UNIVERSITY OF TUNIS EL MANAR

UNIVERSITY OF PAU AND ADOUR COUNTRIES

DOCTORAL THESIS

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of
COMPUTER SCIENCE

Defended by **SABRI ALLANI**

Data Dissemination and Aggregation in Vehicular Adhoc Network

DEFENDED ON 02/11/2018 IN FRONT OF JURY MEMBERS:

HABIB OUNELLI	PROF, UNIVERSITY OF TUNIS EL MANAR	PRESIDENT
WALID BARHOUMI	PROF, UNIVERSITY OF TUNIS CARTHAGE	REVIEWER
DJAMAL BENSLIMANE	PROF, UNIVERSITY CLAUDE BERNARD LYON 1	REVIEWER
NARJES BELLAMINE	PROF, UNIVERSITY OF MANOUBA	EXAMINER
SADOK BEN YAHIA	PROF, UNIVERSITY OF TUNIS EL MANAR	DIRECTOR
RICHARD CHBEIR	PROF, UNIVERSITY OF PAU AND ADOUR COUNTRIES	DIRECTOR

2017-2018

LIUPPA LAB, University of Pau and Adour Countrie ,64600 Anglet, France
LIPAH-LR 11ES14, University of Tunis El Manar, 2092 Tunis

I dedicate this work: To my parents for education, great love, for their patience and sacrifice.
To my sister. To myself and everyone I love.

Acknowledgements

This research work was carried out thanks to the support and encouragement of several people who have each contributed to his way of completing this work. I want to send them a very large thank you.

I want to thank very sincerely my research director Professor Sadok BEN YAHIA for the trust he gave me. His supervision, his advice and his experience allowed me to acquire the essential knowledge for the development and the result of this work.

Next, I want also to thank very sincerely my research director Professor Richard CHBIER for the trust he gave me. His supervision, his advice and his experience allowed me to acquire the essential knowledge for the development and the result of this work.

I also thank my co-supervisor, Dr. Taoufik YEFERNY for his support and guidance. He often encouraged me during the realization of this work, and genuinely makes use of his experience and knowledge. The work with him was of great value for the realization of this research work.

Abstract

Since the last decade, the emergence of affordable wireless devices in vehicle ad-hoc networks has been a key step towards improving road safety as well as transport efficiency. Informing vehicles about interesting safety and non-safety events is of key interest. Thus, the design of an efficient data dissemination protocol has been of paramount importance. A careful scrutiny of the pioneering vehicle-to-vehicle data dissemination approaches highlights that geocasting is the most feasible approach for VANET applications, more especially in safety applications, since safety events are of interest mainly to vehicles located within a specific area, commonly called ZOR or Zone Of Relevance, close to the event. Indeed, the most challenging issue in geocast protocols is the definition of the ZOR for a given event dissemination. In this thesis, our first contribution introduces a new geocast approach, called Data Dissemination Protocol based on Map Splitting(DPMS). The main thrust of DPMS consists of building the zones of relevance through the mining of correlations between vehicles' trajectories and crossed regions. To do so, we rely on the Formal Concept Analysis (FCA), which is a method of extracting interesting clusters from relational data. The performed experiments show that DPMS outperforms its competitors in terms of effectiveness and efficiency. In another hand, some VANET applications, e.g., Traffic Information System (TIS), require data aggregation in order to inform vehicles about road traffic conditions, which leads to reduce traffic jams and consequently CO_2 emission while increasing the user comfort. Therefore, the design of an efficient aggregation protocol that combines correlated traffic information like location, speed and direction known as Floating Car Data (FCD) is a challenging issue. In this thesis, we introduce a new TIS data aggregation protocol called Smart Directional Data Aggregation (SDDA) able to decrease the network overload while obtaining high accurate information on traffic conditions for large road sections. To this end, we introduce three levels of messages filtering: (i) filtering all FCD messages before the aggregation process based on vehicle directions and road speed limitations, (ii) integrating a suppression technique in the phase of information gathering in order to eliminate the duplicate data, and (iii) aggregating the filtered FCD data and then disseminating it to other vehicles. The performed experiments show that the SDDA outperforms existing approaches in terms of effectiveness and efficiency.

Résumé

Dans cette thèse nous traitons la problématique de dissémination et d'agrégation des données dans le contexte d'un réseaux véhiculaires VANET (Vehicular Ad-Hoc Networks). D'une côté la dissémination de données permet d'informer les véhicules mobiles des principaux événements en un temps acceptable , et de l'autre côté l'agrégation permet de résumer plusieurs données émanant de sources différentes concernant le même événement. Le challenge de la dissémination consiste à calculer la zone de relevance d'un événement, de délivrer les messages aux véhicules de cette zone, et de continuer à délivrer les messages en continu aux véhicules de cette zone. Le challenge de l'agrégation consiste essentiellement à sélectionner les messages à agréger et à qualifier les messages provenant de véhicules lointains. Pour résoudre le problème de dissémination nous proposons un nouveau protocole de dissémination des données dans les réseaux VANET. La principale idée de ce protocole est basée sur la définition de zones de relevance ZOR (zone of relevance of a région) pour la mesure de l'intérêt d'une zone par rapport à un évènement donné, et la définition de split Map permettant de décomposer une grande région en un ensemble de ZORs. L'approche de calcul des ZORs est formalisée, elle est basée sur les techniques de greedy pour l'extraction de la couverture pertinente. D'autre part, certaines applications VANET, par exemple le système d'information de trafic (TIS), nécessitent une agrégation de données pour informer les véhicules des conditions de circulation, ce qui réduit les embouteillages et par conséquent les émissions de CO_2. Par conséquent, la conception d'un protocole d'agrégation efficace combinant des informations de trafic corrélées telles que l'emplacement, la vitesse et la direction, appelées données flottantes sur les voitures (FCD), pose un problème complexe. Dans cette thèse, nous introduisons un nouveau protocole d'agrégation de données dans un réseau VANET appelé SDDA (Smart Directional Data Aggregation). Ce protocole est dédié aussi bien à l'échange de données dans un contexte urbain et autoroutier. Le protocole proposé est basé sur une sélection des messages à agréger. Trois principaux filtres ont été utilisés : filtrage basé sur la direction des véhicules, filtrage basé sur la limitation de vitesse, et filtrage basé sur l'élimination des messages dupliqués. Trois algorithmes d'agrégation sont proposés, ils visent à optimiser l'algorithme de SOTIS. Les trois algorithmes traitent des cas de routes unidirectionnelles, bidirectionnelles et les réseaux urbains.

Abbreviation List

DOT : Department of Transportation

DPMS : swift data Dissemination Protocol based on Map Splitting

DSRC : Dedicated Short Range Communication

DTGS : Dynamique Time Stable GeoCast

Dmap : Demond Map

EP : Encounter Probability

FCA : Formal Concept Analysis

FCD : Floating Car Data

GDB : Global data base

HetNets : Heterogenous Networks

ITS : Intelligent Transportation System

IVC : Inter-Vehicle communications

LDI : Location dependent information

P-IVG : Probabilistic Inter Vehicle Geocast

P2P : Peer-to-Peer

RLSMP : Region based location protocol service management protocol

RSU : Road Side Units

SAS-GP : Semantic and Self-Decision Geocast Protocol for Data Dissemination over vanet

SDDA : Smart Directional Data Aggregation

SDN : Software defined Networking

SUMO : Simulation of Urban Mobility

US : United States

V2I : Vehicle-to-Infrastructure

V2V : Vehicle-to-Vehicle

VANET : Vehicle Ad Hoc Network

WAVE : Wireless Access in Vehicular Environment

WiMAX : Worldwide Interoperability for Microwave Access

ZOR : Zone of Relevance

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Context and motivation

Owing to the embedded devices into modern cars and the great progress of wireless technologies, issues related to transportation (e.g., traffic congestion, road safety and driver comfort) are grasping more and more interest. In fact, thanks to these embedded devices, vehicles are able to detect several types of basic events (e.g., accidents, empty parking places, bottling, obstacles, weather conditions, road cut, to mention a few). Interestingly enough, vehicles are also able to exchange such information through a Vehicular Ad-hoc Network (VANET). Indeed, this network enables advanced Intelligent Transportation System (ITS) services including various safety and non-safety applications. ITS standards perform communication in VANETs so as to support transportation and cooperation services [12]. As far as the development in the wireless communication field is generally concerned, the ITS applications are developed on the basis of car-to-car communication standards (i.e., Dedicated Short Range Communications (DSRC) and Wireless Access in Vehicular Environment (WAVE)) [12]. For instance, the DSRC uses the dedicated 75 MHz frequency spectrum (in the range 5.85-5.925 GHz) for Vehicle to Infrastructure (V2I) and Vehicle to Vehicle (V2V) communication [12, 25]. All VANET applications must share this allocated bandwidth as shown in subsection 1.1.

It is important to note that Inter Vehicle Communication (IVC) is enabled for most applications, namely entertainment, enhanced driving, and active safety. Subsequently, it is a must to manage the available limited-bandwidth carefully and efficiently as it is a scarce resource [25]. Thus, communication VANET protocols need a careful design. In this respect, information has to be exchanged between mobile vehicles in an efficient way by avoiding

(a) The huge number of parking places **(b)** The huge number of road segments in a city

Fig. 1.1 Vehicles broadcasting same events

as far as possible the broadcast storm problem [1]. Indeed, a core and challenging issues in VANET are the design of efficient *data dissemination* and *aggregation* protocols. The ultimate goal of the data dissemination protocol is to inform vehicles about important events. The thriving challenge would be to maximize the reachability ratio, i.e., by only informing the interested vehicles, and by avoiding the broadcast storm problem. On the other hand, data aggregation is an interesting approach, allowing to integrate several data about the same event to generate a summary (or aggregate) leading to reduce network traffic. Hence, data aggregation process consists of several steps allowing to merge, update, or delete certain information that are duplicated, similar, or expired. Indeed, the goals of this thesis are to introduce effective and efficient data dissemination and aggregation protocols.

[1]The broadcast storm is known to lead to network saturation as well as conflict and collision issues

1.2 Contributions

The major contributions of this work are shown as follows:

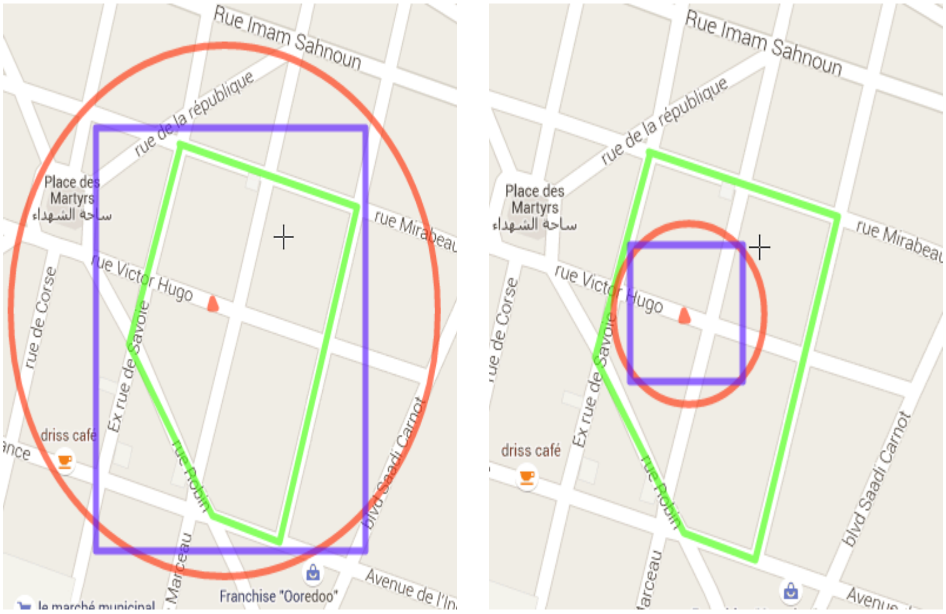
1.2.1 Data dissemination in VANETs

A careful scrutiny of the pioneering vehicle-to-vehicle data dissemination approaches highlights two main categories to do that: broadcast- or geocast-based techniques. It is worthy to mention that the main moan to be addressed in the broadcast techniques relies on the dissemination of messages to all vehicles in the networks without exception, which is costly and useless in various scenario since vehicles are not always concerned by all events. This drawback has contributed to the emergence of another alternative: geocast-based techniques allowing information delivery to vehicles inside a specific region. Indeed, geocasting is the most feasible data dissemination approach for VANET applications, more especially in safety applications, since safety events are of interest to vehicles within a specific area standing close to the event location. Interestingly enough, an geocast-based approach needs to cope with the following requirements:

1. Determine the geocast area, also called *Zone Of Relevance* (ZOR) of an event;
2. deliver the message to all vehicles within the ZOR; and
3. keep the geocasted message alive in the network for a desired delay, such that the disseminated information could reach all the arriving vehicles.

Although the literature witnesses a wealthy number of geocast-based techniques for data dissemination, only few of them consider all of the three aforementioned requirements. Indeed, the most challenging issue for geocast protocols is the definition of the ZOR for an event dissemination. Unfortunately, the existing geocast-based protocols define the ZOR in a simplified manner as: a rectangle [19] [35] [26], or a circle or polygon [1] which is strongly restrictive. For example, in Figure 1.2, the green region is considered as the ZOR of an accident warning; however, the target regions are specified as a circle or a rectangle, which are smaller or larger than the ZOR. Therefore, in the first case (Figure 1.2 **(a)**), a lot of unnecessary messages are exchanged and send to unconcerned vehicles. Whereas, in the second case (Figure 1.2 **(b)**), many concerned vehicles won't receive the message.

In order to cope with the aforementioned requirements, we propose in the first part of our thesis a new geocast approach for urban area, called Data Dissemination Protocol based on Map Splitting (DPMS) . DPMS aims to reach a high reachability ratio as well as a high



(a) The target region is too wide **(b)** The target region is too narrow

Fig. 1.2 Zone of relevance vs. target region specifications

geocast precision by sending messages only to vehicles in the ZOR with a minimum overhead cost. The main originality of DPMS is the computation of the ZORs through the unveiling of strong connections between the set vehicle trajectories and a set of regions. Thus, the determination of the ZORs can be seen as an instantiation of the cover set problem, i.e., finding the minimal coverage of the boolean matrix, in terms of formal concepts, keeping track of the relationship between vehicle trajectories and map regions.

1.2.2 Data aggregation in VANETs

To cope with the limited bandwidth, the various requirements emerging from the applications themselves and the highly dynamic network topology, VANET-based applications requiring information gathering and aggregation (e.g., traffic information systems, weather information systems, parking spaces and travel time predictions) have to deal with the following challenges:

1. How to decide if two or more FCD messages must be aggregated or not?
2. How to select timely data to be aggregated (since not all data need to be collected after a certain time)?
3. How to consider data from far vehicles?

4. How to filter the unnecessary and duplicated FCD messages in order to avoid affecting the accuracy of the shared traffic information?

5. How to take into consideration the road traffic signals and speed limitations?

In the literature, several data aggregation techniques have been proposed [20, 27, 44, 57]. Nevertheless, in a high complex urban and highway network, the large amount of traffic information needs smart filtering and selection criteria after the aggregation process. However, the existing techniques mainly focus on combining the correlated items but none of them tries to use a suppression technique in order to eliminate the duplicated messages. Moreover, they take into consideration neither the vehicle directions nor the road speed limitations in the provided traffic information. To overcome this shortage, we introduce in the second part of our thesis a new data aggregation protocol called Smart Directional Data Aggregation (SDDA) able to decrease the network overload while obtaining high accurate information on traffic conditions for large road sections. To this end, we introduce three levels of messages filtering: (i) filtering all FCD messages before the aggregation process based on vehicle directions and road speed limitations, (ii) integrating a suppression technique in the phase of information gathering in order to eliminate the duplicate data, and (iii) aggregating the filtered FCD data and then disseminating it to other vehicles. The performed experiments show that the SDDA outperforms existing approaches in terms of effectiveness and efficiency.

1.3 List of Publications

The idea and the contributions presented in this thesis are part of several peer reviewed research papers. In this subsection we give the list of our publications grouped by type and sorted by date.

International Journal

- **Sabri Allani**, Taoufik Yeferny, Richard Chbeir (2017). A scalable data dissemination protocol based on vehicles trajectories analysis. Ad Hoc Networks, 71:31–44.

International conferences

- **Sabri Allani**, Taoufik Yeferny, Richard Chbeir, Sadok Ben Yahia (2016). Dpms: A swift data dissemination protocol based on map splitting. In Proceedings - International

Computer Software and Applications Conference (compsac 2016), volume 1, pages 817–822.

- **Sabri Allani**, Taoufik Yeferny, Richard Chbei, Sadok Ben Yahia (2016). A novel VANET data dissemination approach based on geospatial data. In The 7th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2016)/The 6th International Conference on Current and Future Trends of Information and Communication Technologies, pages 572–577.

- **Sabri Allani**, Taoufik Yeferny, Richard Chbei, Sadok Ben. (2018). Smart directional data aggregation in vanet. In Proceedings - International Conference on Advanced Information Networking and Applications (In press).

1.4 Thesis organization

The rest of the thesis is in five chapters:

1. In Chapter 2, we describe the pioneering existing approaches of data dissemination and aggregation in order to show their limitations/drawbacks.

2. In Chapter 3, we thoroughly describe our data dissemination protocol, called Data Dissemination Protocol based on Map Splitting (DPMS) [5] [4] [3], before presenting in the simulation settings and the evaluation of the proposed DPMS protocol. Also, a comparison between DPMS and other main geocast protocols is presented.

3. In Chapter 4, we introduce a new data aggregation protocol called Smart Directional Data Aggregation (SDDA) [2]. Performance evaluation of the introduced protocol is then presented.

4. Chapter 5, concludes this thesis and pins down several future directions.

Chapter 2

Related work

2.1 Introduction

This chapter is devoted to the presentation of vehicular networks (VANETs). Also, we made a comparison between VANETs and Mobile Ad-hoc Networks (MANETs) to emphasize their different characteristics. Next, we will discuss the architectural concepts of communications, the scope of applications, the characteristics, the constraints and the challenges of VANETs. Finally, we present a detailed overview of data dissemination and aggregation in VANET.

2.2 Background information

Since their inception, Vehicle Ad-Hoc Networks (VANETs) have become a very broad area of research taking into consideration the specifications of the city and the networks in relation to ad-hoc mobile networks (MANETs) as well as its applications, architectures, capabilities and its data management mechanisms.

2.2.1 Ad Hoc Networks

Ad hoc networks, are wireless networks able to organize themselves spontaneously and autonomously in the environment in which they are deployed without previously defined infrastructure, created on demand to meet a need specific. The task of network management is distributed over the set of entities communicating over a wireless link, these entities are often called "nodes" [13].

The IETF MANET group provides a more precise definition in the introduction of [13]: "An ad hoc network includes mobile platforms (for example, a router interconnecting different hosts and wireless devices) called nodes that are free to move without constraint."

An ad hoc network is therefore an autonomous system of mobile nodes. This system can operate in isolation or interfere with fixed networks through gateways. In the latter case, an ad hoc network is a network end. Nevertheless, the terminology "ad hoc network" is relatively not very explicit. This is probably why the scientific community is replacing it sometimes by that of "spontaneous network". Figure 2.1 shows an example of an Ad Hoc Network.

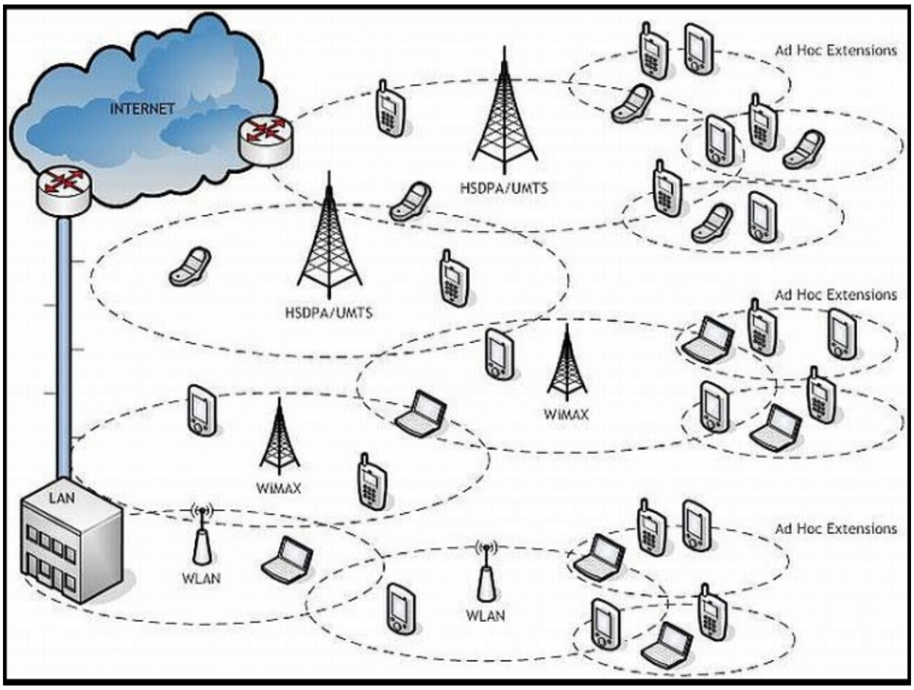


Fig. 2.1 Example of an ad hoc network [23]

In figure 2.1 wireless networks are capable of organizing without previously defined infrastructure. For example from one equipment to another without infrastructure (access point).

2.2.2 Mobile ad hoc network (MANET)

The ad hoc mobile network, usually called MANET (Mobile Ad hoc Network) is an autonomous system consisting of a dynamic mobile node interconnected by wireless links without using any fixed infrastructure and without centralized routing protocols in MANET networks [1] [13]. Nodes are free to move randomly and, therefore, can change the network structure quickly and unpredictably.

[1] https://tools.ietf.org/html/rfc2501

Mobility: The mobility of the nodes is obviously a very characteristic specific to ad hoc networks. This mobility is intrinsic to the functioning of the network. It differs from nomadicity (mobility of only terminal nodes) or from roaming (static equipment but can be moved)[51]. In an ad hoc network, the topology of the network can change rapidly, randomly and unpredictably and the techniques of routing of conventional networks, based on pre-established routes, can no longer function correctly [13].

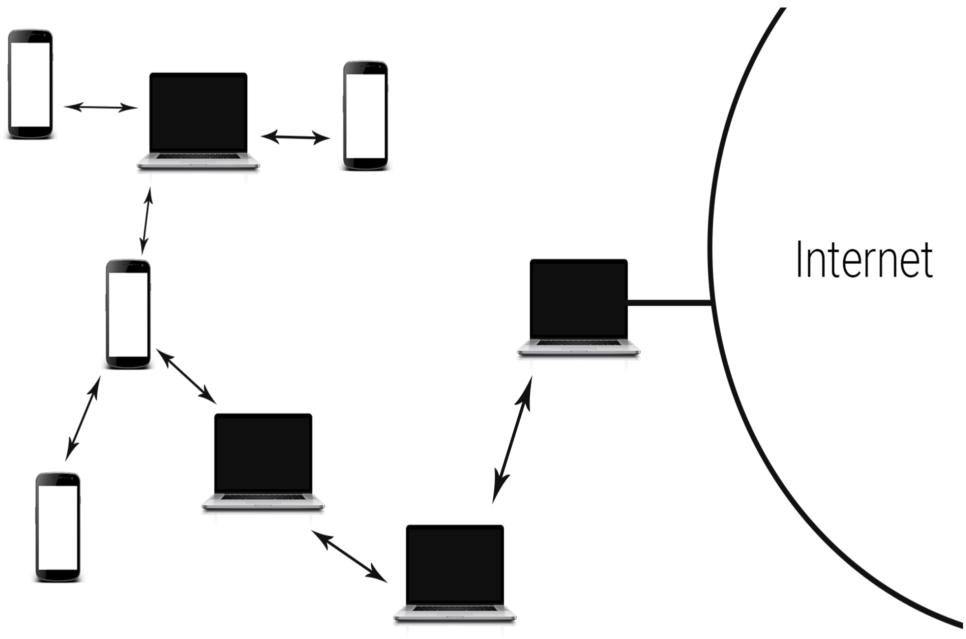


Fig. 2.2 Example of Mobile Ad Hoc Network (MANET) [57]

In figure 2.2 the MANET networks are capable of organizing autonomous and without any previously defined infrastructure. Moreover, the mobility and the speed of nodes is limited in space. For example, mobiles and laptops can share infirmations between them and construct a MANET network.

