 12th ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN '15)*, Cancun, Mexico, 2015, pp. 25–32.

[93] ——, "An integrated VANET-based data dissemination and collection protocol for complex urban scenarios," *Elsevier Ad Hoc Networks*, vol. 52, pp. 28–38, 2016, Modeling and Performance Evaluation of Wireless Ad Hoc Networks. DOI: 10.1016/j.adhoc.2016.07.008.

[94] L. Codeca, R. Frank, and T. Engel, "Luxembourg SUMO Traffic (LuST) Scenario: 24 Hours of Mobility for Vehicular Networking Research," in *7th IEEE Vehicular Networking Conference (VNC 2015)*, Kyoto, Japan: IEEE, 2015. DOI: 10.1109/VNC.2015.7385539.

[95] L. Codeca, R. Frank, S. Faye, and T. Engel, "Luxembourg SUMO Traffic (LuST) Scenario: Traffic Demand Evaluation," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 2, pp. 52–63, 2017. DOI: 10.1109/MITS.2017.2666585.

[96] C. Sommer and F. Dressler, "Using the Right Two-Ray Model? A Measurement based Evaluation of PHY Models in VANETs," in *17th ACM International Conference on Mobile Computing and Networking (MobiCom 2011), Poster Session*, Las Vegas, NV: ACM, 2011.

[97] C. Sommer, S. Joerer, and F. Dressler, "On the Applicability of Two-Ray Path Loss Models for Vehicular Network Simulation," in *4th IEEE Vehicular Networking Conference (VNC 2012)*, Seoul, Korea: IEEE, 2012, pp. 64–69. DOI: 10.1109/VNC.2012.6407446.

[98] J. Barrachina, P. Garrido, M. Fogue, F. J. Martinez, J.-C. Cano, C. T. Calafate, and P. Manzoni, "Road Side Unit Deployment: A Density-Based Approach," *IEEE Intelligent Transportation Systems Magazine*, vol. 5, no. 3, pp. 30–39, 2013. DOI: 10.1109/mits.2013.2253159.

[99] P. Salvo, I. Turcanu, F. Cuomo, A. Baiocchi, and I. Rubin, "LTE Floating Car Data application off-loading via VANET driven clustering formation," in *12th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2016)*, Cortina d'Ampezzo, Italy: IEEE, 2016, pp. 192–199.

[100] ——, "Heterogeneous cellular and DSRC networking for Floating Car Data collection in urban areas," *Vehicular Communications*, vol. 8, pp. 21–34, 2017, Internet of Vehicles. DOI: 10.1016/j.vehcom.2016.11.004.

[101] T. S. Rappaport, Wireless Communications: Principles and Practice, 2nd ed. Prentice Hall, 2009.

[102] E. Damosso, L. Correia, and E. Commission, COST Action 231: Digital Mobile Radio Towards Future Generation Systems : Final Report, ser. EUR (Series). European Commission, 1999.

[103] J. Calabuig, J. F. Monserrat, D. Gozalvez, and O. Klemp, "Safety on the Roads: LTE Alternatives for Sending ITS Messages," *IEEE Vehicular Technology Magazine*, vol. 9, no. 4, pp. 61–70, Dec. 2014. DOI: 10.1109/MVT.2014.2362272.

[104] Z. Hameed Mir and F. Filali, "LTE and IEEE 802.11p for vehicular networking: a performance evaluation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, p. 89, 2014. DOI: 10.1186/1687-1499-2014-89.

[105] ETSI, "LTE E-UTRA Physical layer procedures (3GPP TS 36.213)," ETSI, TS 136 213 V13.2.0, Aug. 2016.

[106] M. Haenggi, Stochastic Geometry for Wireless Networks. Cambridge University Press, 2012.

[107] T. Clausen, C. Dearlove, and B. Adamson, Jitter Considerations in Mobile Ad Hoc Networks (MANETs), RFC 5148, Feb. 2008.

[108] S. Uppoor and M. Fiore, "Characterizing Pervasive Vehicular Access to the Cellular RAN Infrastructure: An Urban Case Study," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2603–2614, Jun. 2015. DOI: 10.1109/TVT.2014.2343651.

[109] ETSI, "Intelligent Transport Systems (ITS); Decentralized Congestion Control Mechanisms for Intelligent Transport Systems operating in the 5 GHz range; Access layer part," ETSI, TS 102 687 V1.1.1, Jul. 2011.

[110] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Elsevier Physics Reports*, vol. 424, no. 4–5, pp. 175–308, Feb. 2006. DOI: 10.1016/j.physrep.2005.10.009.