2.2.3 Vehicle ad hoc network (VANET)

The origin of ad-hoc networks [45] dates back to the beginning of the 1970s with the ALOHA project commissioned by the University of Hawaii. This project been conducted to allow computers in the Hawaiian Islands to be connected to each other by radio waves, in a system of communication to a jump. Directly inspired by ALOHA, the military agency United States DARPA commissioned in 1973 the PRNet project (Packet Radio NETwork) to study packet radio communications in autonomous networks. The use of ad-hoc networks is justified as the installation of an infrastructure proving to be inappropriate, whether for economic reasons or for physical reasons (e.g. geographically difficult access areas), new fields of application have recently been envisaged. They can be used for the communication of backup units, when a natural disaster (such as an earthquake, flood) destroys infrastructure telecommunications and the establishment of a satellite link for each entity in question. Communication represents too high a cost. The first applications of MANETs have been running in military environments or battlefields. Mobile nodes are soldiers, planes, etc. For military applications, MANETs allow to retrieve information in a given hazardous area, or to monitor the movement of the enemy. Another area of application particularly well adapted for ad-hoc networks concerns the sensor networks. In the year 2000, MANETs were deployed in Transportation Systems(ITS). They were particularly used for inter-vehicle communication, which gave birth to VANETs as shown in figure 2.3.

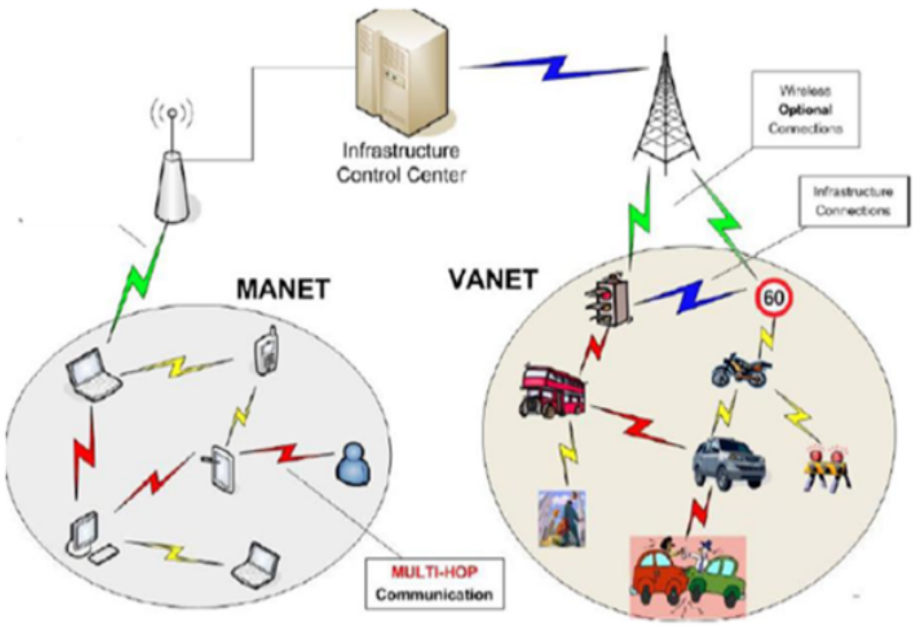


Fig. 2.3 Vehicle Ad Hoc Network (VANET) [56]

Owing to embedded devices into modern cars, issues related to transportation, like traffic congestion, road and safety driver comfort, are grasping the interest. In fact, thanks to these systems, vehicles are able to detect several types of information, e.g., an accident in the road, an empty place in a parking, bottling, obstacles, weather, road cut, to cite but a few. Interestingly enough, vehicles are also able to exchange such information using Inter Vehicle Communications (IVCs) [31], which are based on short-range wireless technologies, to build a Vehicular Ad hoc Network (VANET). In this respect, a core and challenging issue in vehicular networks is the design of an efficient dissemination protocol that informs vehicles about interesting safety and non-safety events. The thriving challenge would be to maximize the delivery ratio by avoiding as far as possible the broadcast storm problem. The latter is known to lead to network saturation as well as conflict issues and collision. A careful scrutiny of the pioneering vehicle-to-vehicle dissemination approaches highlight that they can be split into two categories : event-driven, scheduled, or non-demand approach. In the remainder, we will argue for choosing the event-driven approach, where data dissemination is carried out through broadcast or geocast techniques. Worth of mention, the main mean that can be addressed to the broadcast stands in the costly dissemination of messages to all the vehicles standing in the neighborhood. This drawback has led us to naturally opt for the geocasting technique, which is the most feasible data dissemination approach for VANET applications [39, 50].

Communication architectures

The main purpose of VANETs is to ensure communication and exchanged information between vehicles, as well as between vehicles and road elements, such as intersection lights and traffic signs. So we can distinguish three types of communications:

- **Vehicle Vehicle Communications (V2V)**: This architectural mode is based on a simple inter-mobile communication that does not involve axes infrastructures. In this type of communication, a VANET becomes a special case of MANETs.

- **Vehicle Communications to Infrastructure (V2I)**: In this approach, the exchange information involves infrastructure points of the RSU (Road Side Units) route [50]. Examples include road signs, intersection lights, satellites, antennas, Internet, etc.

- **Hybrid Communications (V2X)**: The combination of two architectures (V2V) and (V2I)which means a hybrid communication. In fact, the use of vehicles as a relay of communication permits not only extending the limited ranges of the infrastructure but also saving the use of RSUs. A special case of this mode of communication is Vehicular Sensor Networks (VSN). Figure 2.4 shows an example of hybrid V2X communication in VANETs.

Vehicular communication environment

The recovery points for routing in the VANETs come out not only of the fast mobility of vehicles, but also of the spatial and temporal diversity, the density of the traffic and the

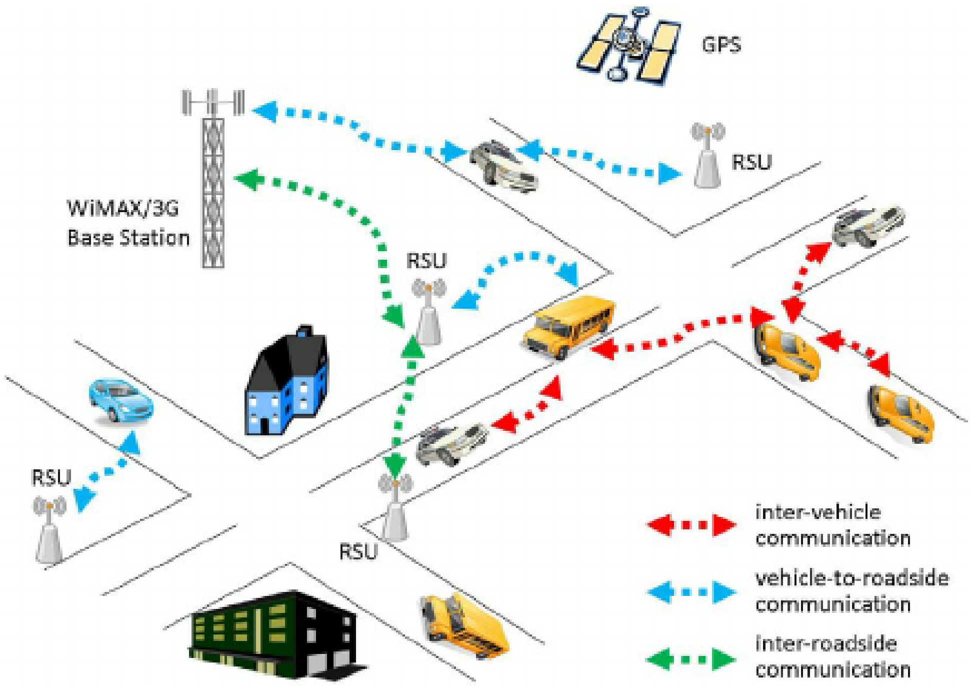
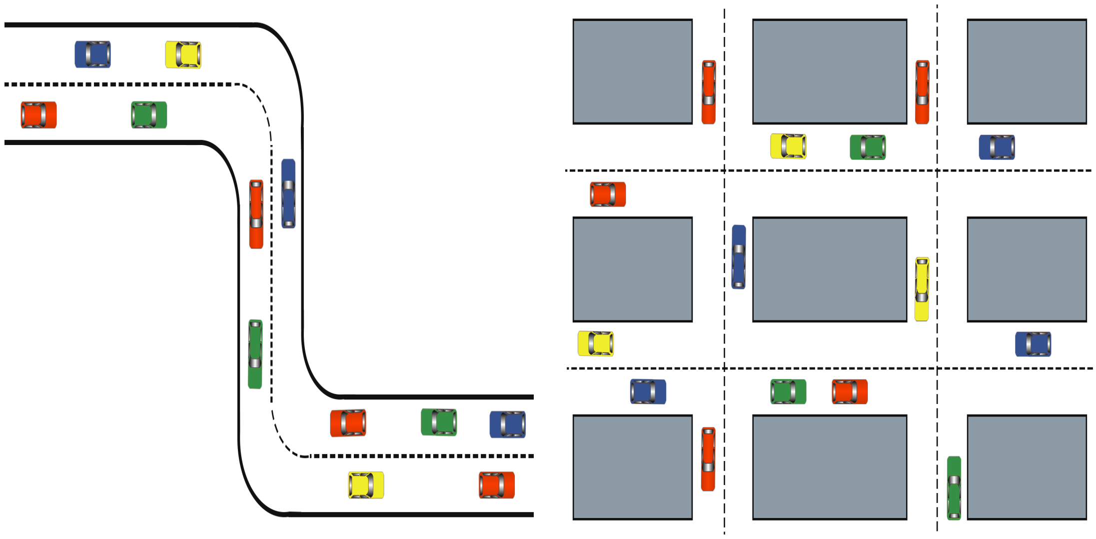


Fig. 2.4 Example of hybrid V2X communications in VANETs [56]

communication environment, which are sometimes unfavorable to the establishment of communications. VANETs require the consideration of greater environmental diversity. Due to the mobility of vehicles, it is possible to move from an urban environment characterized by many obstacles to the propagation of signals, having a highway environment possessing a specific characteristics. Generally speaking, mobility environments are divided into two parts: highways and cities. On highways, vehicles can only move in two directions over several tracts. Nevertheless, in the cities, vehicles can turn very often and RoadSide Units (RSUs) are much more present. Environmental diversity is illustrated in figure 2.5.

Figure 2.5 shows the tow types of the vehicular communication environment, where figure 2.5 (a) represents a highway environment and figure 2.5 (b) represents an urban environment.

Because the network topology is highly dynamic and the environment imposes constraints that must be taken into account when routing data, the routing function in VANETs is a fundamental problem. Therefore, given its complexity, this problem is largely dealt with by researchers. To solve it, several protocols have been proposed. In fact, since VANETs is only a particular version of MANETs, a lot of work has been done to use dedicated dissemination protocols for them. However, other work has shown that the MANET protocols are not suitable for the vehicular context. Thus, dedicated protocols for VANETs have been proposed [25]. In what follows, we begin by presenting the main protocols used for MANETs. Then, we expose the data dissemination and aggregation solutions conceived for VANETs.



(a) Highway environment **(b)** Urban environment

Fig. 2.5 The vehicular communication environment

2.3 Data dissemination in VANET

VANET security applications aim at improving drivers and passengers safety on roads by notifying any dangerous situation. Generally, these applications are based on data dissemination, which are most of the time periodic. This is to enable the state of the road and surrounding vehicles. The VANET data dissemination protocols can be categorized as infrastructure-based and infrastructure-less [1]. The infrastructure-based protocols [30, 21, 43] use RoadSide Units (RSU) [50] in junctions and along the roads to store and disseminate VANET messages. These protocols generally achieve good results. However, they rely on costly infrastructures. In this respect, infrastructure-less protocols have been recently attracting more interest by the scientific community due to their capability to disseminate information without relying on a costly infrastructure. They can be categorized into either broadcast or geocast data dissemination protocols [40]. In the following, we discuss the most recent ideas including intelligent broadcasting and geocasting techniques. Applications widely range from emergency messaging to exchanging traffic information.

2.3.1 Broadcast-based data dissemination

Broadcasting techniques are frequently used in VANETs for data sharing, traffic information, weather, entertainment and commercial announcements, with the aim to disseminate information to all vehicles, without exception, using blind or moderated flooding mechanisms.

Within a blind flooding, a vehicle broadcasts each received or detected information to all neighboring vehicles. This approach can increase reachability by informing all interested vehicles. However, it undoubtedly leads to network congestion, conflict and collision issues, often named the broadcast storm problem [47] [52].

In the literature, several research studies have attempted to improve this by adopting various suppression techniques, which are probabilistic (e.g., weighted p-persistence) [47], timer-based (i.e., Slotted 1-persistence) [47], or hybrid (*i.e.*, Slotted p-persistence) [47]. These techniques have tempted to reduce the broadcast redundancy and the packet loss ratio by decreasing the number of vehicles spreading the same message while ensuring a high reachability.

In weighted p-persistence [47], a forwarding probability is assigned to each neighboring vehicle according to its distance to the message broadcaster. A higher forwarding probability is assigned to the vehicles that are located farther away from the broadcaster. After a fixed waiting time (e.g., 2 ms), the receiver rebroadcasts with the assigned probability if it does not receive duplicate copies of the message as shown in figure 2.6.

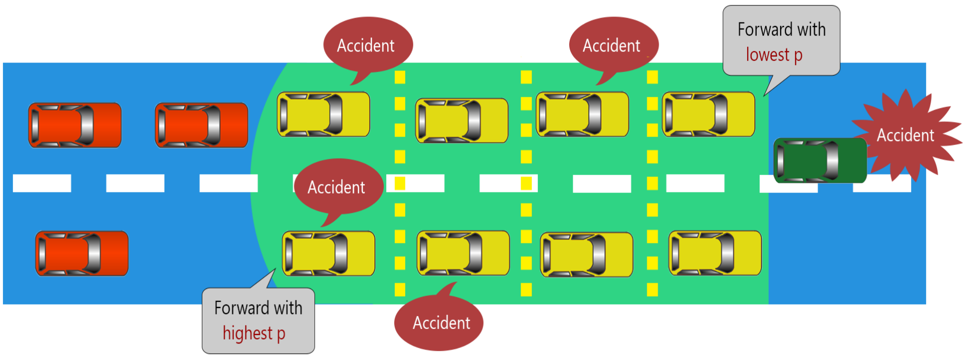


Fig. 2.6 Weighted p-persistence scheme

In Slotted 1-persistence [47], different waiting time slots are assigned to the neighboring vehicles depending on their locations. A shorter waiting time is assigned to the vehicles located in the farthest region from the broadcaster. Upon receiving a message, the receiver checks the packet ID and rebroadcasts it with a probability 1 at the assigned time slot if it receives the packet for the first time and has not received previsouly any duplicates before its assigned time slot; otherwise, it discards the packet as shown in figure 2.7.

In Slotted p-persistence [47], a forwarding probability and a slotted waiting time are used. Indeed, a higher probability and a shorter slotted waiting time are assigned to the vehicles located in the farthest region from the broadcaster. Hence, the receiver rebroadcasts with the pre-determined probability p at the assigned time slot if it does not receive any message echo from the neighboring vehicles as shown in figure 2.8.

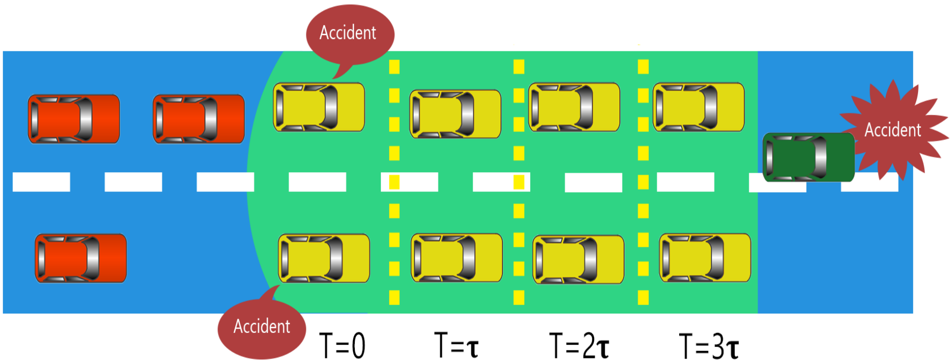


Fig. 2.7 Slotted 1-persistence scheme

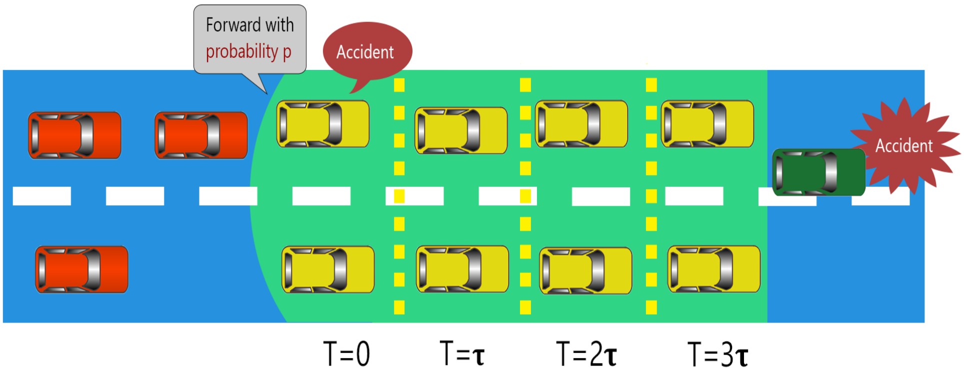


Fig. 2.8 Slotted p-persistence scheme

In [47], the authors conclude, after performing several simulations, that the Slotted-1 persistence is the most efficient technique. In [38], the authors address the multi-directional dissemination issue in an urban scenario within a dense network. Therefore, they introduce an optimized version of the Slotted 1-persistence where the main optimization is to give a high priority of dissemination to the farther vehicles having the same direction as the event message.

However, in various scenarios (particularly safety event dissemination), broadcast techniques are unsuitable since only some vehicles near the event location could be concerned. Geocasting techniques have emerged as an interesting alternative since they can help overcoming the broadcast storm problem more effectively [40]. In the following, we provide an overview of main recent works that have tackled the problem of data geocasting in vehicular networks.

2.3.2 Geocast-based data dissemination

Geocast protocols aim at disseminating information only to vehicles inside a specific geographical area, called ZOR or Zone of Relevance [8]. It is worth mentioning that, in geocast protocols and after determining the ZOR – probabilistic, timer-based, priority-based or hybrid suppression broadcasting techniques can be used to disseminate the message inside the ZOR. The vehicles receiving the message outside the specified area simply ignore the message [28].

Figure 2.9 shows some of pioneering geocast protocols in each category.

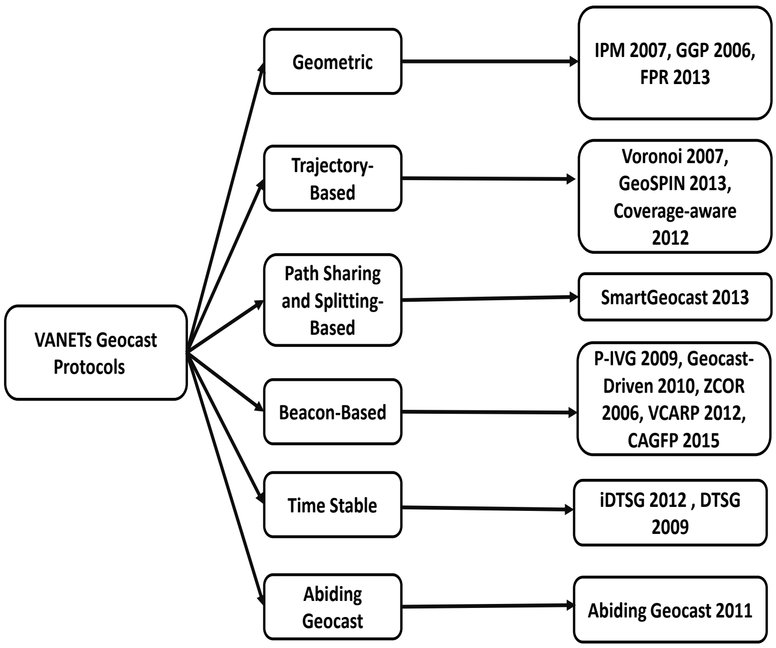


Fig. 2.9 Categories of geocast protocols in VANETs

The SmartGeocast [55] approach is the most worth of cite one falling within the category *Path, Sharing and Splitting-Based approaches*. In the latter, two phases are included: geocasting initialization and geocasting maintenance. The former step relies on the definition of landmarks, tokens and messages using a digital map, event type and event position. The concept of path splitting is used when landmarks is split as head lighthouse (HLH) or tail lighthouse (TLH). The HLH is located at the upstream crossroad where an abnormal event occurs, whereas the TLH is located at the crossroad close to a target region. The concept of path sharing is used whenever multiple messages through multiple paths are aggregated into a single message and delivered through a single path, leading to reduced bandwidth consumption. During the maintenance phase, the vehicle that holds aggregated messages should retransmit them to other vehicles based on the concept of regional autonomy. Hence, this splitting of the area into small target areas can reduce the message redundancy and maintenance cost.

In *Beacon-Based approach*, the information is disseminated to ZOR periodically based on a one-hope time slot. The beaconing operation can be done before or during geocasting the message, in order to exchange several information (e.g., position, signal power, etc.) between neighboring vehicles. Ibrahim et al. [19] introduced a beacon-based approach named *probabilistic Inter-Vehicular Geocast (P-IVG)*. The latter selects the farthest vehicles, standing within the transmission range, for further dissemination of the message. When several nodes have an equivalent distance and starts the same timer, P-IVG is based on vehicle density information.

In *Time Stable Geocast*, the message must be kept alive inside the specific geocast area, however, to deal with such thing a time stable geocast period. Worth of cite, Rahbar et al. and Kheawchaoom et al. [35] [26], respectively, introduced two versions of Dynamic Time-Stable Geocast protocol (DTSG). The main focus was to guarantee the message delivery with a low cost, such that the protocol should dynamically update the waiting stable time. The first version of DTGS includes two phases: a pre-stable period and a stable one. The first one comes back to geocasting the message to the specific region. Whenever a vehicle detects such critical event, it immediately broadcasts it and keeps rebroadcasting. This process comes to an end as far it receives the same message from a rely vehicle on the opposite side.The most distant vehicle is selected for the relaying mission based on a stable period of time. The ultimate goal of this stable period is to keep alive the message within the geocast region whenever it is still woth of pertinence for this region.

In *Abiding Geocast*, the information is only disseminated to the geocast area for a specific duration of time in order to keep alive all messages and disseminate them to all vehicles [29]. Hence, three basic approaches are proposed: server-based, neighbors approach and neighbor's election-based approach: Neighbor's approaches are the most used in abiding geocast. They use both unicasting and single-hop broadcasting for data dissemination. The unicasting is used to transmit the new message to the specified location utilizing infrastructures while broadcasting is used to disseminate the message within the specified location. In Yu et al. [29], the opposite side vehicles are exploited to make an optimized version of [53] safety abiding geocast protocol. Indeed, a dynamic adjustment of waiting time after the next broadcast and the selection of the best opposite lanes vehicles to rebroadcast the message in such critical dangerous event can deal with the broadcast storm problem [47] and avoid unnecessary broadcast.

In this respect, Ibrahim et al. [19] introduce a probabilistic and beacon-based approach named p-IVG (*probabilistic Inter-Vehicular Geocast*). The p-IVG approach tackles the drawbacks of timer based broadcasting approach (e.g., the Slotted 1-persistence [47]) in dense networks. In fact, in a dense network, a very large number of vehicles within a given area (e.g., ZOR) have almost the same re-broadcast probability or waiting time, so they re-broadcast the packet at the same time. This would lead to a local spatial broadcast storm. To overcome this drawback, the authors of the p-IVG proposed to set the re-broadcast decision probabilistic based on the surrounding vehicle density. Indeed, as far as several vehicles have an equivalent distance and start the same timer, p-IVG relies on vehicle density information to set the waiting time at each vehicle in order to solve the spatial broadcast storm problem. However, in p-IVG, vehicles exchange beaconing messages to determine the density of the surrounding vehicles, what undoubtedly leads to an extra network overload. Furthermore, the p-IVG only focuses on the highways scenario and does not pay attention to the ZOR determination issue. Indeed, it assumes the existence of a certain number of vehicles, within a rectangular area near to the event location, where the messages should be transmitted.

Rahbar et al. [35] and Kheawchaoom et al. [26] introduce, respectively, two versions of the Dynamic Time-Stable Geocast protocol (DTSG) for the highways scenario. The main focus is to provide low cost message delivery and to keep the message alive within the ZOR for a specific period of time (e.g., event life time). DTSG includes two phases: a pre-stable period and a stable one. The first one consists of geocasting the message to the specific region. Whenever a vehicle detects such a critical event, it immediately broadcasts (and keeps rebroadcasting) the event. This process ends when the vehicle receives the same message from a rely vehicle on the opposite side. After this period, the protocol moves to the stable period; i.e., the protocol tries to maintain the message within the target region for a specified period of time, whenever it is still worth of pertinence for this region. Like p-IVG, DTSG only considers the highways scenario and defines the ZOR as a rectangular shape.

Allal et al. [1] split the dissemination area, which is equivalent to a ZOR, into a set of sub-ZORs to deal with the temporal network fragmentation issue. To do so, they use simple geometrical shapes (e.g., circles, rectangles, triangles and polygons) to cover these sub-ZORs. In order to lower the message overhead and the processing time, the authors provide a technique allowing to determine whether some sub-ZORs are in the same direction so to send them the same message. It is worth mentioning that the authors delegate the affectation of the ZOR to a dedicated authority (as road safety services) which would provide the coordinates or designations of stretches of roads where vehicles could be affected by an event.

In the aforementioned geocast protocols, some approaches have defined the ZOR as a rectangle (e.g., [19] [35] [26] to cite but a few). Whereas, in some other approaches, it has been defined as a circle or a polygon [1]. Although Jochle et al. [22] conclude that a circular geocast area performs better in most of the scenarios, the ZOR specification remains a thriving challenge in geocast protocols.

Another interesting direction of research is based on exploiting vehicle trajectories and/or road map topology in furtherance of geocasting messages. In this respect, Delot et al. [14] introduce a dissemination protocol that follows a *forward-if-relevant* principle by which each vehicle receiving the event would decide whether the event should be further disseminated. Indeed, in this protocol, a vehicle decides to rediffuse or not the event based on the concept of encounter probability, denoted EP. The latter is an estimation of the likelihood that a vehicle will meet an event. If the computed EP is greater than or equal to a certain diffusion threshold, the message is considered as relevant enough to be rediffused by the receiving vehicle. Otherwise, the message is simply ignored. Doing so, an event is propagated to the neighboring vehicles while the event is considered relevant in the area, which leads to Dynamic Dissemination Areas (DDA). The main limitation of such approach is related to the fact that it does not keep the event alive within the DDA and the suppression technique, used to deal with the broadcast storm issue, is not efficient enough. For example, in a dense network, a large number of neighboring vehicles can have a high EP value. Consequently, all of them will rebroadcast the event which would undoubtedly lead to a local broadcast storm.

Alsubaihi et al. [6] introduce the SAS-GP protocol, which initially executes an algorithm to determine locally the semantic geocast area. Indeed, the geocast area is defined as all possible paths that lead up to the event location. Then, the protocol disseminates the information in three phases: spread, preserve, and assurance. In each phase, an appropriate timer is set based on the distances between the nodes, the transmission range, the predicted propagation, and the transmission delays of the medium. In the SAS-GP protocol, the ZOR might be a very large area, especially in an urban scenario with a complex road map, since it is defined as all possible paths leading to the event location. Consequently, several non-interested vehicles would receive the event message (since the vehicles heading the event location are not necessarily interested in the event).

To sum up, the aforementioned studies cannot effectively cope with all of the three predefined requirements of geocast protocols. For the sake of ensuring a high reachability ratio as well as a high geocast precision, we advocate that only appropriate messages needs to be sent to vehicles interested in the event with a minimum overhead cost. In this paper, we introduce a novel geocast protocol that highlights the following sighting features:

- Precised detection of the ZOR by applying a data analysis technique on vehicles' trajectories. Actually, in our protocol, the map is split into a set of regions. Then, for each region, we determine the set of regions composing its ZOR. Hence, a centralized database is dedicated to store the result of map decomposition and ZOR association.

- Better management of the event lifetime in the ZOR. Indeed, the geocasted message should remain in the ZOR for a desired delay, such that arriving vehicles can be informed. Only few approaches (e.g., [35] and [26]) addressed this issue by rebroadcasting the event during a certain delay, which significantly increases the network overload. To tackle this issue in our protocol, when a vehicle enters in a new region, it retrieves events received by other vehicles in the current region. By doing so, pertinent events are kept alive inside the region, without the need of rebroadcasting event messages as the existing geocast approaches do.

- Better event dissemination. In fact, we avoid the broadcast storm during the dissemination of events to vehicles within the ZOR by using a moderated Slotted-1 persistence technique. This suppression technique avoids the broadcast storm problem and guarantees a high *reachability*.

2.4 Data aggregation in VANET

In larger road networks, urban and highway, there is a huge amount of traffic information that quickly exceeds reasonable limits. In this respect, an appropriate aggregation mechanism can reduce the communication cost while obtaining useful aggregated information. Several related approaches have been proposed in the literature. A typical kind of Mobile Ad hoc Networks (MANETs) is Vehicular Ad hoc Networks (VANETs) [12]. In these latter, every network node, which essentially represents one vehicle, can communicate with an existing infrastructure or with other vehicles, for instance Road Side Units (RSU). In the recent past, there has been a fast development of the technology needed to support VANETs. Such a technology has been opted for by several pioneering automobile manufacturers [12], such as Honda, Toyota and the US Department of Transportation (DoT). That has helped novel ideas for VANET-based applications to emerge, namely for safety and non safety information. The latter includes internet access, P2P file sharing and multimedia streaming [34]. The former includes notifying traffic conditions and upcoming roadway hazards, as well as collecting and disseminating weather information [34]. The US DoT has mainly focused on safety applications in vehicular networks including both collision warning and assistance [12]. Intelligent Transportation System (ITS) standards have carried out communications in VANETs to support cooperation and transportation services [12]. To ameliorate ITS effectiveness and information exchanging, advanced technologies have been incorporated, like gesture and image recognition. To develop wireless communication, ITS applications have been evolved regarding car-to-car communication standards [12]. For instance, the DSRC uses the dedicated 75 MHz frequency spectrum (in the range 5.85-5.925 GHz) for Vehicle to Infrastructure (V2I) and Vehicle to Vehicle (V2V) communication [12]. All VANET applications must share this allocated bandwidth. In addition to that, Inter Vehicle Communication (IVC) technology is application-enabled for active safety enhanced driving, and entertainment. As a consequence, the careful and efficient management of the available limited bandwidth is a must, on account of being a scarce resource. As a result, the communication VANET protocols require a careful design. Thus, in the aim of coping with the limited bandwidth, multiple needs have emerged from the highly dynamic network topology and the applications themselves. In particular, applications based on VANETs, which necessitate information gathering and aggregation, must consider the following challenges: i) Which FCD message to aggregate? ii) when to aggregate data? iii) How to deal with information got from far vehicles? iv) How to separate out unneeded or duplicated FCD messages? v) How to account for speed limitations and road traffic signals ? data lot of information aggregation techniques in the literature have been put forward [20, 27, 44, 57]. Data aggregation use cases are summarized in table 2.1 . Nevertheless, within highly complex urban and highway networks, a big quantity of traffic data need selection criteria and smart filtering after the aggregation process. On the other hand, all existing techniques mainly have been chiefly focused on combining correlated items. However, no technique has eliminated duplicated messages. In the target of coping with the previous requirements, Smart Directional Data Aggregation (SDDA), a new aggregation protocol, is introduced [18, 27] . It is properly able to deal with traffic data in highway and urban conditions. This protocol rightly selects the FCD messages to aggregate. For this task, SDDA provides 3 filters: (1) The first one utilizes the vehicle's directions. Indeed, every vehicle aggregates not just information corresponding to its direction but also stores, carries and forwards uninteresting data. (2) The second one is carried out by using road speed limitations. Therefore, in case an average speed of received FCD messages surpasses maximum maximal permitted speed, it will be brushed aside, hence being replaced by the maxal permitted speed value. Accordingly, there will be an increase in aggregation accuracy. (3) The third one can be done by integrating a suppression technique [48] in to the purpose of eliminating, in the aggregation phase, the duplicated inputs through the provision of a high accurate time slot. Actually, the latter is based on the range and the distance to be waited for by each vehicle before rebroadcasting an FCD message.

Applications	**Events**	**Values**	**Mentioned in literature**
Traffic information system	Traffic jam	speed and positions known as Flooding Car Data (FCD)	Very Often
Weather information systems	Ice, rain , etc	Average temperature, visibility of road, Rain degree	Somtimes
Road condition warnings	Broken road, road constraction , icy Road, etc	Road id, Location, condition and street adresse	Somtimes
Parking Spaces	Available parking space	Number and location of free parking spaces	often
Trip Travel Time Prediction	Trip paths and time	Travel time and short pathes	Somtimes

Table 2.1 Data aggregation use cases

In [46], the authors put forward Self-Organizing Traffic Informational Systems (SOTIS) where vehicles would periodically exchange their speed and position on the road. Each vehicle in SOTIS would calculate average speeds from neighboring vehicles before rebroadcasting the aggregated data. It would basically provide an outline of the current movement conditions through the use of periodic road information broadcasting. On the one hand, the main SOTIS limitation was related to the broadcast of the periodic Floating Car Data (FCD) since it did not ensure or even indicate how duplicated [FCD] messages that came from the same road segment could be aggregated together. On the other hand, the aggregation of the same duplicated FCD messages would reduce the accuracy of the final aggregated records. Figure 2.10 depicts the self-organizing traffic informational systems in SOTIS.

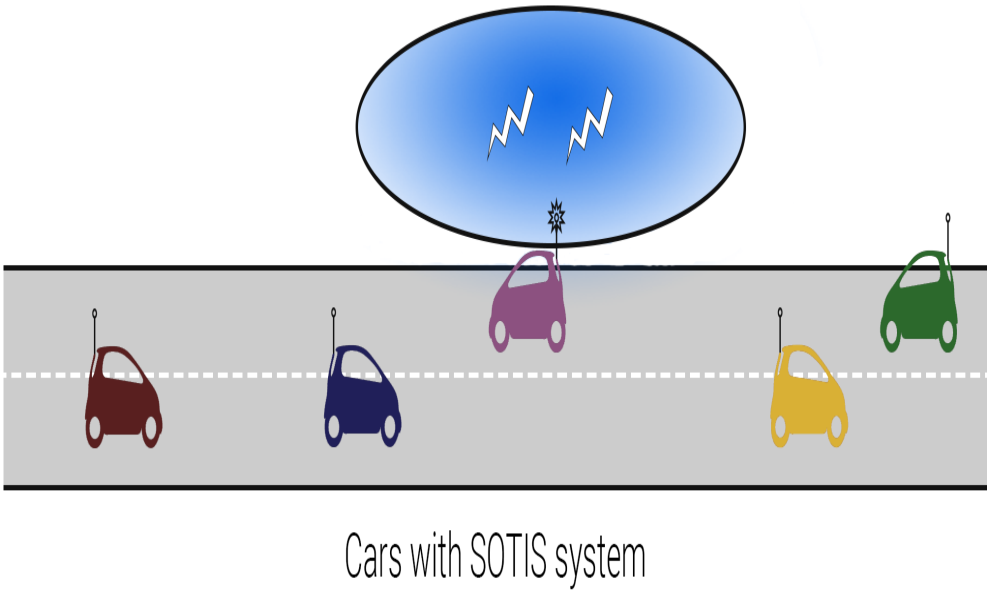


Fig. 2.10 Self-Organizing Traffic Informational Systems (SOTIS) [46]

Traffic View is another interesting approach described in [34]. It proposes a similar periodic approach based on broadcasting beacons that contained the FCD, like the average speed of a current road segment and its traffic density. The main difference between SOTIS [46] and Traffic View was the computation of the average speed since the aggregation process in Traffic View was an accumulative of average speeds in the road (starting from the quickest vehicle to the slowest one). On the contrary, within SOTIS a vehicle in the center could aggregate all neighboring vehiculars' data within its range. It was blatant that the aggregated record in Traffic View consisted of just one time-stamp value, one position and one speed, as well as vehicular IDs. Hence, this would use better the bandwidth when transmitting messages to every individual vehicle. Figure 2.11 depicts data aggregation with TrafficView.

CASCADE was suggested in [20, 18] as an optimized Traffic View version [34]. It allowed compressing syntactic data in the aim of optimizing the use of a wireless channel and at the same time guaranteeing accurate aggregated information. Furthermore, it would divide a road into 12 rows (16m $\times$ 126m) of a cluster leading to a 1.5 km visibility (named a local view). Thus, when vehicles in a local view cluster shared their FCD messages with

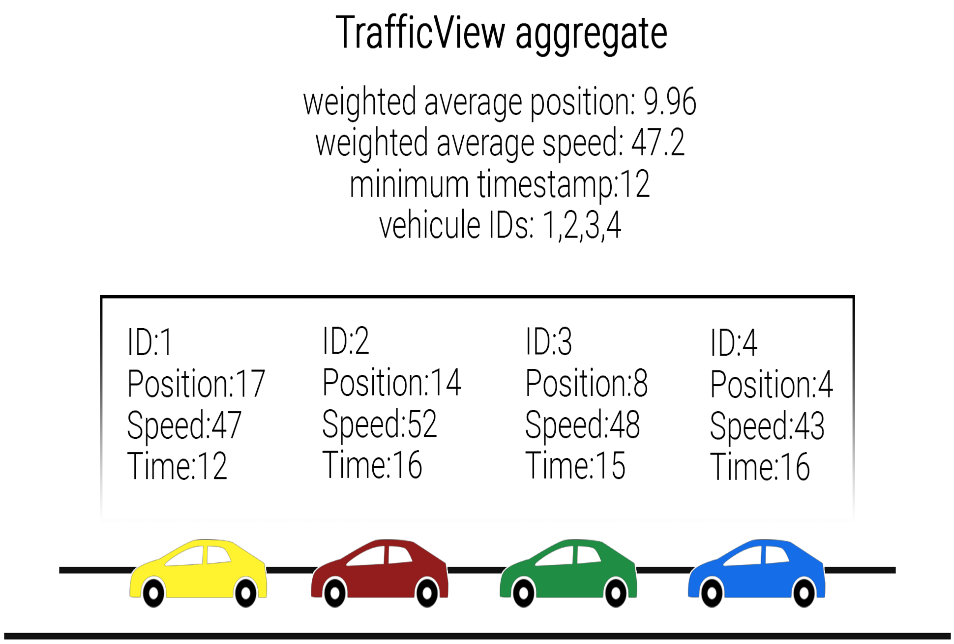


Fig. 2.11 Data aggregation with TrafficView [34]

another cluster, they had an extended view of the whole road segment. Figure 2.12 represents the data aggregation and compression with CASCADE.

Tsai introduced in [44] a hybrid Aggregating Data Dissemination (ADD) approach that combined both the V2V and V2I models. This approach also aggregates the number of available free parking spaces in a big region. To do that, the author split a map into a grid structure of square regions. The geodesic distance between these regions was in fact the Road Side Unit (RSU) communication range. The author defined four data aggregation levels in every region. In the first aggregation level, each vehicle would send its parking place, id, position and speed to the RSU center of the region. The RSU would aggregate all received data in the second level before rebroadcasting it to all vehicles in the region. In the third level, the vehicles in the extreme regions would share their traffic information, with the RSU sink. Finally, this later would aggregate, in the fourth level, all information coming from various regions before rebroadcasting it.

Kumar and Dave introduced in [27] a new multi-criteria decision-making for data aggregation. As a matter of fact, the proposed approach assisted a vehicle to decide about the relevance between data, for instance vehicle speed, vehicle direction and free parking space. Hence, the suggested system could decide if two or more input data were similar enough (syntactically or semantically) to be aggregated or not. To achieve this, the authors represented the knowledge base as a KD-tree data structure in order to check the relevance

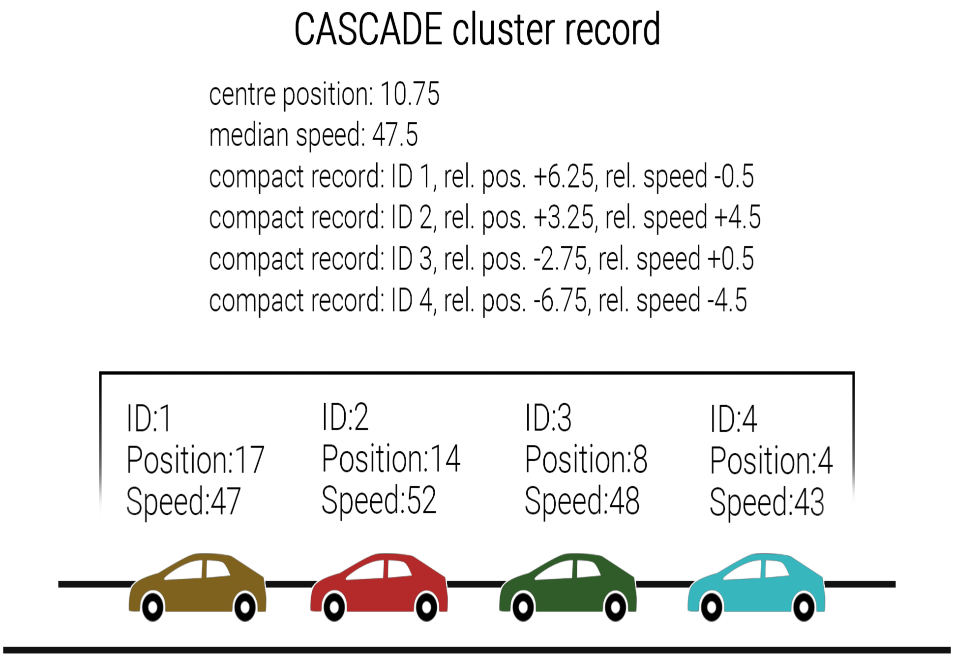


Fig. 2.12 Data aggregation and compression with CASCADE [20, 18]

between nodes using the graph characteristics. Although interesting, the provided approach had major drawbacks but mainly it considered the location of aggregates and ignored all other properties (e.g., vehicle directions, maximum allowed speed, etc.). Indeed, this aggregation decision would only consider data coming from the same road segment.

Time aggregated graphs were introduced by George et al. in [17, 16]. They allowed formalizing the road networks and the spatio-temporal properties of the road as a graph data model which would generally support shortest-path query graph algorithms. These time-aggregated graphs could annotate the properties of edges and nodes with the intervals during the time of vehicle presence. However, to decide whatever multiple items could be aggregated or not, the spatio-temporal model was only based on the vehicle travel time, the event life time, and the location. Thus, the major drawback of such a model was that one could not exactly estimate the vehicle travel time, since it could change its direction at any time.

In the same vein, Zeki et al. introduced in [54] an aggregation structure for events produced and exchanged in vehicular networks. More precisely, the proposed structure was based on a spatio-temporal model having two levels. The first one is a physical level consisting of a repository shared between all vehicles to share information without loss. The second one was a logical level where each driver would define his/her preferences for what information (s)he was interested in. This model would manipulate the same shared knowledge base between all vehicles. The ability of being duplicate insensitive and being able to guarantee lossless exchanged information was a significant characteristic of their data structure. In addition, the storage space, mainly needed for the aggregation structure, was particularly limited to include a governable number of temporal dimensions. The main limitation of this solution was the maintenance and the privacy issues of such a shared knowledge base [32]. Moreover, the overhead level was still high due to the problem of duplication and the absence of data filtering before the aggregation process.

In [36, 37], the aggregation of the FCD was only based on the geographical characteristics of the area. For that purpose, the authors introduced the Region-based Location protocol Service Management Protocol (RLSMP) which aimed to reduce the updated positions as well as reducing the number of messages generated to locate car positions. Although the aggregation solution clearly reduced the network overload, it resulted in: i) more packet collisions and consequently more re-transmissions essentially due to the fact that the exchanged packets had a big size, and ii) longer delays owing to the processing carried out on the data.

In [49], the authors proposed a road real-time visual car navigation system, where a driver can obtain location-dependent information (LDI). In this system, each vehicle has a Demand map (Dmap), a data set representing the geographical distribution of the strength of demands for LDI. Each vehicle shares a subset of a Dmap with other vehicles using flooding method. However, unneeded or duplicated LDI may be broadcasted in the network then network resources may be wasted and overload ratio will rise up.

In [58], the decision to aggregate, carry or forward FCD was based on the current collected traffic information. The proposed approach relies on data centers to store FCD. Indeed, the authors formulate the FCD gathering issue as a scheduling optimization problem and demonstrate that it is a NP-hard problem. Thus, the filtering process is solved by a genetic-based heuristic algorithm called RIDE and a dynamic programming solution. The main limitation of RIDE is the high cost of deploying data centers. Also, the amount of data to be transferred could be large and the network communication overhead must be considered.

In [15], Xiaoyu et al. introduced an adaptive software-defined networking (SDN) scheme of vehicle clustering for aggregated traffic information based on the fifth generation (5G)-VANET. This scheme is based on base stations and access points as a local database (LDB) to store the clustering and FCD information. The information gathered from multiple LDBs constitutes the global database(GDB). After that, The SDN controller enables the sharing between the heterogeneous network (HetNets) through the separation of the control and data planes. Nevertheless, this collection scheme raises many challenge. First, the timeliness of FCD is crucial and data must be collected, aggregated and disseminated within a tolerable delay. Last, in urban scenario, the amount of data could be large and the network overload must be considered.

To sum up, the aggregation process in the aforementioned approaches consists of three phases known as decision, fusion and dissemination:

1. The decision phase, where decision regarding the selection of data items to be fused is made

2. The fusion phase, which is related to the function of fusion. Therefore, all the similar data will be merged in one record

3. The dissemination phase, where aggregated data are broadcasted to other vehicles.

Consequently, the existing aggregation schemes have several limitations, but mainly:

- Security: If the aggregation has the ability to decrease bandwidth consumption problems, it might make security issues harder to manage (e.g., the encryption and the decryption of multiple aggregated and compressed packets) [32].

- Scalability: The existing schemes have medium aggregation time as well as low scalability. This is owing to the fact that when the number of the duplicated exchanged messages goes up, the number of collision problems rises as well [27].

- Genericity: Only few approaches [44, 27] propose a generic model for both aggregation and dissemination mechanisms. In fact, such a combination and synchronisation is of paramount importance to avoid the broadcast storm problem [48] and to decrease the network overhead.

- Filtering: The input items are not fully filtered out (many duplicated items and irrelevant items are not neglected) leading to a high level of network overload.

A comparative study is presented in table 2.2 of the existing aggregation methods, while evaluating them according to the security, the scalability, the genericity and the filtering. In the following table we focus only on the V2V Floating Car Data aggregation methods.

Approche	Security	Scalability	Genericity	Filtering
SOTIS	Easy to manage	low	yes	no
TrafficView	Easy to manage	medium	yes	no
CASCADE	Hard to manage because of package compression	medium	yes	no

Table 2.2 Comparative study between the V2V Floating Car Data aggregation methods

2.5 Conclusion

In a VANET vehicular network, vehicles are equipped with on-board sensors that allow them to detect several types of information (ie road accident, empty space in a parking lot, traffic jams, obstacles, work, cut-off road, etc.). This information can be exchanged using wireless communication technologies. This has spawned several VANET applications related to safety, transport efficiency and driver assistance. Indeed, these applications collect the information with the onboard sensors, then treat and disseminate them to other vehicles. As they travel, vehicles exchange a large volume of information. Due to the special features, the highly dynamic topology with intermittent connectivity of the vehicular network, the information has to be exchanged in a very efficient way with a minimal delay while avoiding the saturation of the network and the problems of conflict or collision (Broadcast Storm Problem). To overcome these problems, approaches for dissemination and aggregation of data have been proposed.

In VANET, the dissemination consists in disseminating information via multi-hop communications to interested vehicles by limiting the routing time, ensuring good coverage and avoiding as possible the broadcast storm problem. In another hand, data aggregating is an interesting approach, allowing to integrate multiple information about the same event to generate a summary (or aggregate) to reduce network traffic. It enables vehicles to merge, update, or delete certain information that are duplicated, similar, or expired.

Chapter 3

Proposed data dissemination protocol for VANET

3.1 Introduction

In this chapter, we thoroughly describe our data dissemination protocol DPMS. Indeed, the main originality of DPMS is the formalization of the ZORs through the unveiling of strong connections between the set vehicle trajectories and a set of regions. Thus, the determination of the ZORs can be seen as an instantiation of the cover set problem [33], i.e., finding the minimal coverage of the boolean matrix, in terms of formal concepts, keeping track of the relationship between vehicle trajectories and map regions.

The remainder of this chapter is organized as follows. In Section 3.2, we thoroughly describe the main idea of our dissemination protocol DPMS. In Section 3.3 the simulation settings and the evaluation of the proposed DPMS protocol are presented. Also, a comparison between DPMS and other main geocast protocols is presented here. The last section concludes this chapter.

3.2 DPMS: A swift data Dissemination Protocol based on Map Splitting

In the following, we first define the main concepts and functions used in our approach before detailing the architecture, the map splitting and the dissemination protocol.

Definition 1 ***Road (rd)***: *is a path that connects two junctions. We denote by rd^* the set of junctions connected to the road rd. For example, let rd be a road connecting two junctions j_1 and j_2, then $rd^* = \{j_1, j_2\}$.*

In the sequel, we say that two roads rd_i and rd_j are connected if they share at least one common junction, i.e., $rd_i^* \cap rd_j^* \neq \emptyset$ ■

Definition 2 ***Region (r)***: *is a finite set of connected roads. Formally,*

$$r = \{rd_1, rd_2, ..., rd_n\}.$$

Where:

- *rd_i denotes a road in the region r*
- *$\forall\ rd_i \in r\ \ \exists\ rd_{j \neq i} \in r$ such that $rd_i^* \cap rd_j^* \neq \emptyset$ ■*

Definition 3 ***City (c)***: *is a finite set of connected regions. Formally,*

$$c = \{r_1, r_2, ..., r_n\}.$$

Where:

- *r_i denotes a region in the city c*
- *$\forall\ r_i \in c\ \ \exists\ r_j \in c$ such that $\exists\ rd_p \in r_i\ \ \exists\ rd_{k \neq p} \in r_j$ and $rd_p^* \cap rd_k^* \neq \emptyset$ ■*

The following functions are used in our approach.

Definition 4 ***Zone Of Relevance (ZOR)***:
The zone of relevance of a region r, denoted $ZOR(r)$, is the set of regions having a certain interest regarding events occurring in region r. Formally, let R be a set of regions and $\mathscr{P}(\mathscr{R})$ the powerset of R, we define $ZOR(r)$ as follows:

$$\begin{aligned} ZOR\colon R &\rightarrow \mathscr{P}(\mathscr{R}) \\ r &\mapsto \{r_1, r_2, ..., r_n\}. \end{aligned}$$

where $r_i\ \ (1 \leq i \leq n)$ is a region having a certain interest regarding what happens in region r ■

Definition 5 ***(Split Map)***
The split map of a city c, denoted $SP(c)$, is the set of regions composing the city c with their corresponding ZOR. Formally, let C be a set of cities and R a set of regions, we define $SP(c)$ as follows:

$$SP: C \to R \times \mathscr{P}(\mathscr{R})$$
$$c \mapsto \{(r_1, ZOR(r_1)), \ldots, (r_n, ZOR(r_n))\}.$$

where each couple $(r_i, ZOR(r_i))$ denotes a region in the city c and its corresponding ZOR
■

In the following section, we thoroughly describe the architecture of our protocol as well as its main functions.

3.2.1 Architecture

As depicted in Figure 3.1, the architecture of our proposal, which consists of two tiers of nodes:

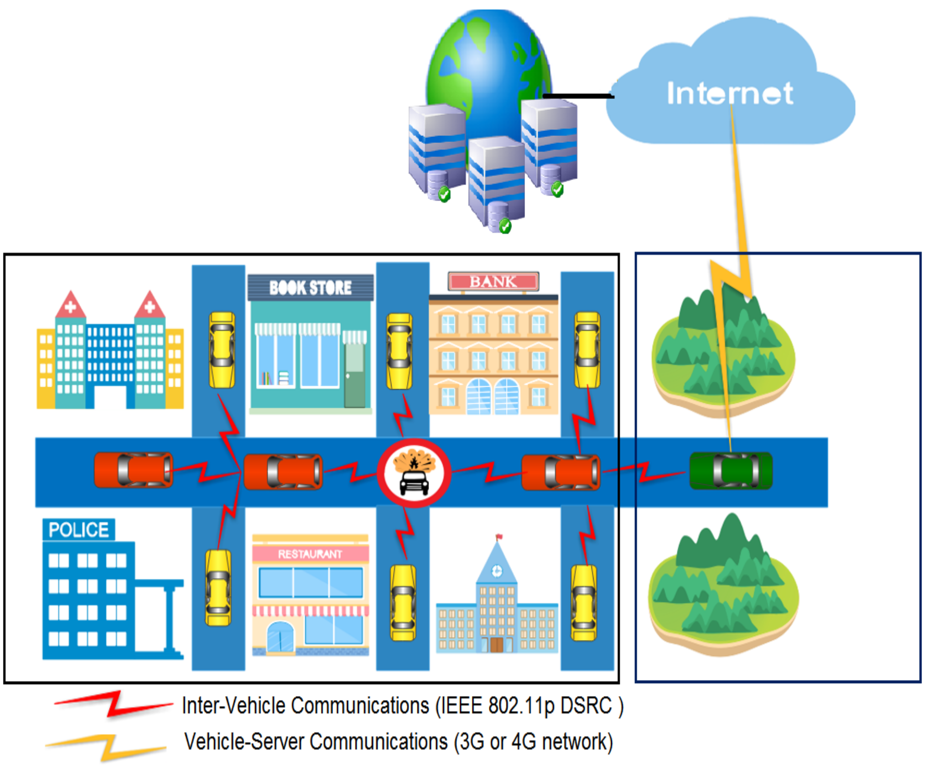


Fig. 3.1 General architecture

- **Server**: It offers a storage and processing services. Indeed, its main function is to split the map of a given city c into a set of regions R, and then to compute the ZOR of each region. The result of map splitting and ZOR computation (i.e., the split map of the city c, $SP(c)$) is then stored in a shared knowledge base. To do this, the server firstly sends a query to vehicles driving in the city c to gather their log files (i.e., each file contains the set of vehicle trajectories, where each trajectory is a set of roads crossed by the vehicle). Thereafter, it runs a dedicated algorithm (as described in subsection 3.2.2) to compute the ZOR of each region in the city based on the collected trajectories. In this study, we assume that updating the map is done every T period (defined manually) from scratch. Of course, this can be done in a more sophisticated way to improve the quality of the ZORs. However, we would like to leave this interesting problem to a dedicated study.

- **Vehicle**: It disseminates events using our DPMS protocol (as described in subsection 3.2.3). In DPMS, the data dissemination process is done only through V2V communications using IEEE 802.11p DSRC technology [24]. However, DPMS requires the split map $SP(c)$, shared by the server, of the city c in which the vehicle is moving. Hence, to avoid vehicle/server communications during the event dissemination time, the vehicle downloads the split map $SP(c)$, via cellular network technology (such as 3G or 4G) whenever it enters in an unvisited city (e.g., the green car in Figure 3.1). In our study, we assume that vehicles are able to determine their respective positions on the road using a Global Positioning System.

3.2.2 Map Splitting and ZOR computation approach

While, in the literature, ZORs are often assumed to be of any form and are still chosen according to the scenarios and motivation needs of the authors, the efficient ZOR detection is a thriving challenge leading to increase the reachability ratio and to overcome the broadcast storm problem. In this respect, we introduce here a new method for the ZOR computation that firstly splits the map of a given city c into a set of regions, then, for each region r_i, it determines its zone of relevance, $ZOR(r_i)$. For example, in Figure 3.2, our method splits the map into 4 regions r_1, r_2, r_3 and r_4. Thereafter, it associates for the events that arose in region r_1 the $ZOR = \{r_1, r_2, r_3, r_4\}$, which closely match the green region that is considered as the ZOR in Figure 1.2.

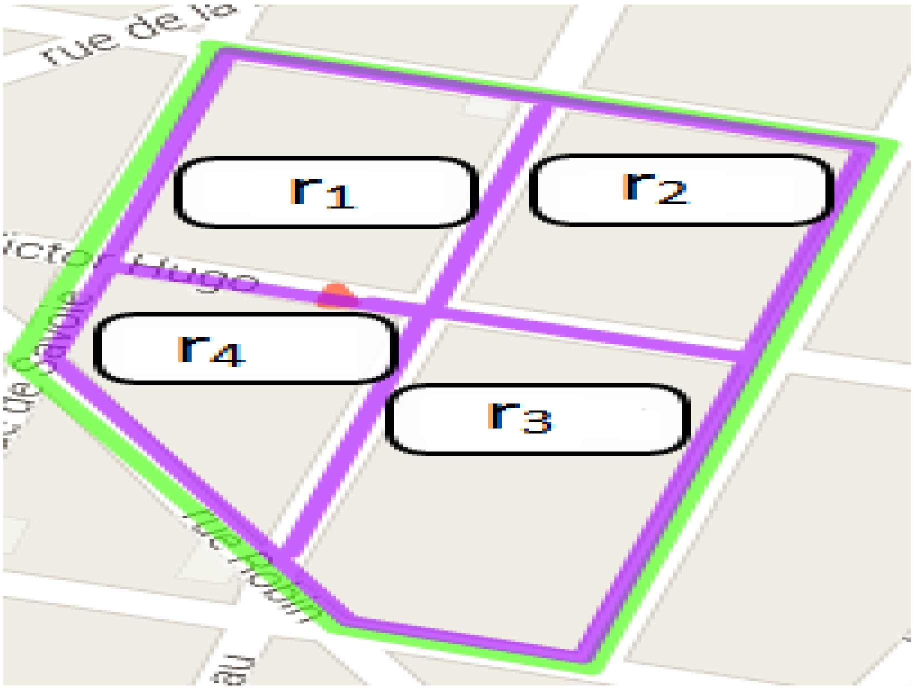


Fig. 3.2 Map Splitting and ZOR detection

To compute the ZOR, our method exploits the links between vehicles' trajectories and regions. To do so, we rely on the Formal Concept Analysis (FCA) [10], which is a method of extracting interesting clusters from relational data. The method is based on a formalization of a philosophical view of conceptual knowledge. The basic notion in the FCA is a formal concept which consists of two sets: extent – a set of all objects sharing the same attributes, and intent – a set of all the shared attributes. The basic input data for the FCA is a table, called a formal context, in which the rows represent objects and the columns represent attributes. The entries of the table contain 'yes/no' information saying whether the corresponding objects has the corresponding attributes or not. One of the main outputs of the FCA is a concept lattice – a hierarchy of formal concepts present in the formal context. The extents and intents of formal concepts are formed by a particular pair of operators induced by the formal context.

In the following, we recall some basic definitions related to FCA, which is a branch of applied applied mathematics. Based on a mathematization of concept and concept hierarchy, it activates mathematical methods for conceptual data analysis and knowledge processing .

Key notions

Definition 6 (FORMAL CONTEXT)
A formal context is a triplet $\mathscr{K} = (\mathscr{V}, \mathscr{R}, \mathscr{I})$, *where* $\mathscr{V}$ *represents a finite set of vehicles'*

trajectories, $\mathscr{R}$ is a finite set of regions (or attributes), and $\mathscr{I}$ is a binary (incidence) relation (i.e., $\mathscr{I} \subseteq \mathscr{V} \times \mathscr{R}$). Each couple $(v, r) \in \mathscr{I}$ expresses that the vehicle trajectory $v \in \mathscr{V}$ is in the region $r \in \mathscr{R}$ ■

Example 1 *Table 3.1 illustrates a formal context, where $\mathscr{V} = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ is a set of vehicle trajectories and $\mathscr{R} = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$ is a set of regions.*

	r_1	r_2	r_3	r_4	r_5	r_6	r_7
v_1	0	0	1	0	1	1	0
v_2	0	0	1	0	1	1	1
v_3	0	0	1	0	1	0	1
v_4	0	0	1	0	0	1	0
v_5	0	0	1	1	0	1	0
v_6	0	1	0	0	0	1	1
v_7	1	1	0	0	0	1	1

Table 3.1 Example of formal context

Worth mentioning, the link between the power-sets $\mathscr{P}(\mathscr{R})$ and $\mathscr{P}(\mathscr{V})$ is defined as follows:

Definition 7 **(GALOIS CONNECTION)**
Let $\mathscr{K} = (\mathscr{V}, \mathscr{R}, \mathscr{I})$ be a formal context. The application ψ associating to the set of vehicle trajectories $V \subseteq \mathscr{V}$ the set of regions $R \subseteq \mathscr{R}$, which are common to all the vehicle trajectories V, is defined as follows:

$$\psi : \mathscr{P}(\mathscr{V}) \to \mathscr{P}(\mathscr{R})$$
$$V \mapsto \psi(V) = \{r \in \mathscr{R} \mid \forall v \in V, (v, r) \in \mathscr{I}\}$$

In a dual way, the application ϕ is defined from the power-set of regions to the power-set of vehicle trajectories as follows:

$$\phi : \mathscr{P}(\mathscr{R}) \to \mathscr{P}(\mathscr{V})$$
$$R \mapsto \phi(R) = \{v \in \mathscr{V} \mid \forall r \in R, (v, r) \in \mathscr{I}\}$$

The coupled applications (ψ, ϕ) form a Galois connection between the power-set of $\mathscr{V}$ and of $\mathscr{R}$ [9] ■

Owing to the definition of these operators, we are able to define a formal concept.

Definition 8 (FORMAL CONCEPT)
A pair $\langle V,R\rangle \in \mathscr{V} \times \mathscr{R}$ of mutually corresponding subsets, i.e., $V = \psi(R)$ and $R = \phi(V)$, is called a formal concept, where V is called extent and R is called intent ■

Roughly speaking, V is the maximal set of vehicles passed through the regions set R.

Example 2 *For example, the formal concept $\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$ shows that $\{v_1v_2v_4v_5\}$ is the maximal set of vehicles passed through the regions r_3 and r_6.*

In the following, we present the pseudo-concept notion, which will be used in the remainder.

Definition 9 (PSEUDO-CONCEPT)
The pseudo-concept associated to the couple (v,r), denoted PC_{vr}, is the union of all the formal concepts containing the couple (v,r). Formally,

$$PC_{vr} = \{(o,i) \mid (o,i) \in \phi(r) \times \psi(v) \subseteq \mathscr{R} \mid o \in \phi(r) \wedge i \in \psi(v)\}$$

■

Example 3 *With respect to the formal context shown by Table 3.1, Figure 3.3 (Up) sketches the pseudo-concept associated to the couple (v_1,r_3) : $PC_{v_1r_3}$=$(\{v_1v_2v_3v_4v_5\},\{r_3r_5r_6\})$. Figure 3.3 (Bottom) shows the four formal concepts extracted from $PC_{v_1r_3}$, i.e., $\langle\{v_1v_2v_3v_4v_5\},\{r_3\}\rangle$, $\langle\{v_1v_2v_3\},\{r_3r_5\}\rangle$, $\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$ and $\langle\{v_1v_2\},\{r_3r_5r_6\}\rangle$.*

The main thrust of the notion of formal concepts stands for the fact that they vehicle a conceptual structure from data.

Such a structure consists of units, which are formal abstractions of concepts of human thoughts allowing meaningful and comprehensible interpretation. In our context, a formal concept conveys the strong connection between a set of vehicle trajectories and a set of regions. For example, from the formal context sketched by Table 3.1, we may extract the following formal concept: $\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$. The latter indicates that four trajectories, out of 7, are supporting connection between the regions r_3 and r_6. Hence, if an event appears in the region r_3, then the vehicles standing within the region r_6 are very likely to be concerned.

Interestingly enough, the computation of ZORs comes down to extracting these formal concepts. Nevertheless, the overwhelming number of formal concepts that may be drawn is a hindrance towards a larger utilization of the FCA. In fact, an interesting tackle of this issue is to find coverage of a formal context by a minimal number of formal concepts [33]. In the following, we present how we compute ZORs, inpired by the work conducted in [33].

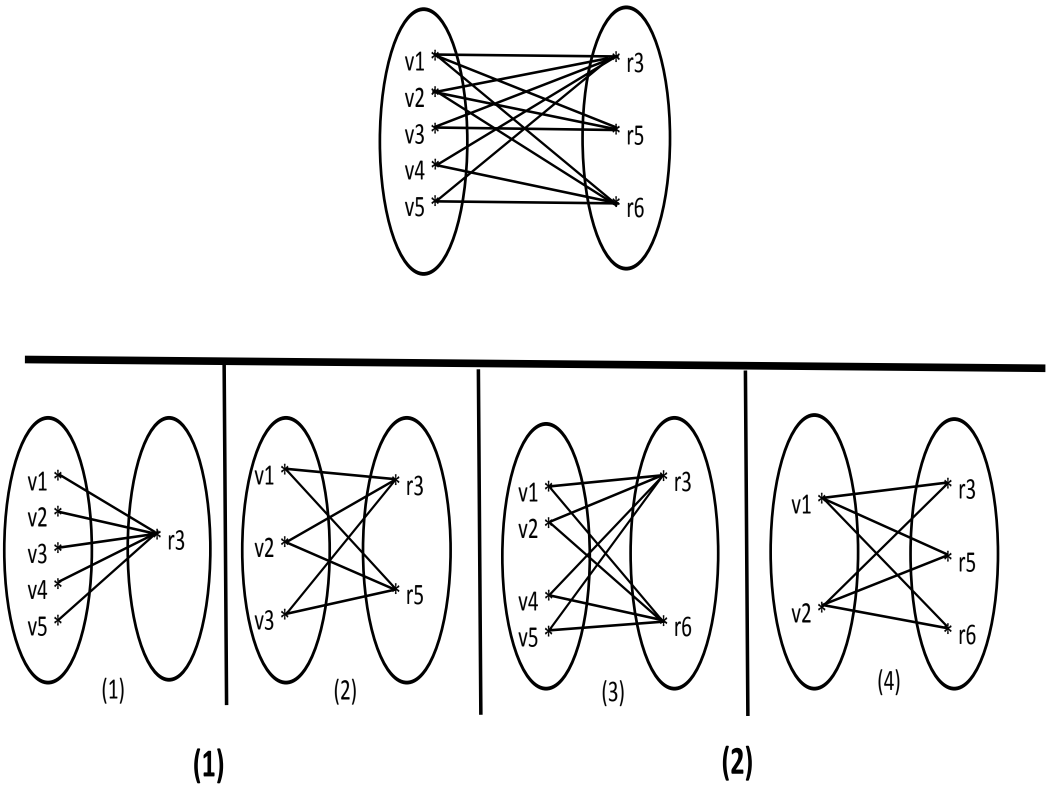


Fig. 3.3 Pseudo-concept

ZOR computation

The algorithm that we introduce is based on a greedy approach, for the extraction of a pertinent coverage of a binary relation. The guiding idea of our approach is to give priority to the formal concept having the highest number of vehicle trajectories, and consequently having a reduced intent part. In the ZORFINDER algorithm, we propose to rely on the following selection metric $\mathscr{M}$, where $|.|$ is the cardinality operator :

$$\mathscr{M}(\langle V,R\rangle) = (|V| \times |R|) - (|V| + |R|)\,. \quad (3.1)$$

The aim behind this selection metric is to achieve a maximal trade off between the respective cardinalities of the intent and extent parts. Plainly speaking, the larger the intent and extent parts are, the higher the interestingness of the formal concept is. The rationale behind this is that by maximizing the cardinality of the extent part as well as the intent one, we maximize the number of covered couples.

Example 4 *If we consider the selection function 3.1 defined above, the pseudo-concept $PC_{v_1r_3}$ associated to the couple (v_1,r_3) and the four formal concepts extracted from $PC_{v_1r_3}$: **(1)** $\langle\{v_1v_2v_3v_4v_5\},\{r_3\}\rangle$; **(2)** $\langle\{v_1v_2v_3\},\{r_3r_5\}\rangle$; **(3)** $\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$; **(4)** $\langle\{v_1v_2\},\{r_3r_5r_6\}\rangle$, then the most pertinent one is equal to $\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$, since :*

- $\mathscr{M}(\langle\{v_1v_2v_3v_4v_5\},\{r_3\}\rangle) = -1$;
- $\mathscr{M}(\langle\{v_1v_2v_3\},\{r_3r_5\}\rangle)$=1;
- $\mathscr{M}(\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$=2; *and*
- $\mathscr{M}(\langle\{v_1v_2\},\{r_3r_5r_6\}\rangle)$=1.

In this example, the idea is to select the most interesting formal concept including the couple (v_1,r_3). This selection has to be carried out through the set of all formal concepts including the couple (v_1,r_3). With respect to the gain function (3.1), the formal concept $\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$ is retained. The extent part, composed of a maximal set of vehicles, is strongly connected to the intent part, composed of an associated maximal set of regions.

In the following, we introduce a new approach, called ZORFINDER, to build a coverage of appropriate formal concepts. As a greedy approach, to set the coverage works by selecting at each stage the set that covers the greatest number of uncovered elements, the ZORFINDER iteratively sweeps the uncovered couples of the given formal context.

The associated pseudo-concept, PC_{vr}, is computed from the couple (v,r), and then two cases are considered:

1. If PC_{vr} is reduced to a formal concept, then it is considered as appropriate and is added to the coverage;

2. Otherwise, the algorithm proceeds to extract all the formal concepts from PC_{vr}, and computes for each one its gain value. The formal concept that maximizes the selection metric $\mathscr{M}$ is considered as appropriate and is added to the coverage.

Example 5 *If we consider the formal context presented in Table 3.1, the ZOR set denoted $\mathscr{ZOR}_{\mathscr{M}}$ associated to the selection metric $\mathscr{M}$ is composed of the following formal concepts:*

1. $\langle\{v_1v_2v_4v_5\},\{r_3r_6\}\rangle$,

2. $\langle\{v_1v_2\},\{r_3r_5r_6\}\rangle$,

3. $\langle\{v_2v_3\},\{r_3r_5r_7\}\rangle$,

4. $\langle\{v_6v_7\},\{r_2r_6r_7\}\rangle$,

5. $\langle\{v_5\},\{r_3r_4r_7\}\rangle$,

6. $\langle\{v_7\},\{r_1r_2r_6r_7\}\rangle$.

Note that the ZOR associated to a given region r_i is equal to the union of all the intent parts of the formal concepts of $\mathscr{ZOR}_{\mathscr{M}}$, including r_i in their intent respective parts. Formally, $\text{ZOR}(r_i) = \{\cup r_j \mid \langle V,R\rangle \in \mathscr{ZOR}_{\mathscr{M}} \wedge r_j, r_i \in R\}$. For example, we have $ZOR(r_2) = \{r_1r_2r_6r_7\}$. In the following, we present a thorough description of our DPMS protocol.

3.2.3 DPMS protocol

The flowchart diagram in Figure 3.4 depicts the different operations of the DPMS protocol. Indeed, DPMS is permanently listening to three different events. It executes a set of operations whenever one of those events is detected. In what follows, we detail how DPMS handles each event:

1. *Vehicle is moving*: When a vehicle v is moving, DPMS repetitively runs the following operations. It firstly determines the road rd on which v is located using GPS. Thereafter, it checks if v has the split map of the current city c (i.e., c is the city containing the road rd), noted $SP(c)$. When v does not have $SP(c)$, it downloads it from the server and stores it locally. Indeed, vehicles only need to download the split maps of unvisited cities. After that, v must cheek if it enters or not a new region r of the city c. In this

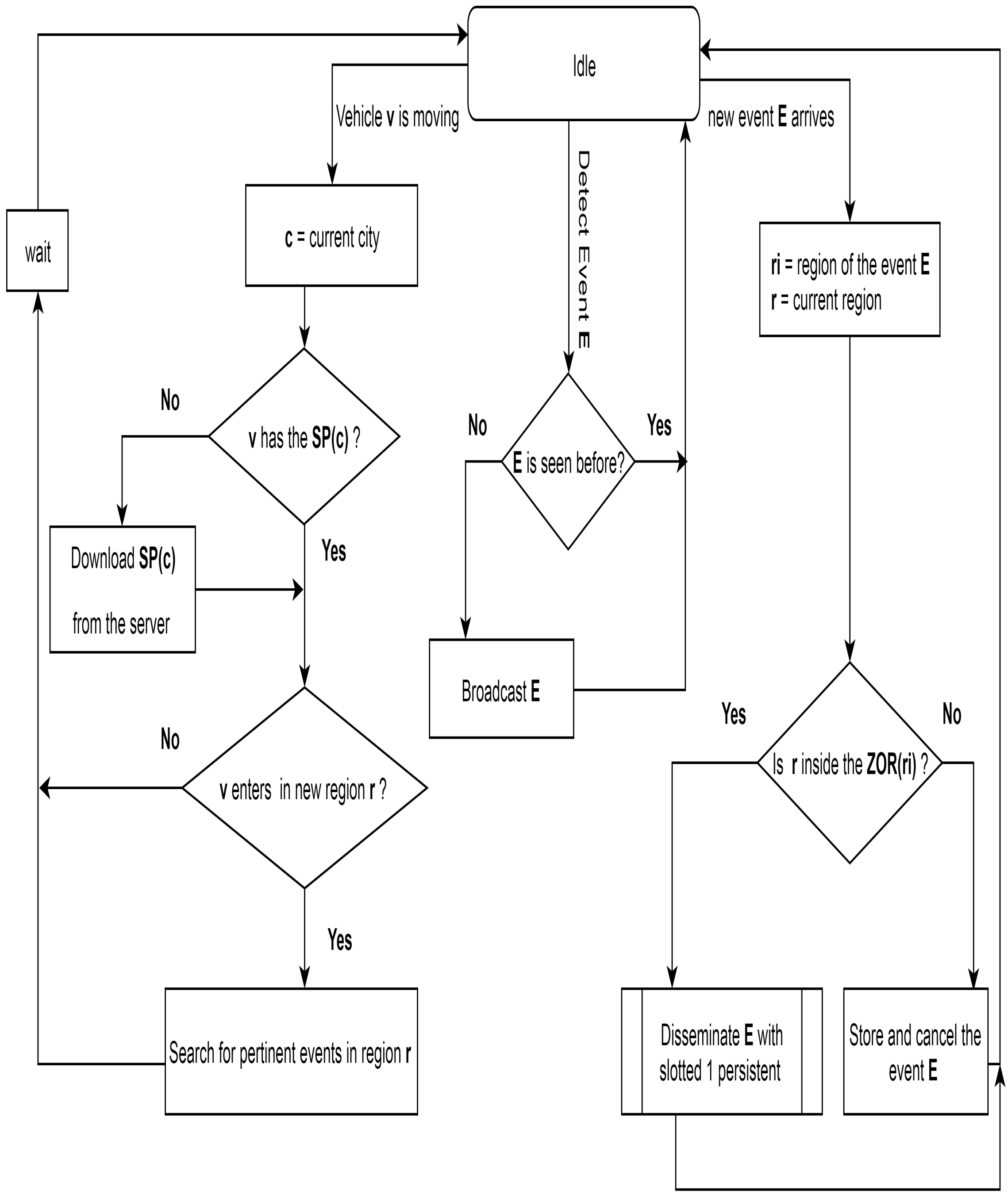


Fig. 3.4 Protocol Flowchart Diagram

case, it runs the *SearchForPertinentEvents* operation in order to retrieve relevant events received by the nearest vehicle in the current region. By doing so, we significantly shorten the time needed to deliver the event to the concerned vehicles. Furthermore, the relevant events are kept alive inside the region, without the need of rebroadcasting event messages as existing geocast approaches do.

2. *Event is detected*: Whenever a vehicle detects a new event, it generates an event message according to the following structure [sender ID, message ID, event region ID, event type, event time][1]. Thereafter, the vehicle broadcasts the generated message to all neighboring vehicles.

3. *New event arrives*: Whenever a vehicle v receives an event message E from another vehicle, it firstly determines the region r_i of the event, i.e., the region in which the event has occurred, from the event message. Thereafter, it checks if the current region r of the vehicle belongs to the zone of relevance $ZOR(r_i)$ of the event region r_i. In this case, the Slotted 1-Persistence suppression technique is used for disseminating the event to vehicles within the ZOR. Otherwise, the event will be simply stored then ignored.

As we mentioned in the related work section, the Slotted 1-Persistence is a timer-based suppression technique. Indeed, using this technique, if a vehicle i receives a packet from a vehicle j, it will firstly calculate a waiting time slot $T_{S_{ij}}$. Then, it will re-broadcast the event if it does not receive any duplicate packet during the waiting time slot; otherwise, it discards it. Given all of the relative distances between the vehicles i and j (D_{ij}), the average transmission range R and the predetermined number of slots Ns, $T_{S_{ij}}$ is computed as follows:

$$T_{S_{ij}} = N_s(1 - [\frac{min(D_{ij}, R)}{R}]) \times \tau \qquad (3.2)$$

where τ is the estimated one-hop delay, which includes the medium access delay and the propagation delay.

[1]The size of the message is less than 2321 Bytes,and the maximum allowed size of the message is 802.11p standard [41]

3.3 Performance evaluation

In this section, we present the performance evaluation of our proposal versus the Slotted 1-Persistence [35]. We choose the Slotted 1-Persistence as a baseline approach thanks to its high reachability ratio [47, 40]. It is important to note that for the sake of providing a fair comparison, we limited broadcasting within the Slotted 1-persistence to vehicles inside a circular region considered as the ZOR of the disseminated event (i.e., circular region is the most efficient shape according to Jochle et al. [22]). This region is defined by a center point p, which is the geographic coordinate of the event location and a radius r.

In addition, we point out that the introduced approach relies on a split map to determine for each region its ZOR. To assess the impact of the region definition on the effectiveness and the efficiency of our DPMS proposal, we study the following scenarios that take into consideration the granularity of the concept of region:

- $DPMS_1$: Each region is defined as a set (or a cluster) of roads where the distance between them is less than $300m$ (i.e., the range of the DSRC protocol [24]) such that the event could reach all vehicles inside the region without rebroadcasting it by the neighboring vehicles.
- $DPMS_2$: Each road is considered as a region.

Simulation settings

The network simulation is performed by OMNeT++ [7] along with the physical layer modeling toolkit MiXiM [2], which makes it possible to employ accurate models for radio interference, as well as shadowing by static and moving obstacles. Added to that, the simulation of Urban Mobility is performed thanks to SUMO [23], which is a microscopic and continuous road traffic simulator. With those two well-established simulators, the nodes simulated by OMNeT++ can interact with SUMO to simulate the influence of the IVC on road traffic and mobility. In the remainder, we take advantage of these two simulators included in the Veins simulation framework [3]. It provides realistic models for $802.11p$ DSRC, PHY and MAC layers. The PHY and MAC parameters are defined according to the basic specifications of the **802.11p standard**.

The simulation settings are summarized in Table 3.2. In the MAC layer, we set the transmission power to $40mW$ to achieve approximately $300m$ of interference range. In

[2] *http://mixim.sourceforge.net/*

[3] Veins is an open source simulation framework for Inter-Vehicular Communication (IVC) that combines both event-based network micro-simulation model as well as road traffic simulator. It is available at *http://veins.car2x.org/*

addition, we vary the amount of vehicles driving on our map from 100 to 1000, ranging from low traffic usually occurring during night times and higher traffic in the afternoon.

Frequency band	5.9 GHz
Transmission power	40 mW
Transmission range	300m
Bandwith	10 MHz
Slot time	13 us
Slot number	5
Average vehicle's speed	80 km/h
Number of vehicles	100 - 1000
Density of vehicles	20 - 200 vehicles $/km^2$
Data message size	2313 bytes
Data message frequency	0.5 Hz

Table 3.2 Simulation settings

3.3.1 Scenario

For our performance evaluation, we select real-world road topologies from three cities at the governorate of Bizerte in the north of Tunisia. We consider the three road topologies depicted in Figure 4.5, representing portions of the urban areas of Corniche, Zarzouna and Menzel Bourguiba cities. Table 3.3 summarizes the characteristics of each map.

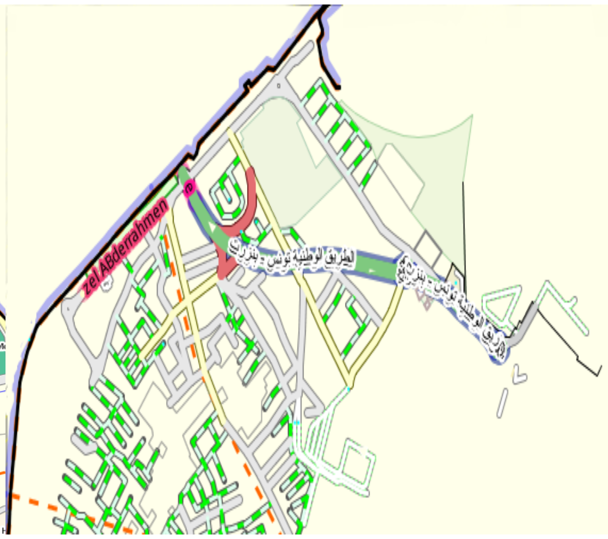
Map	Number of roads	Number of Junctions	Dimensions
Corniche	503	214	2.2km x 2.5 km
Zarzouna	681	312	2.2km x 2.5 km
Menzel Bourguiba	919	468	2.2km x 2.5 km

Table 3.3 Road topologies characteristics

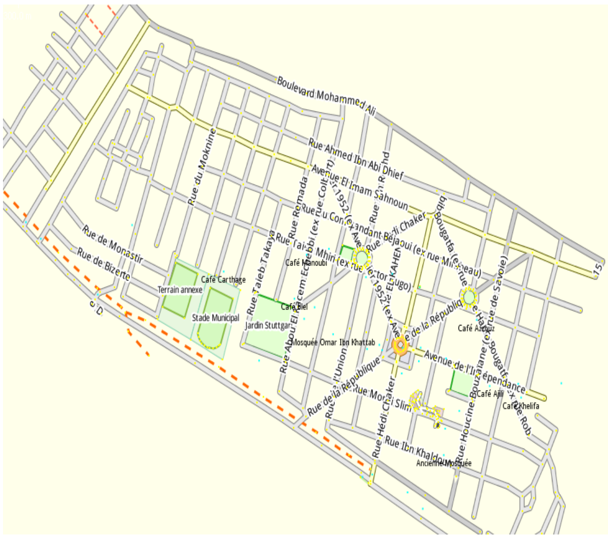
3.3.2 Evaluation metrics

The assessment of the performances of our protocol is carried out through two global metrics, namely the effectiveness and efficiency of the dissemination protocol, which are detailed below.

(a) Corniche



(b) Zarzouna



(c) Menzel Bourguiba

Fig. 3.5 Example of different road topologies

Effectiveness assessment

We consider that a dissemination protocol is effective whenever it guarantees a high geocast **Reachability** and a high **Precision**.

Reachability: It means the average delivery ratio of dissemination, where the message must reach all intersected vehicles of such an event *e* (called accuracy of message transport in[22]). Formally, the reachability metric is defined as follows:

$$Reachability(e) = \frac{|IIV|}{|IV|} \tag{3.3}$$

where *IIV* stands for the set of concerned informed vehicles, *i.e.*, only relevant vehicles for an event *e*, and *IV* stands for the set of concerned vehicles in an event *e*. The average reachability is defined as follows:

$$AverageReachability = \frac{\sum Reachability(e))}{NumberOfEvents} \tag{3.4}$$

Precision: This metric assesses to what extent the protocol is able to only inform relevant vehicles that are actually concerned by a given event *e*. Hence, the challenge will be to obtain higher values of geocasting which is in a close connection with the quality of computing the geocasting area [4]. Formally, the precision metric is defined as follows:

$$Precision = \frac{|IIV|}{|AIV|} \tag{3.5}$$

where *IIV* stands for the set of concerned informed vehicles, i.e. only relevant vehicles for an event *e*, and *AIV* stands for the set of all informed vehicles, i.e. relevant as well as irrelevant vehicles for an event *e*. The average precision is defined as follows:

$$AveragePrecision = \frac{\sum Precision(e))}{NumberOfEvents} \tag{3.6}$$

F-score: The F-score is often used in the fields of information retrieval, machine learning. In our case, we define the F-score as the harmonic mean of precision and reachability, *i.e.*,

$$F - score = \frac{2 \times (Precision \times Reachability)}{Precision + Reachability} \tag{3.7}$$

Efficiency assessment

We consider that the dissemination protocol is efficient whenever it flags out a minimum network **Overload**, a minimum network **Latency**, and a minimum **Packet Loss**. These metrics are explained here:

[4]Geocast is a special case of multicast where data should be only disseminated to a special geographic area.

Overload: The overload metric stands for the total number of sent packets. Interestingly enough, the ultimate goal of any dissemination protocol is to avoid the overload problem [47] by looking for minimizing the number of message transmissions in the network.

Latency: it refers to the amount of time which is needed to deliver a message to an interested vehicle. The average latency, *AL*, is defined as follows:

$$AL = \frac{\sum(t_i - T)}{NumberOfInterestedVehicles} \tag{3.8}$$

where t_i stands for the arrival time of the event message to a vehicle i and T is the time stamp of the occurrence of the event.

Packet Loss: It refers to the number of lost packets versus that of sent packets.

3.3.3 Results

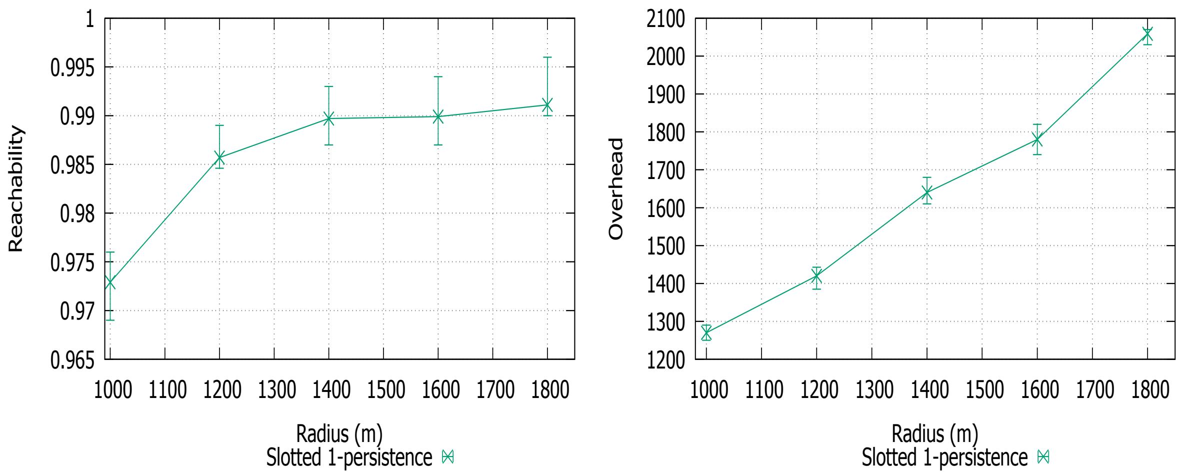
Effectiveness and Efficiency of Slotted 1-persistence under different values of ZOR radius

In order to define the best ZOR radius of the baseline protocol, we compute its reachability and overload according to different values of the ZOR radius.

Figure 3.6, with the map of Menzel Burguiba city, depicts that the reachability and the overload of Slotted 1-persistence grow as far as the radius of the ZOR increases. Indeed, we observe that within a radius of 1400*m* the reachability becomes stable (around 0.99). However, the overload sharply rises as long as the ZOR radius values go up. Consequently, we can deduce that within this value of the ZOR radius, Slotted 1-persistence achieves a maximal trade-off between effectiveness and efficiency. Hence, in the remainder, we set the ZOR radius of Slotted 1-persistence to 1400*m* and we compared its effectiveness and efficiency against our protocol.

Effectiveness of dissemination protocol

Figures 3.7-3.9 show the evolution of the reachability, the precision, and the F-score under different traffic densities (i.e., number of vehicles in the network) within different road topologies (i.e., maps of Corniche, Zarzouna and Menzel Bourguiba cities). As expected,

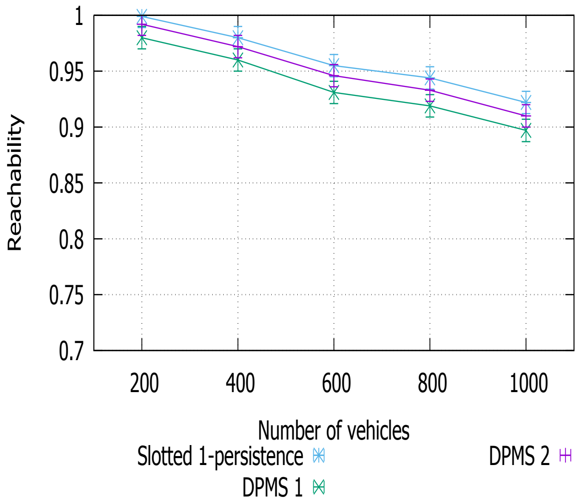


(a) Reachability (b) Overload

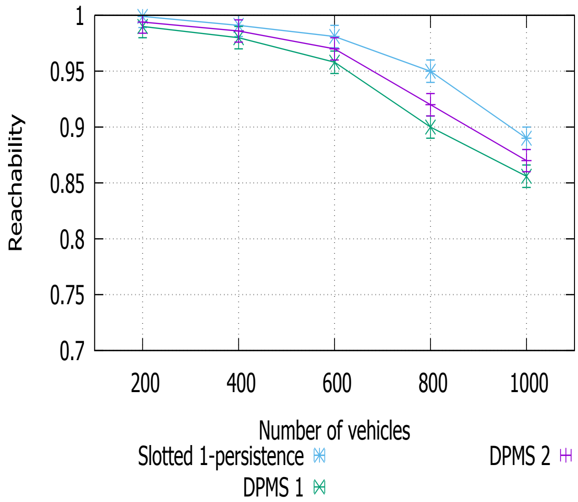
Fig. 3.6 Variation of the average Reachability and Overload of Slotted 1-persistence w.r.t. the variation of ZOR's radius

the values of the different effectiveness metrics decrease for both protocols as far as the number of vehicles increases. Indeed, the higher the number of vehicles is, the lower the probability to reach the interested vehicles. Moreover, we observe that the reachability of the Slotted 1-persistence protocol is slightly sharper than our protocol (Figure 3.7). This is owing to the fact that within the Slotted 1-persistence protocol, the message will be sent nearly to all the vehicles in the network (i.e., interested or not), since the circular region with a radius equal to 1400*m* will include the whole map in most cases. Hence, a higher number of vehicles is notified, increasing consequently the reachability metric. Nevertheless, our protocol overcomes this drawback thanks to a high geocasting precision that only targets interested vehicles and keeps a low overload value. Actually, Figure 3.8 demonstrates that our protocol has a high geocasting precision under different network densities whenever compared to the Slotted 1-persistence. Hence, it increases the overall precision of the Slotted 1-persistence by around 176% (Figure 3.10). This encouraging performances are owed to the fact that our protocol closely matches the ZOR. Therefore, within our ZOR computation method, a less number of uninterested vehicles receive the events leading to a higher precision than the circular ZORs do. In addition, Figure 3.9 flags out that our protocol performs a better trade-off reachability/precision than the baseline protocol does. In fact, our protocol raises the overall F-score of the Slotted 1-persistence by around 161% (Figure 3.10).

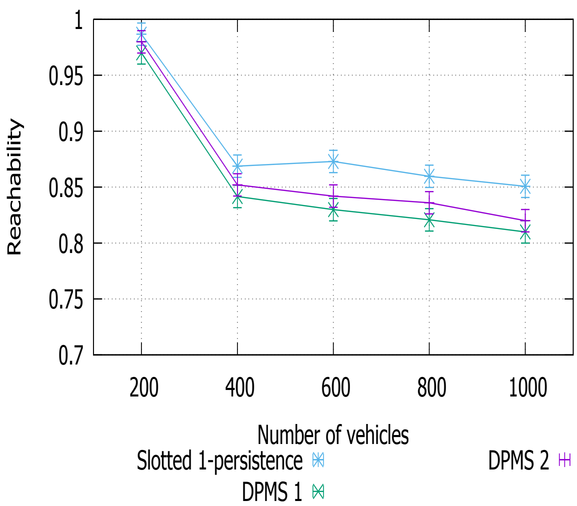
Studying the impact of the road topology on the effectiveness of the dissemination protocols is of paramount importance. Indeed, Figures 3.7-3.9 show that the effectiveness of



(a) Corniche



(b) Zarzouna

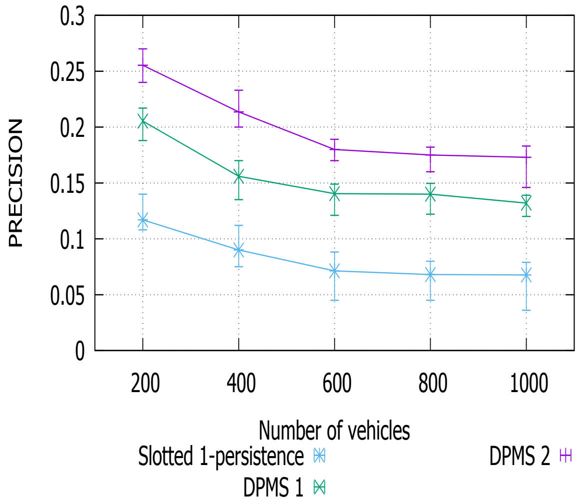


(c) Menzel Bourguiba

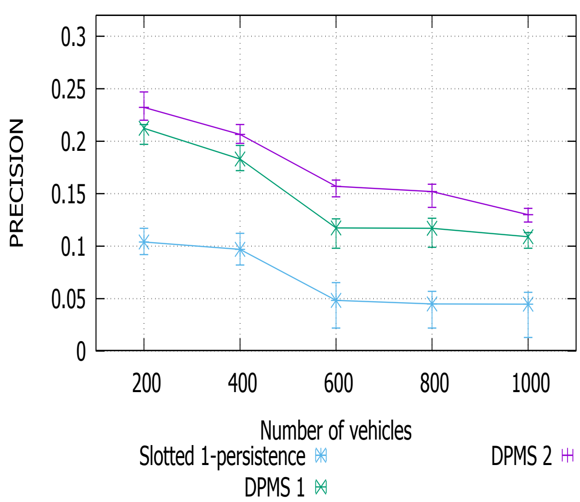
Fig. 3.7 Variation of the average Reachability w.r.t. the variation of the number of vehicles under different maps

the different protocols varies slightly according to the road topology. We observe that the reachability, the precision, and the F-score are better within the map having a less number of roads and junctions. This is due to the fact that in such a city the vehicles usually drive in a smaller area leading to an increase in the effectiveness of the dissemination protocol.

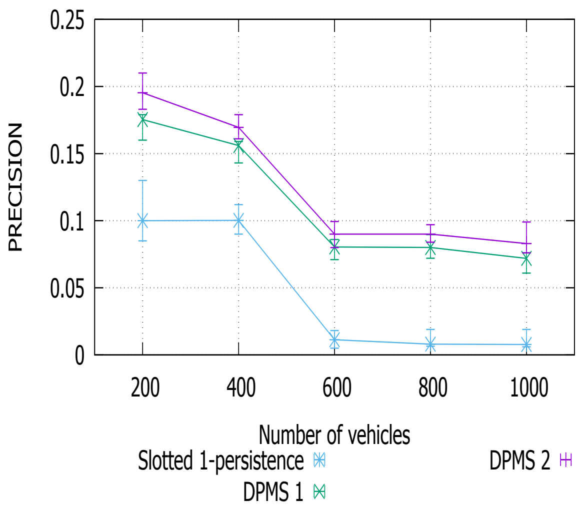
Studying the effect of varying the granularity of the region concept is worth of interest. Indeed, we remark that $DPMS_2$ performs better than $DPMS_1$ in terms of effectiveness under different variations in the number of vehicles. Actually, $DMPS_2$ increases respectively the



(a) Corniche



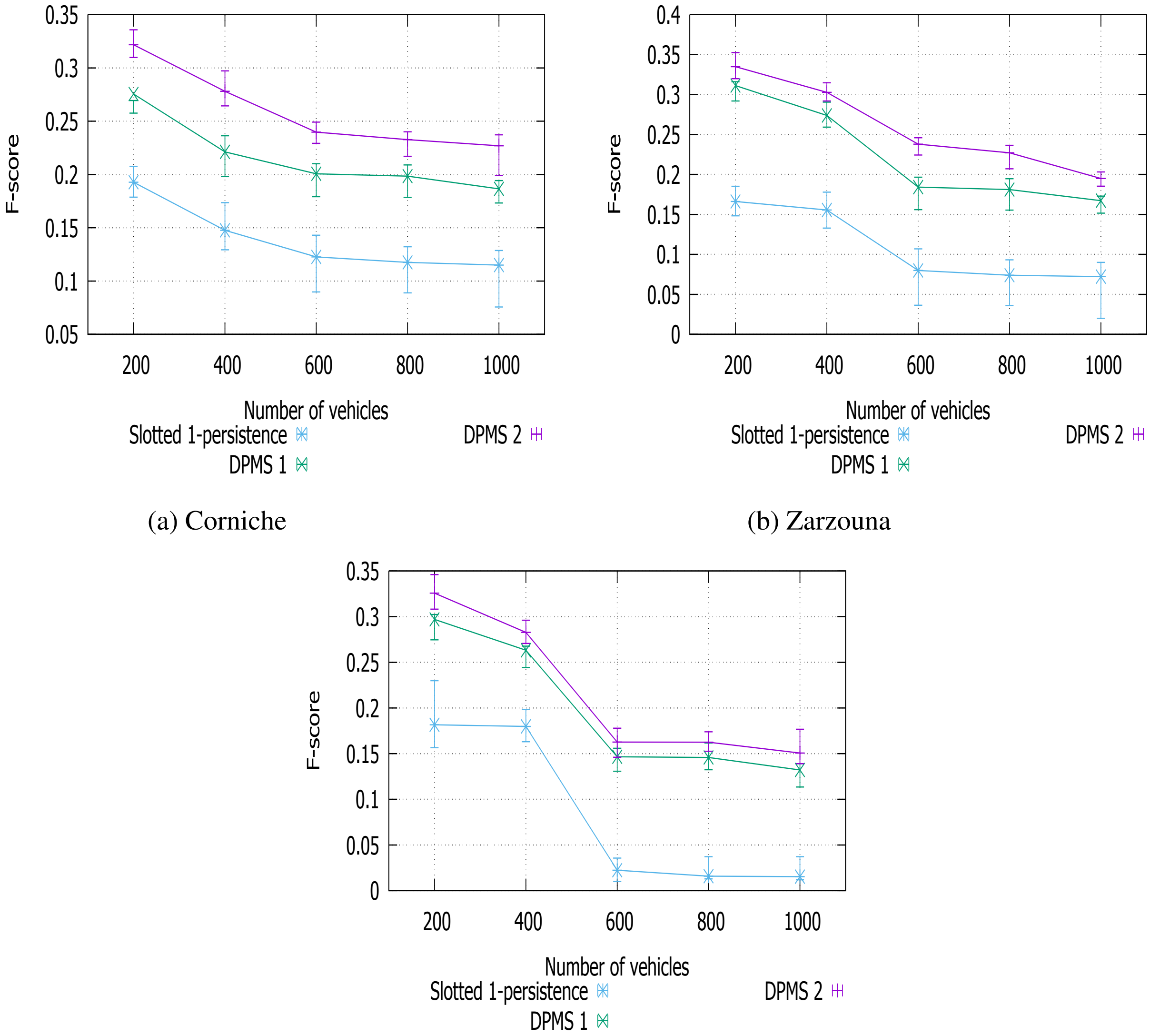
(b) Zarzouna



(c) Menzel Bourguiba

Fig. 3.8 Variation of the average Precision w.r.t. the variation of the number of vehicles under different maps

average reachability, the precision, and the F-score of $DPMS_1$ by around 1.3%, 11% and 10% respectively (Figure 3.10). This is explained by the fact that in $DPMS_2$, we associate for each road its ZOR (i.e., since each region is considered as a road), what leads to a higher precision in bounding the ZOR. Consequently, this fact will undoubtedly increase the effectiveness of the geocast protocol. The downside of this finer granularity stands for a higher complexity of the ZOR computation process giving birth to challenging scalability issues.



(a) Corniche

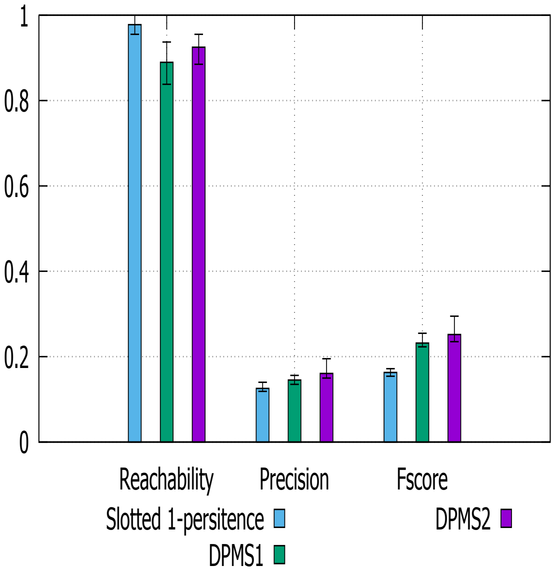
(b) Zarzouna

(c) Menzel Bourguiba

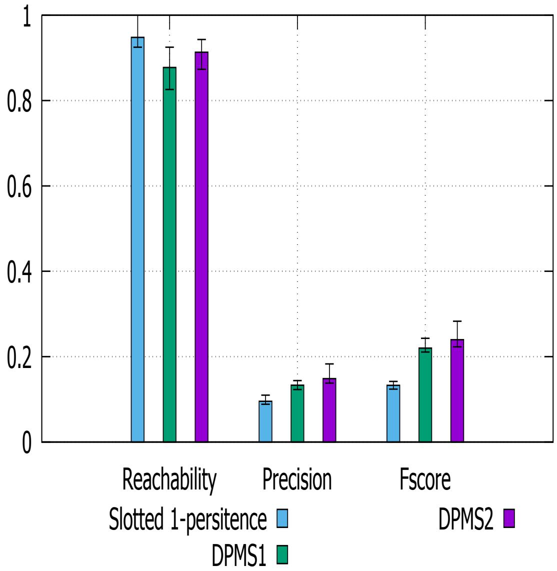
Fig. 3.9 Variation of the F-score w.r.t. the variation of the number of vehicles

Efficiency of dissemination protocol

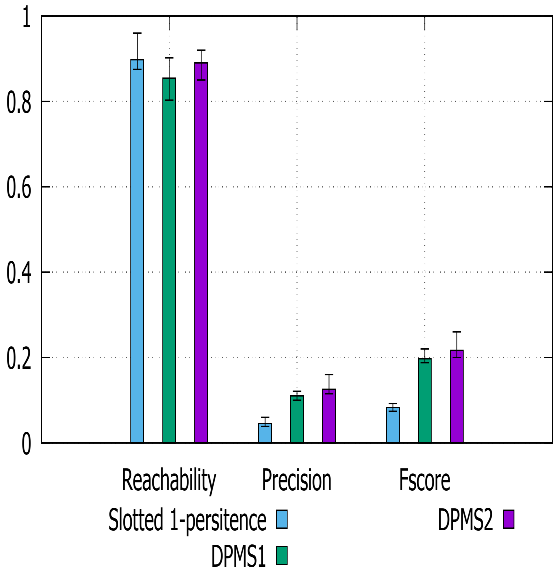
Figures 3.11-3.13 show the evolution of the overload, the latency, and the packet loss of our protocol compared to the Slotted 1-persistence w.r.t. the number of vehicles in the network according to different road topologies (i.e., maps of Corniche, Zarzouna and Menzel Bourguiba cities). We observe that the efficiency of both protocols decreases as far as the number of vehicles in the network increases. Indeed, raising the traffic density drastically increases the number of vehicles broadcasting the messages, which leads to a rise in the overload, the latency, and the packet loss ratio. In addition, both of Figure 3.11 and Figure



(a) Corniche



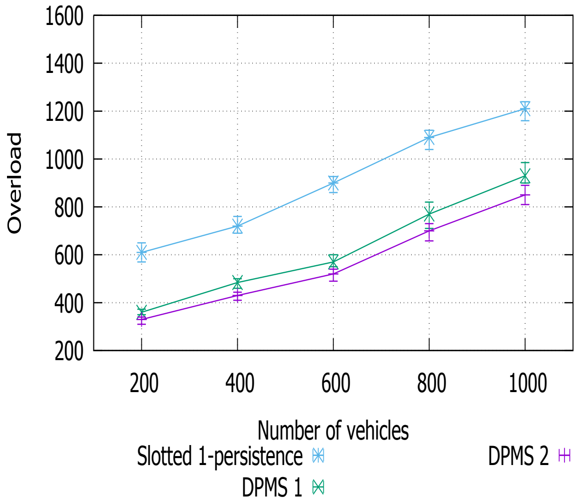
(b) Zarzouna



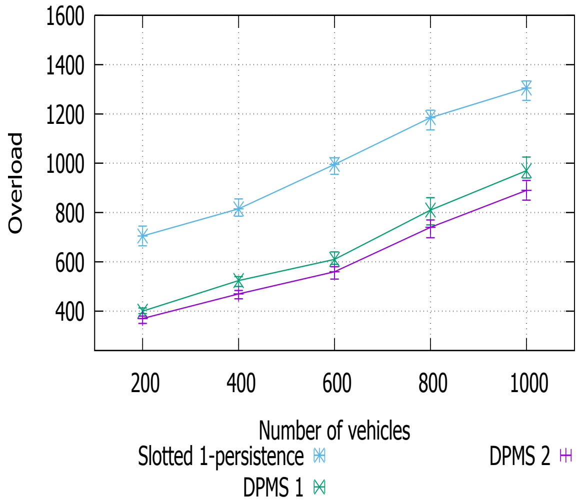
(c) Menzel Bourguiba

Fig. 3.10 Average Reachability, Precision and F-score

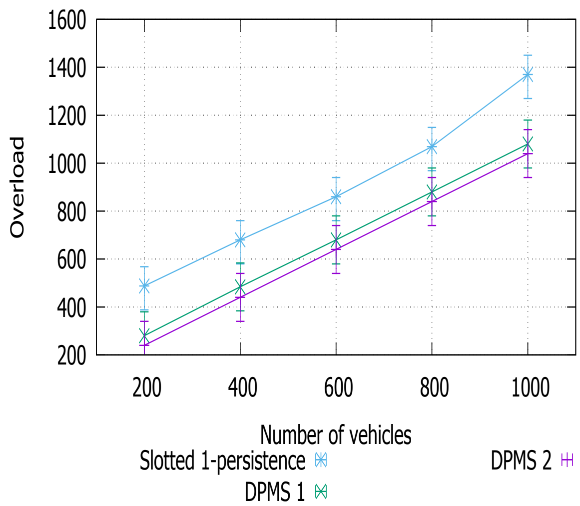
3.12 demonstrate that our protocol flags out a minimum network overload as well as a packet loss than the Slotted-1 persistence protocol does for all the considered variations in the traffic density and within different road topologies. Indeed, our protocol achieves a less 28% overload and respectively a less 43.7% packet loss than its competitor. This is thanks to the fact that within our ZOR computation algorithm, the dissemination area is too narrow than



(a) Corniche



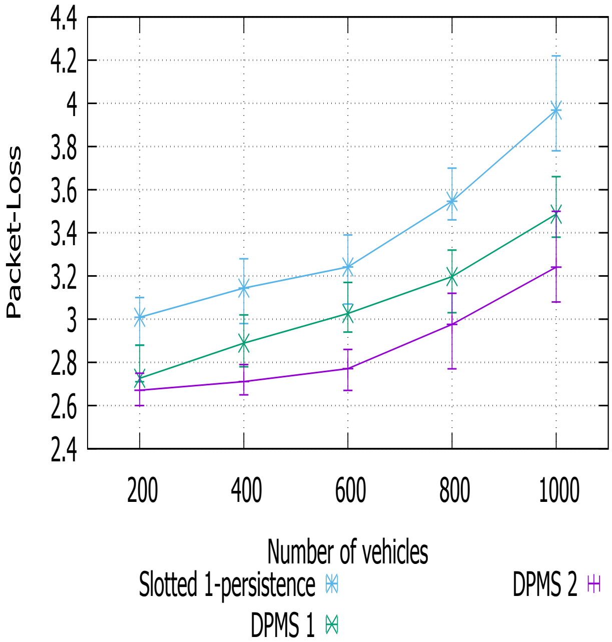
(b) Zarzouna



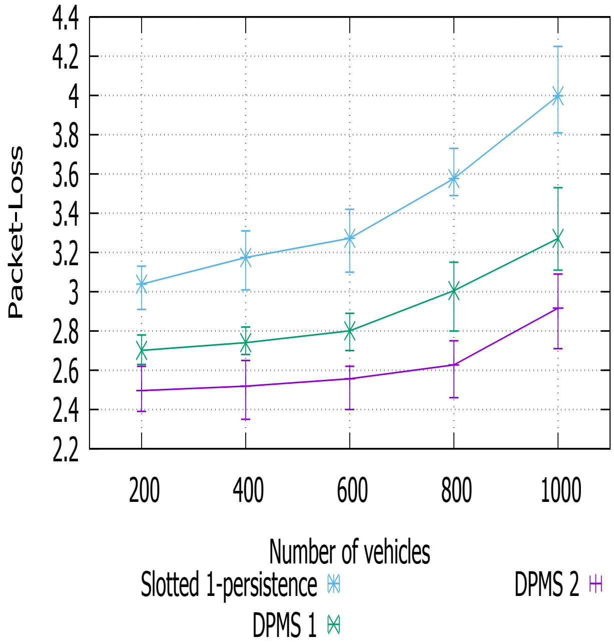
(c) Menzel Bourguiba

Fig. 3.11 Variation of the Overload w.r.t. the variation of the number of vehicles

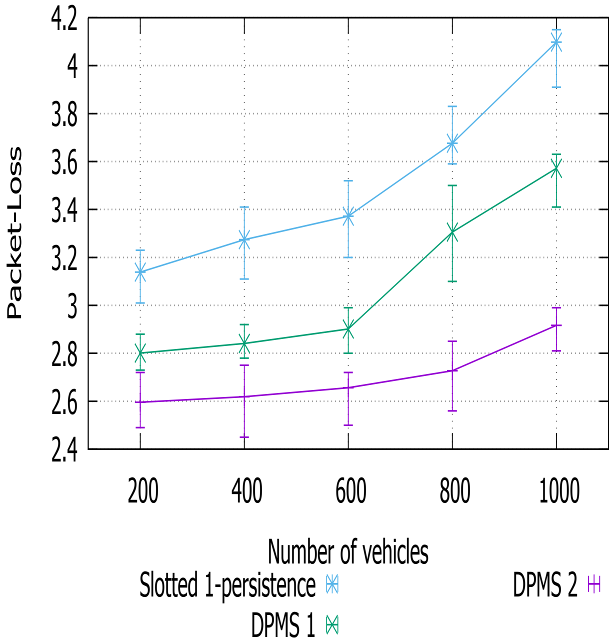
the circular one. Consequently, only a small number of vehicles inside the ZOR disseminate the messages, which helps to significantly decrease the overload and the packet loss ratio. It is also worth mentioning that the latency of our protocol is 10% less than the Slotted 1-persistence protocol. In fact, in our protocol, a vehicle can be informed about relevant events whenever it contacts an informed vehicle inside the ZOR without having the need to rebroadcast event messages like the Slotted 1-persistence does. Additionally, we also remark that there is a slight difference in the efficiency of protocols under different road topologies.



(a) Corniche



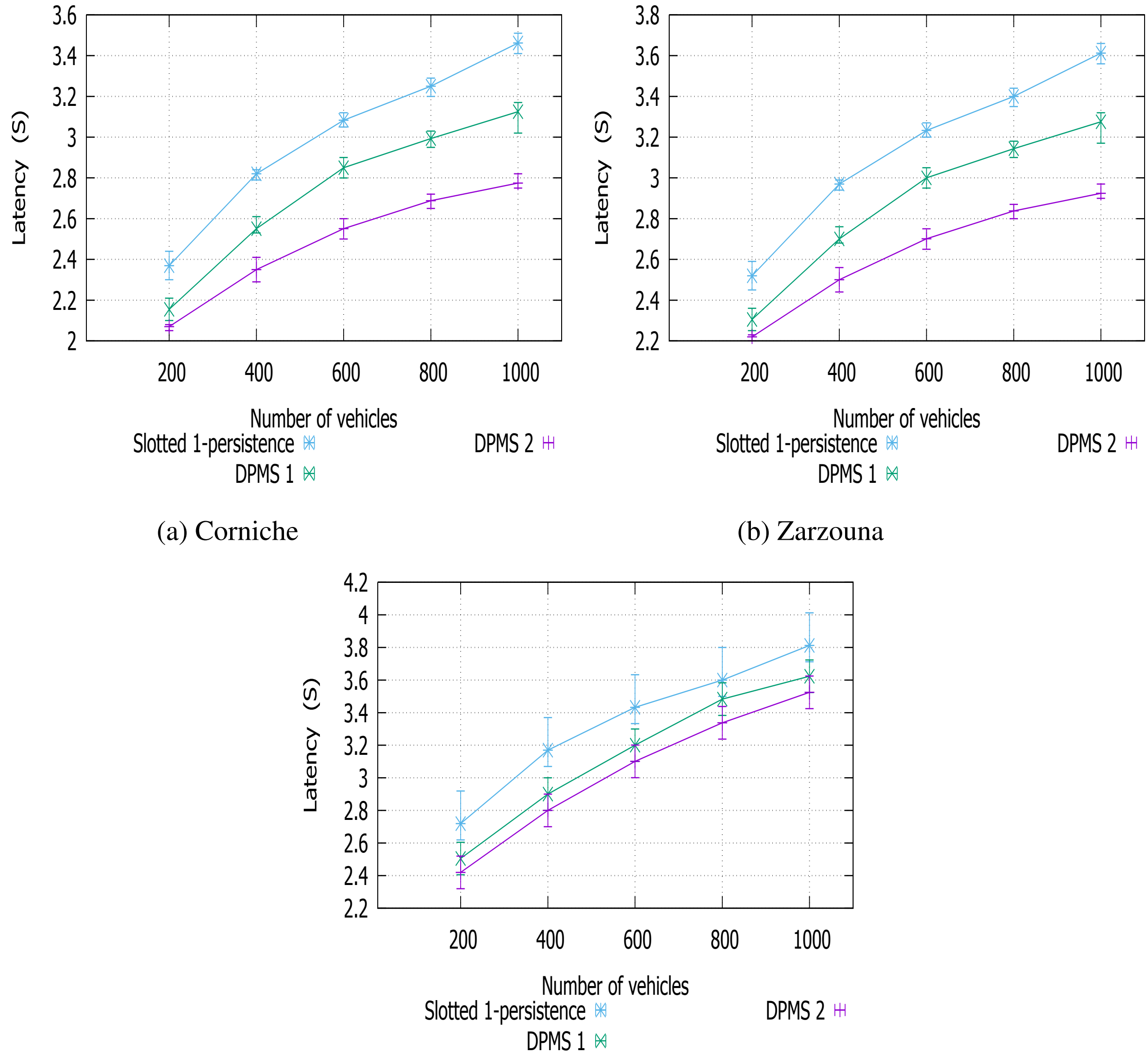
(b) Zarzouna



(c) Menzel Bourguiba

Fig. 3.12 Variation of the Packet loss w.r.t. the variation of the number of vehicles

Finally, Figures 3.11-3.13 depict that by splitting the map into a small region (e.g., a road considered as a region), $DPMS_2$ outperforms $DPMS_1$ in terms of efficiency under different vehicle densities.



(a) Corniche

(b) Zarzouna

(c) Menzel Bourguiba

Fig. 3.13 Variation of the average Latency w.r.t. the variation of the number of vehicles

3.4 Conclusion

In this chapter, we have introduced DPMS as a MAP-splitting-based dissemination protocol to disseminate information about VANET safety events. The main thrust of our protocol stands for an adequate targeting of the zone of relevance of the disseminated messages. Doing so has allowed us to meet our goals, namely reaching a high delivery ratio as well as a high geocast precision. Extensive experimental work has shown that DPMS has obtained very encouraging results versus those got by pioneering approaches of the literature.

Chapter 4

Proposed data aggregation protocol for VANET

4.1 Introduction

A careful scrutiny of existing data aggregation approaches in the literature highlights that they mainly focus on combining the correlated items but none of them tries to use a suppression technique in order to eliminate the duplicated messages. Moreover, they did not take into consideration neither the vehicle directions nor the road speed limitations in the provided traffic information. To cope with these aforementioned requirements, we introduce here a new aggregation protocol called Smart Directional Data Aggregation (SDDA) that can deal properly with traffic information in both urban and highway conditions. It consists in properly selecting relevant FCD messages that should be aggregated. To do so, SDDA provides three filters: (i) The first uses vehicle directions. In fact, each vehicle will aggregate only data that corresponds to its direction but only store, carry and forward uninteresting data. (ii) The second is done through road speed limitations, thus, if the average speed of the received FCD messages is higher than the maximum allowed speed then it will be ignored and replaced by the maximum allowed speed value. Doing so, we increase the aggregation accuracy. (iii) The last is performed is by integrating a suppression technique [48] in order to eliminate the duplicated inputs in the aggregation phase by providing a high accurate time slot based on the distance and the range that all vehicles should wait before rebroadcasting the FCD message.

The rest of the chapter is organized as follows. In Section 4.1, we describe our proposed protocol for smart data aggregation. Section 4.4 is dedicated to the simulation settings as well as the evaluation of our suggested protocol. The last section concludes this chapter.

Before describing our aggregation protocol, we present in the following subsection several definitions and concepts.

4.2 Preliminaries

To better understand our solution, we start by defining the main concepts used in our approach:

Definition 10 ***Lane (l)****: It is a one-way path with a paved surface that connects two spatial points on the map. Formally, l:$\langle$ Id, S, E, $\overrightarrow{SE}$, $Speed_{min}$, $Speed_{max}$, Status, Location $\rangle$ where:*

- *Id: represents the identifier of the lane*
- *S and E respectively represent the start and end points which the lane is connecting. Each point is represented by spacial coordinates (e.g., (x,y))*
- *$\overrightarrow{SE}$: represents the directed vector segment from the point S to E*
- *$Speed_{min}$ and $Speed_{max}$ represent the minimum and maximum speed limits , respectively*
- *Status: indicates the lane situation (e.g., closed, open, restricted, etc.)*
- *Location: is the geographical coordinates of the lane in the map.* ⧫

Definition 11 ***Road (r)****: It contains at least one or several lanes having the same or different directions. Formally, r: $\langle$ Id, L, Type, Name, Network Coverage $\rangle$ where:*

- *Id: represents the road identifier*
- *L: contains the set of lanes included in the road*
- *Type: is used to indicate if it is a highway, urban, street, etc.*
- *Name: is used to describe the road (e.g., street name)*
- *Network Coverage: indicates the types of network covered within the road (e.g., GSM, WIFI, 3G, 4G, etc.)* ⧫

Definition 12 ***Vehicle (v)****: A vehicle is defined in our approach as follows: v:⟨ Id, Driving, $Speed_{max}$, Positioning System, Brand, Type, Size, Environment, DSRC Range, Destination ⟩ where:*

- *Id: represents the vehicle identifier*
- *Driving: is the set of the driving settings (e.g., Preferred path, Deriving mode: Economic/sport, etc.)*
- *$Speed_{max}$: represents the vehicle maximum speed*
- *Positioning System: refers to the geographical GIS-system used by the car (e.g., Google Maps, OpenStreetMaps, Bing Maps, etc.)*
- *Brand: refers to the automobile manufacturer of the vehicle (e.g., BMW, Jeep, Toyota, etc.)*
- *Type: is used to indicate the type of the car (e.g., Mini-Van, Sport, Light Truck, etc.)*
- *Size: indicates the size of the vehicle*
- *Environment: refers to the vehicle speed, weather, and the geographic location*
- *DSRC Range: is the signal power of dedicated wireless short-range communication technology (e.g., 300m, 400m, etc.)*
- *Destination: is the location to which a vehicle travels.*

A vehicle can perform three actions:

- *Broadcast: It disseminates FCD messages using the suppression broadcasting technique defined in [48]*
- *StoreCarryandForward: A vehicle can store, carry and then rebroadcast the same message using the rebroadcasting technique defined in [48]*
- *CalculateAverageSpeed: It calculates the average speed using the aggregation function defined in section 4.3*
- *Receive: It receives all types of messages sent using the DSRC protocol [12]*
- *CalculateDirection: A vehicle is able to determinate its direction using a positioning system (e.g., Google Maps) based on its location and its destination*

- *LocateLane: A vehicle is able to locate the current driving lane based on its geographic location.* ♦

In our solution, we assume that our aggregation protocol runs on the top of the MAC-layer, so it requires no modification in IEEE 802.11p standard [12]. Also, we acquire that only one type of messages can be generated and sent, namely the Floating Car Data (FCD) known also as Floating Cellular Data [46].

Definition 13 ***FCD message (f)****: The adopted FCD messages header structure is therefore defined as a 4-tuple: f: ⟨ SenderId, SenderPosition, AverageSpeed, Destination ⟩ where:*

- *SenderId: is the unique identifier of vehicle that sends the message*
- *SenderPosition: contains the spatial position of the sender*
- *AverageSpeed: is the average speed (computed using a function defined in Section 4.3)*
- *Destination: contains the future location of the vehicle sending the FCD message.* ♦

The size of the message is less than $2,321$ Bytes, which is the maximal allowed size as defined by 802.11p standard [12].

4.3 Aggregation protocol

Our aggregation algorithm deals with the following three scenarios: a unidirectional road, a bidirectional road, and an urban scenario. The aggregation function used in our protocol is the same average function used in SOTIS [46]. We choose the main average speed aggregation function of SOTIS since our contribution can be seen as an optimisation of SOTIS. However, as mentioned in the related work, only three approaces, SOTIS [46], TrafficView [34] and CASCADE [18], focus only on the aggregation of the vehicles speed without combining other driver preferences. The average speed aggregation function of SOTIS is defined as follows [46]:

$$\hat{V}_{r,new} = \hat{V}_{r,prev} + \hat{V}_r \tag{4.1}$$

where $\hat{V}_{r,new}$ is the new average speed for the road r, $\hat{V}_{r,prev}$ is its previous average speed, and $\hat{V}_r$ stands for the average speed of the vehicles on the road r. Each vehicle has three aggregation cases: unidirectional-road, bidirectional-road and urban-city. Algorithm 1 depicts the behavior of a vehicle v upon receipt of an FCD message.

Algorithm 1 Aggregation algorithm of a vehicle v

VehilceLocation ← v.Environment.Location
switch *VehilceLocation* **do**
  case *Unidirectional Road*
    Use Algorithm 2
  case *Bidirectional Road*
    Use Algorithm 3
  case *Urban City*
    Use Algorithm 4

4.3.1 Unidirectional road case

On an unidirectional road, all vehicles drive in the same direction. Indeed, vehicles ahead must collect traffic information, aggregate and disseminate it to other vehicles located behind. Figure 4.1 depicts the unidirectional road scenario where lanes contain different speed limitations. Algorithm 2 contains the pseudo-code of processing this case. Briefly, when a

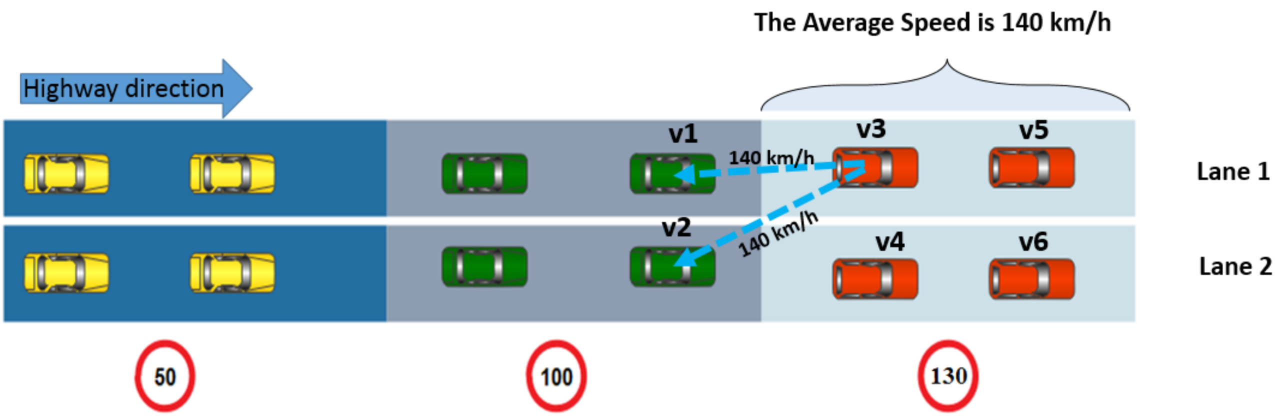


Fig. 4.1 Unidirectional Road Case

vehicle v receives an FCD message, it checks if the message comes from a farther sender on the same road or not. If yes, it computes the new updated average speed $\hat{V}_{r,new}$ based on Equation 4.1. If the received average speed is greater than the maximum allowed speed of the current road r, the vehicle receiving the message will keep its previous average speed $\hat{V}_{r,prev}$ and rebroadcast the value of the road maximum speed. Doing so, the propagated FCD message will warn other vehicles' drivers whose ahead vehicles are driving very fast and exceed the maximum allowed speed. In Figure 4.1, only vehicles v_1 and v_2 will receive and deal with the FCD message sent form vehicle v_3 since they are in the broadcast range and drive behind it.

Algorithm 2 Aggregation algorithm in unidirectional road with speed limitations

```
P ← v.Environment.Location
VD ← v.Destination
VDirection ← v.CalculateDirection(P,VD)
VSpeed ← v.Environment.Speed
LaneSpeedMax ← v.LocateLane.SpeedMax
procedure VEHICLE ON UNIDIRECTIONAL ROAD(fcd)
    PS ← fcd.SenderPosition
    DS ← fcd.Destination
    FAD ← fcd.getAverageSpeed()
    SenderDirection ← v.CalculateDirection(PS,DS)
    SenderDirection ← v.CalculateDirection(PS,DS)
    if Same(SenderDirection, VDirection) and VSpeed ≤ SenderSpeed then
        if SenderSpeed ≤ LaneSpeedMax then
            AverageSpeed ← v.CalculateAverageSpeed(VSpeed, FAD)
            v.Broadcast (AverageSpeed)
        else
            v.Broadcast (LaneSpeedMax)
        end if
    end if
end procedure
```

However, vehicles v_1 and v_2 will ignore the received average speed, 140 km/h, and broadcast the road maximum allowed speed (of 130 km/h). Doing so, all the disseminated traffic information will follow the legal speed.

4.3.2 Bidirectional road case

On a bidirectional road, vehicles drive in two opposite directions. Indeed, vehicles in the opposite directions must collect traffic information, and then aggregate and disseminate it to other opposite vehicles, as shown in Figure 4.3. The exchange of traffic information between the two opposite directions has a paramount importance. In essence, such an exchange would inform and warn all drivers about the traffic conditions ahead, which leads to avoid traffic jam and road accidents. Figure 4.2 illustrates the bidirectional road scenario where lanes contain different speed limitations. Indeed, vehicle v_4 will share its FCD average speed (90 km/h) with vehicle v_1, which will broadcast this information to other behind vehicles since the average speed is less than the maximum allowed speed. Algorithm 3 contains the pseudo-code of processing this case. In fact, the main difference between this case and the previous one is that in a bidirectional scenario, vehicles can accept FCD messages that come form the opposite side or vehicles ahead in the same lane and ignore other messages that come from behind. Actually, the vehicles in the opposite side have a larger overview on

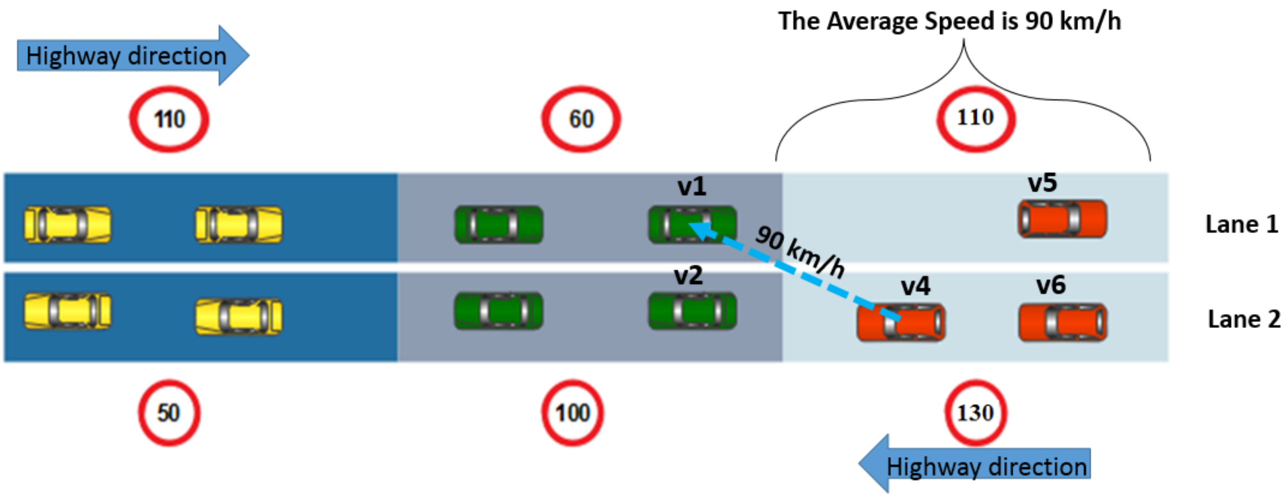


Fig. 4.2 Bidirectional road case

the opposite traffic conditions, since they have been passed in front of it. Doing that allows guaranteeing that the FCD contains an aggregated value of the whole opposite lanes. Figure 4.3 depicts the propagation of an FCD message on a bidirectional road.

4.3.3 Urban network case

In urban network, vehicles drive in many different directions. Thus, vehicles can meet over cross roads and junctions. Indeed, they must exchange their traffic information, as shown in Figure 4.4. However, using a blind aggregation method will affect the precision and the accuracy of the exchanged traffic information. This is owning to the fact that vehicles can share and aggregate the average speed of other vehicles that are not going to the same direction. Our aggregation model can solve this issue based on three filters:

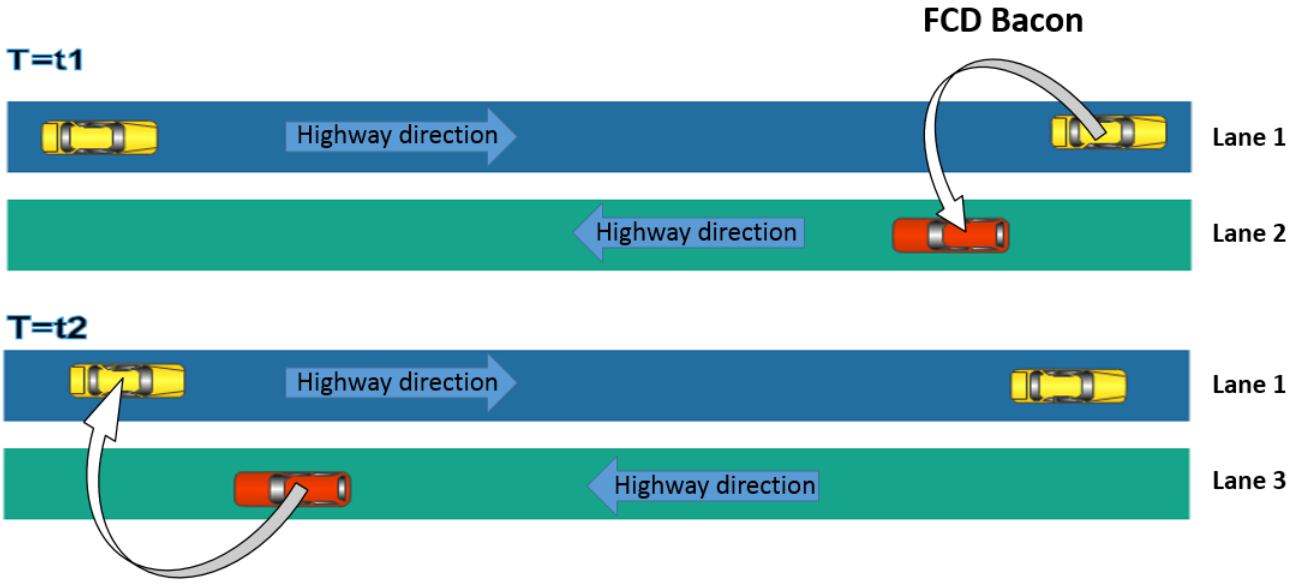


Fig. 4.3 FCD message propagation on a bidirectional road

Algorithm 3 Aggregation algorithm on a bidirectional road with speed limitations

$P \leftarrow$ v.Environment.Location
$VD \leftarrow$ v.Destination
$VDirection \leftarrow$ v.CalculateDirection(P,VD)
$VSpeed \leftarrow$ v.Environment.Speed
$LaneSpeedMax \leftarrow$ v.LocateLane().SpeedMax
procedure VEHICLE ON BIDIRECTIONAL ROAD(fcd)
    $PS \leftarrow$ fcd.SenderPosition
    $DS \leftarrow$ fcd.Destination
    $FAD \leftarrow$ fcd.getAverageSpeed()
    $SenderDirection \leftarrow$ v.CalculateDirection(PS,DS)
    if $NOT\ Same(SenderDirection\ ,\ VehicleDirection)$ **then**
        if $SenderSpeed \leq LaneSpeedMax$ **then**
            $AverageSpeed \leftarrow$ v.CalculateAverageSpeed(VSpeed,FAD)
            v.Broadcast *(AverageSpeed)*
        else
            v.Broadcast *(LaneSpeedMax)*
        end if
    end if
end procedure

Algorithm 4 Aggregation algorithm in urban network

```
P ← v.Environment.Location
VD ← v.Destination
VDirection ← v.CalculateDirection(P,VD)
VSpeed ← v.Environment.Speed
LaneSpeedMax ← v.LocateLane().SpeedMax
procedure VEHICLE IN URBAN CITY(fcd)
    PS ← fcd.SenderPosition
    DS ← fcd.Destination
    FAD ← fcd.getAverageSpeed()
    SenderDirection ← v.CalculateDirection(PS,DS)
    if Same(SenderDirection , VDirection) then
        if SenderSpeed ≤ LaneSpeedMax then
            AverageSpeed ← v.CalculateAverageSpeed(VSpeed, FAD)
            v.Broadcast (AverageSpeed)
        else
            v.Broadcast (LaneSpeedMax)
        end if
    else
        StoreCarryandForward (fcd)
    end if
end procedure
```

1. A direction filter that ensures the aggregation of coherent FCD messages targeting the same road and direction

2. A suppression technique filter that ignores all duplicated FCD messages. In the current version of our protocol, we adopt the famous slotted 1-persistence broadcast protocol [48]

3. An aggregation filter that ignores all the received FCD messages that exceed the road maximum speed. In fact, if the average speed of any FCD message is greater than the maximum allowed speed, the received vehicle will ignore it and broadcast the maximum speed instead.

In urban networks, when a vehicle gets an FCD message, it checks firstly the direction of the sender. Then if the received average speed is less than the road maximum speed limit, the vehicle will aggregate and disseminate the received FCD message. Otherwise, the received message will be ignored.

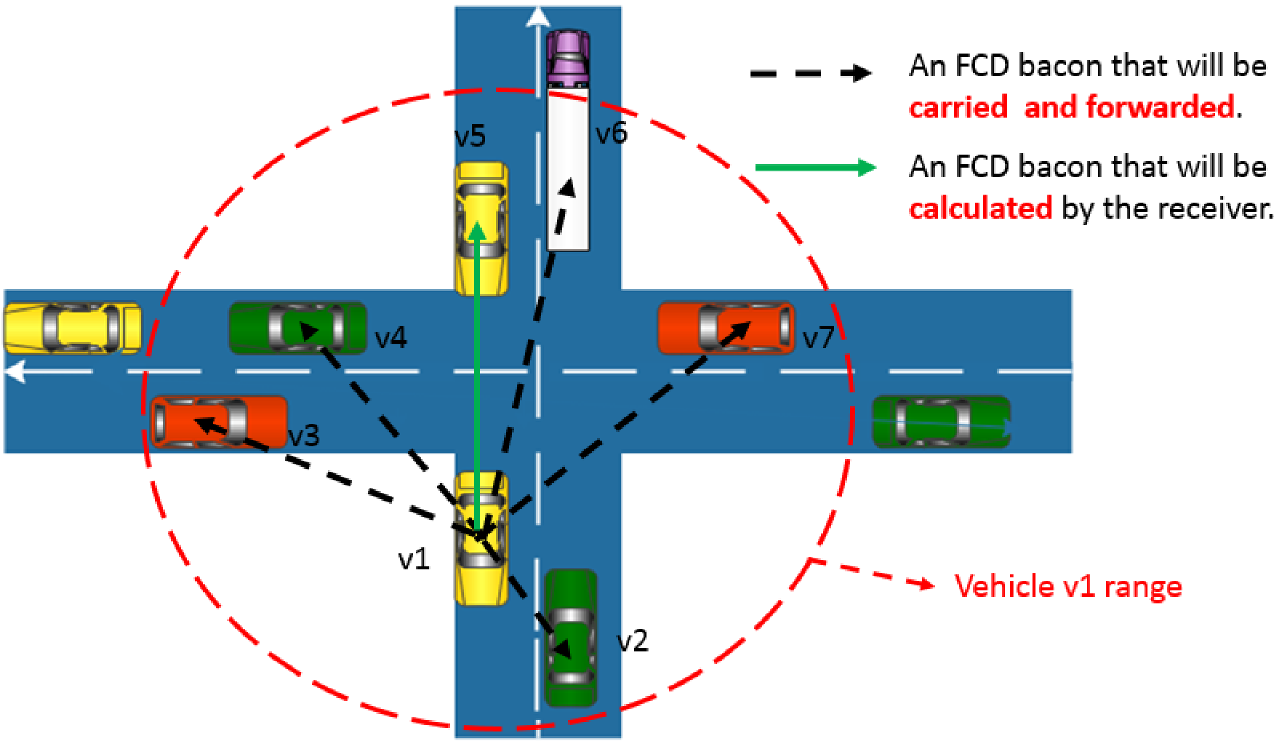


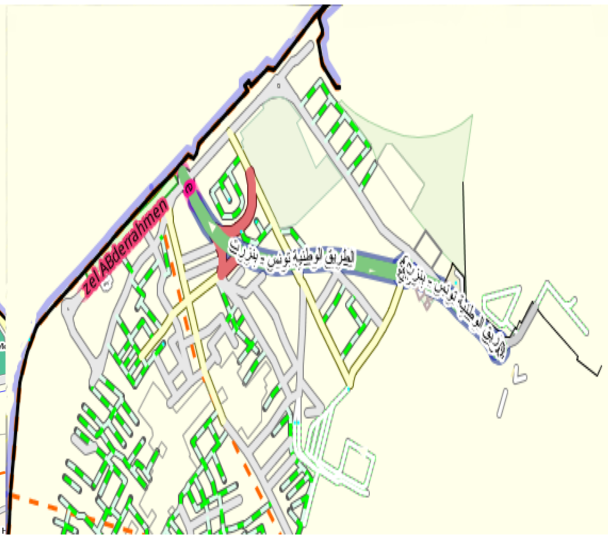
Fig. 4.4 Urban scenario case

Indeed, when a vehicle v receives a message, not in its direction, it will store, carry, and then forward it to other vehicles that may be going to this direction. Doing so, the traffic information will be disseminated to all vehicles in the road intersections. Algorithm 4 depicts a vehicle behavior upon receiving an FCD message in an urban situation.

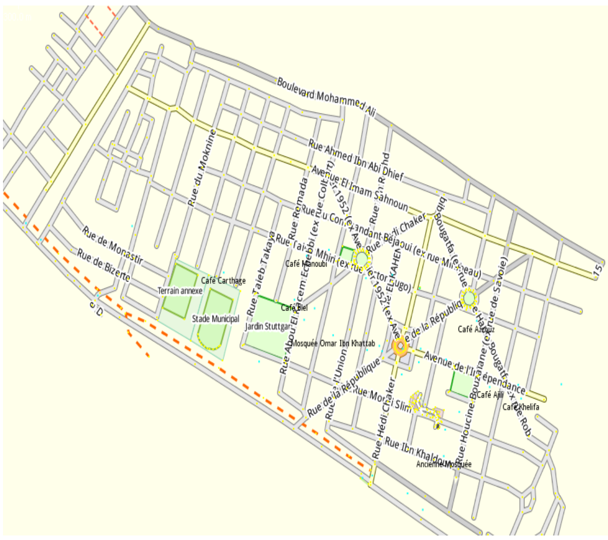
4.4 Performance evaluation

In this section, we present the performance evaluation conducted to evaluate our aggregation protocol. We choose SOTIS [46] and TrafficView [34] as baseline approaches since they focus only on the aggregated FCD without combining other data (like safety and non-safety events). CASCADE [20] will not be considered in our evaluation since it uses a compression algorithm on the top of TrafficView [34] to optimize the MAC-Layer utilization (this is out of scope of our study here but can also be applied by our approach after aggregation). We carry out our experiments using the Veins simulator framework[1]. Veins is an open source framework for Inter-Vehicular Communication (IVC) that combines both a road traffic micro-simulation model and an event-based network simulator. For our performance evaluation, we select real-world road topologies from three cities at the governorate of Bizerte in the north of Tunisia. We consider the three road topologies depicted in Figure 4.5, representing portions of the urban areas of Corniche, Zarzouna and Menzel Bourguiba cities. Table 4.1 summarizes the characteristics of each map. We vary the amount of vehicles driving on the map from 200 to 1000, ranging from low traffic to high traffic.

(a) Corniche



(b) Zarzouna



(c) Menzel Bourguiba

Fig. 4.5 Example of different road topologies

[1] *http://veins.car2x.org/*

Map	Number of roads	Number of Junctions	Dimensions
Corniche	503	214	2.2km x 2.5 km
Zarzouna	681	312	2.2km x 2.5 km
Menzel Bourguiba	919	468	2.2km x 2.5 km

Table 4.1 Road topologies characteristics

4.4.1 Experimental settings

The road traffic simulation is performed by SUMO [23] while the network simulation is performed by OMNeT++ [7] along with the physical layer modeling toolkit MiXiM[2], allowing the employment of accurate models for radio interference, as well as shadowing the use of static and moving obstacles.

Physical layer	Frequency band	5.9 GHz
	Transmission power	30 mW
	Bandwith	10 MHz
Link Layer	Bit rate	6 Mbit/s
	CW	15.1023
	Slot time	13 us
	SIFS	32 us
	DIFS	58 us
	Routing protocol	AODV/GPSR
Scenarios	message size	2312 Bytes
	Message frequency	0.5 Hz
	#Runs	30 times

Table 4.2 Simulation settings

With these two well-established simulators, nodes simulated by OMNeT++ 5.1.0 can interact with SUMO to simulate the influence of IVC on road traffic and mobility. We take advantage of these two simulators included in Veins to provide realistic models for 802.11p DSRC, PHY and MAC layers. The PHY and MAC parameters are defined according to the basic specifications of the 802.11p standard defined in [42]. The simulation settings are summarized in Table 4.2. In the MAC layer, we set the transmission power of a vehicle to 30mW to achieve approximately 300m of interference range.

[2] *http://mixim.sourceforge.net/*

4.4.2 Evaluation metrics

The assessment of the performances of our protocol is carried out through the following metrics:

1. **Overload:** It stands for the total number of sent packets. Interestingly enough, the ultimate goal of any aggregation protocol is to avoid the overload problem [12] by looking for minimizing the number of messages exchanged in the network. The average overload is defined as follows:

$$Overload = \frac{\sum sent\ packet}{\sum vehicle} \tag{4.2}$$

2. **Co2Emission:** refers to the amount of Co2 getting out from vehicles and affecting the environment [11]. Indeed, the ultimate goal of any traffic information system, and especially any aggregation protocol, is to reduce the total Co2 emmision [11] by looking for minimizing as much as possible traffic jams and decreasing the active waiting time of vehicles in cross roads and on highways. The average Co2 emission is defined as follows:

$$Co2Emission = \frac{\sum Vehicle\, Co2Emission}{\sum vehicle} \tag{4.3}$$

3. **Latency:** It is the time needed to deliver the aggregated message to an interested vehicle. The average latency, AL, is defined as follows:

$$AL = \frac{\sum (t_{v_i} - T)}{\sum Interested\ Vehicle} \tag{4.4}$$

where t_i stands for the arrival time of the event message to a vehicle v_i, and T is the time-stamp of the event.

4. **FCD Duplication Ratio:** It is the number of messages that are already sent and aggregated for a given message f. It is defined as follows:

$$Duplication\ \ ratio = \frac{\sum Duplicated\ f}{\sum f} \tag{4.5}$$

5. **Aggregation Precision:** It assesses to what extent our aggregation protocol is able to only aggregate an appropriate FCD message f for a given road r without duplication and to take into consideration the road maximum speed. Hence, the challenge will be

to obtain precision values of the average speed propagated to vehicles compared to the real road traffic state. It is defined as follows:

$$Precision(f) = \frac{|IIV|}{|AIV|} \tag{4.6}$$

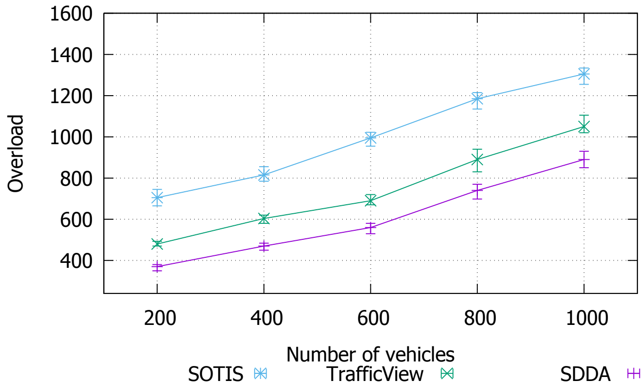
where *IIV* stands for the set of interested informed vehicles (i.e., only appropriate vehicles for a message f), and *AIV* stands for the set of all informed vehicles that will aggregate the received FCD message f (i.e., interested as well as not interested vehicles for an FCD f). The average precision is defined as follows:

$$AveragePrecision = \frac{\sum Precision(f)}{\sum f} \tag{4.7}$$

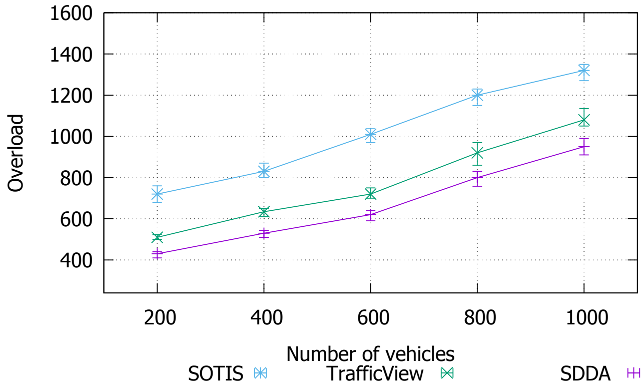
4.4.3 Results

As expected, using our aggregation protocol, the overload level is decreased (cf. Figure 4.6). This is owed to the fact that the number of duplicated messages is eliminated and the propagation of unnecessary aggregated FCD messages to uninterested vehicles is reduced. As a result, our protocol SDDA enables keeping a low overload. Moreover, Figure 4.7 shows that SDDA has a low Co2 Emmision value in different network densities. It is worth mentioning that our protocol decreases the Co2Emmision value in comparison to SOTIS and TrafficView by almost 70% since vehicles will have a precise traffic information and will avoid traffic jams. Figures 4.6 4.7 also demonstrate that the overload and Co2Emmision values will rise for all protocols when the number of vehicles increases. This interesting performance is related to the fact that our strategy permits ignoring all the unnecessary FCD messages when vehicles move based on direction and speed limitation filers. Moreover, Figure 4.9 indicates that the number of duplicated messages in our proposed solution is 99% less than other solutions thanks to the usage of the Slotted-1 persistence suppression technique [48] being able to eliminate all the duplicated messages. Finally, Figure 4.10 depicts that the latency of our SDDA strategy is slightly less than that of other strategies. This is due to the high overload level of SOTIS and TrafficView which increase the network collisions that affects the latency time. [39].

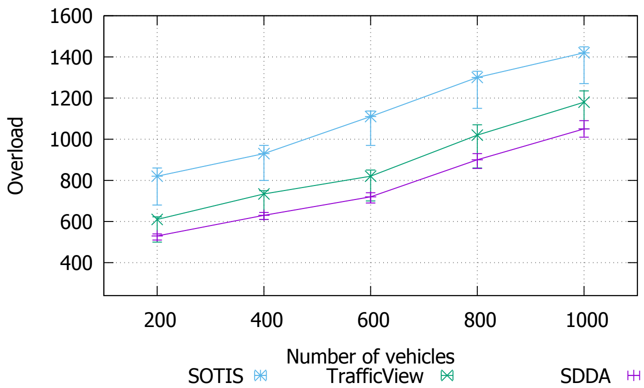
As expected, the packet loss ratio in Figure 4.11 is inversely proportional to the traffic density. In fact, it decreases as far as the number of vehicles grows. This can simply be explained by the fact that when raising the number of vehicles in the network, the communication overload as well as the message reception errors go up, which leads to increase the ratio of lost packets. Consequently, this fact will undoubtedly increase the number of vehicles that will not receive the messages or will receive damaged packets (unreadable, corrupted, etc.). To sum up, the simulation results highlight that our strategy performs much better than the baseline strategies from different perspectives (overload, co2emission, aggregation precision, packet loss and latency). Also, it is important to note that our strategy is genetic and can be adapted to any other aggregation context like road conditions, commercial advertisements and trip travel time.



(a) Zarzouna

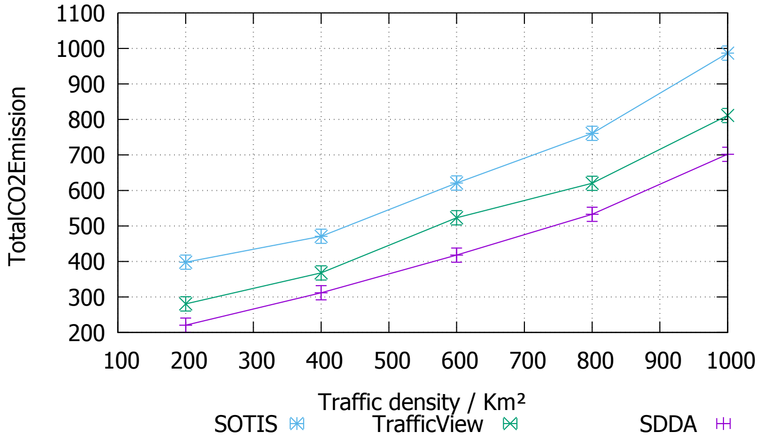


(b) Cornich

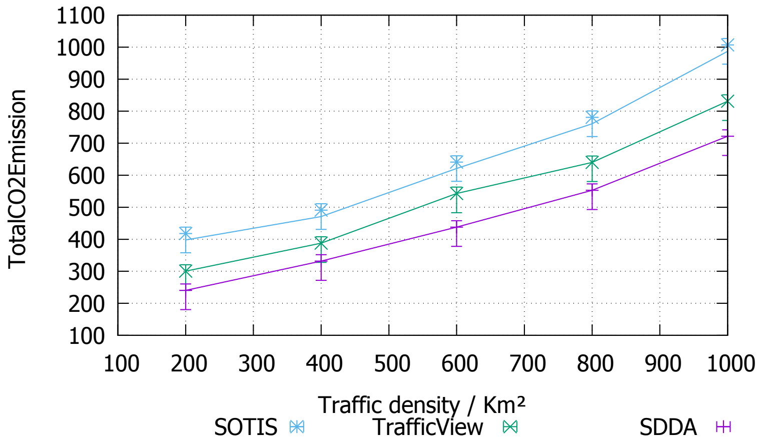


(c) Menzel Bourguiba

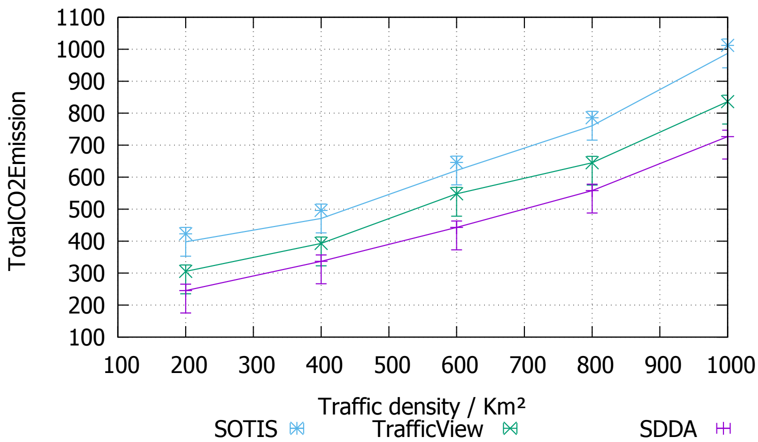
Fig. 4.6 Variation of average overload values w.r.t. number of vehicles



(a) Zarzouna

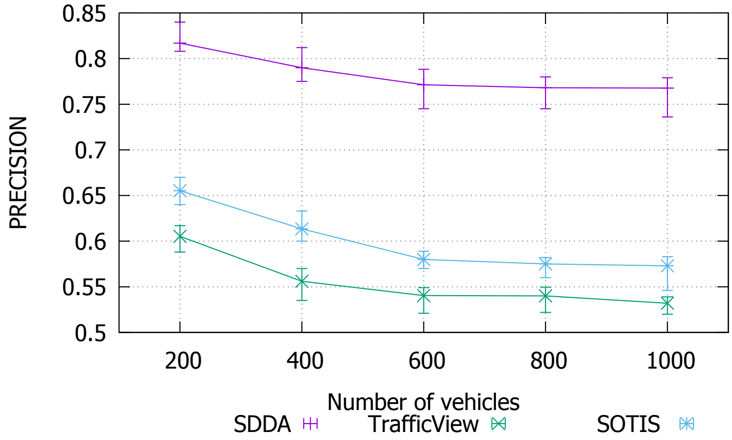


(b) Cornich

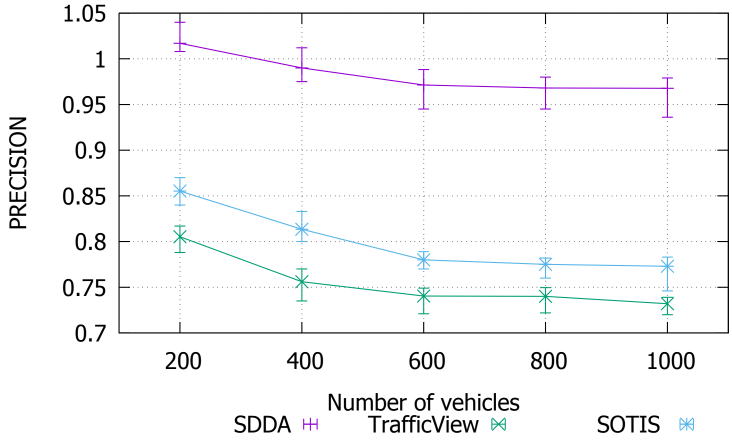


(c) Menzel Bourguiba

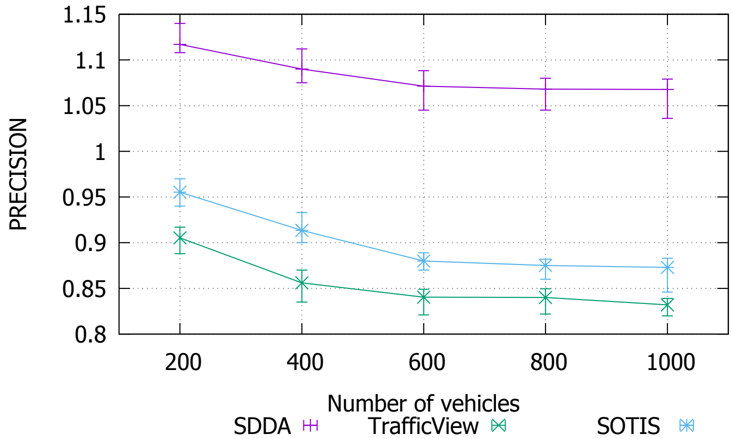
Fig. 4.7 Variation of average Co2 emission values w.r.t. number of vehicles



(a) Zarzouna

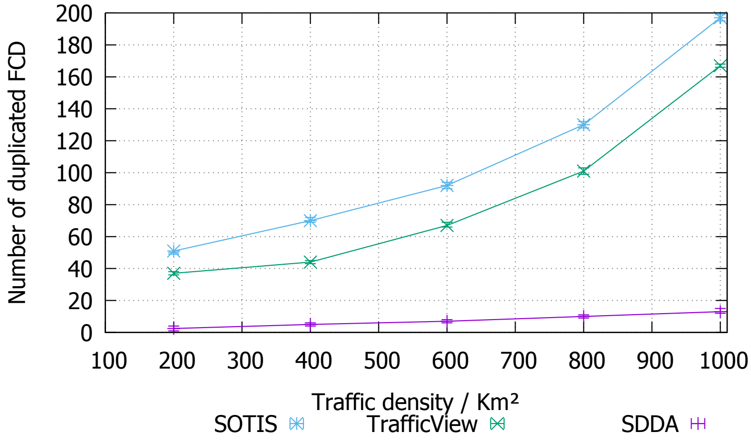


(b) Cornich

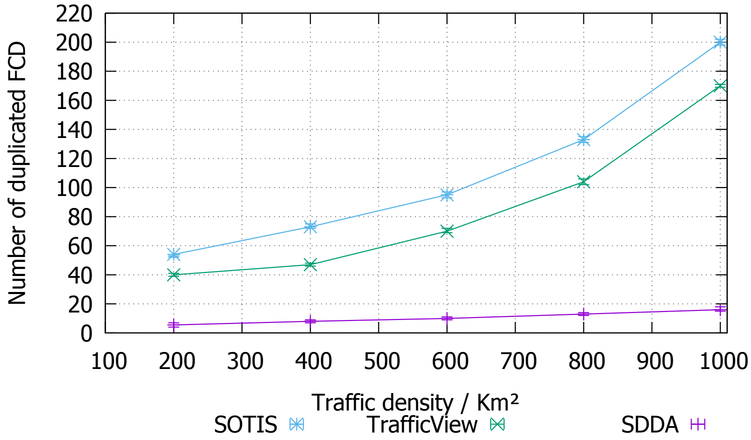


(c) Menzel Bourguiba

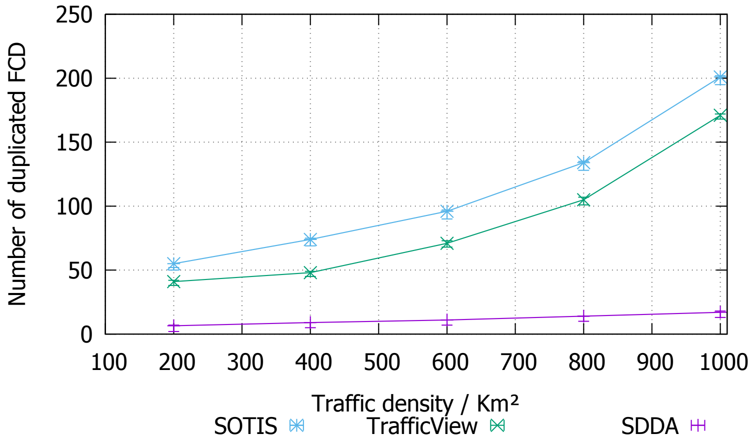
Fig. 4.8 Variation in average aggregation precision FCD values w.r.t. number of vehicles



(a) Zarzouna

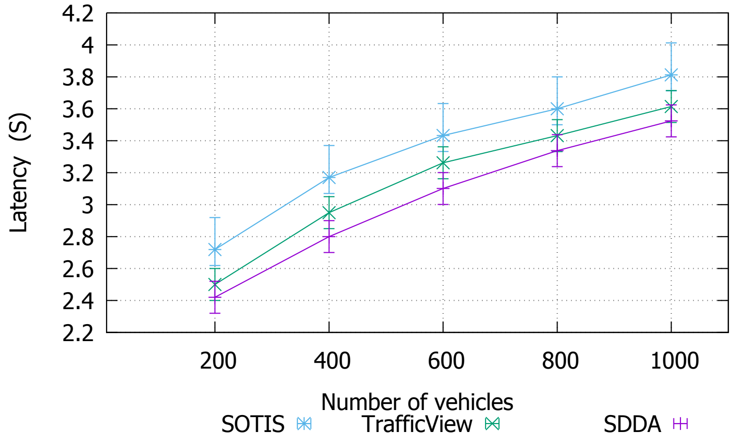


(b) Menzel Bourguiba

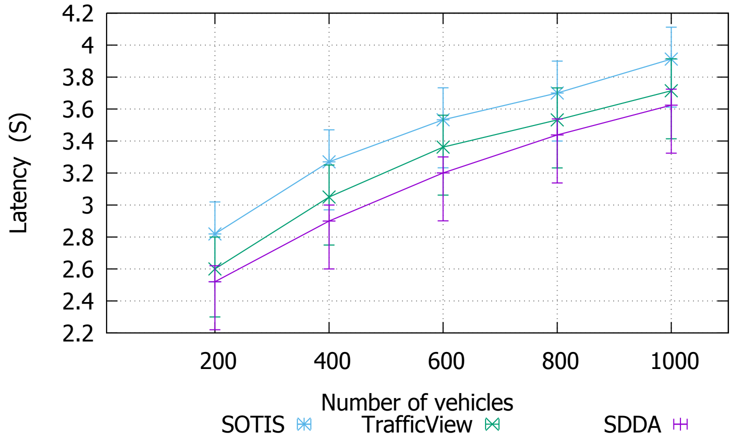


(c) Menzel Bourguiba

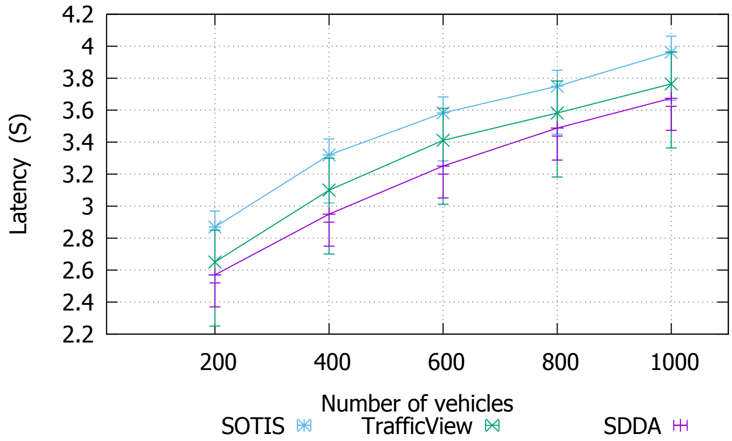
Fig. 4.9 Variation in average aggregation number of duplicated FCD values w.r.t. number of vehicles



(a) Zarzouna

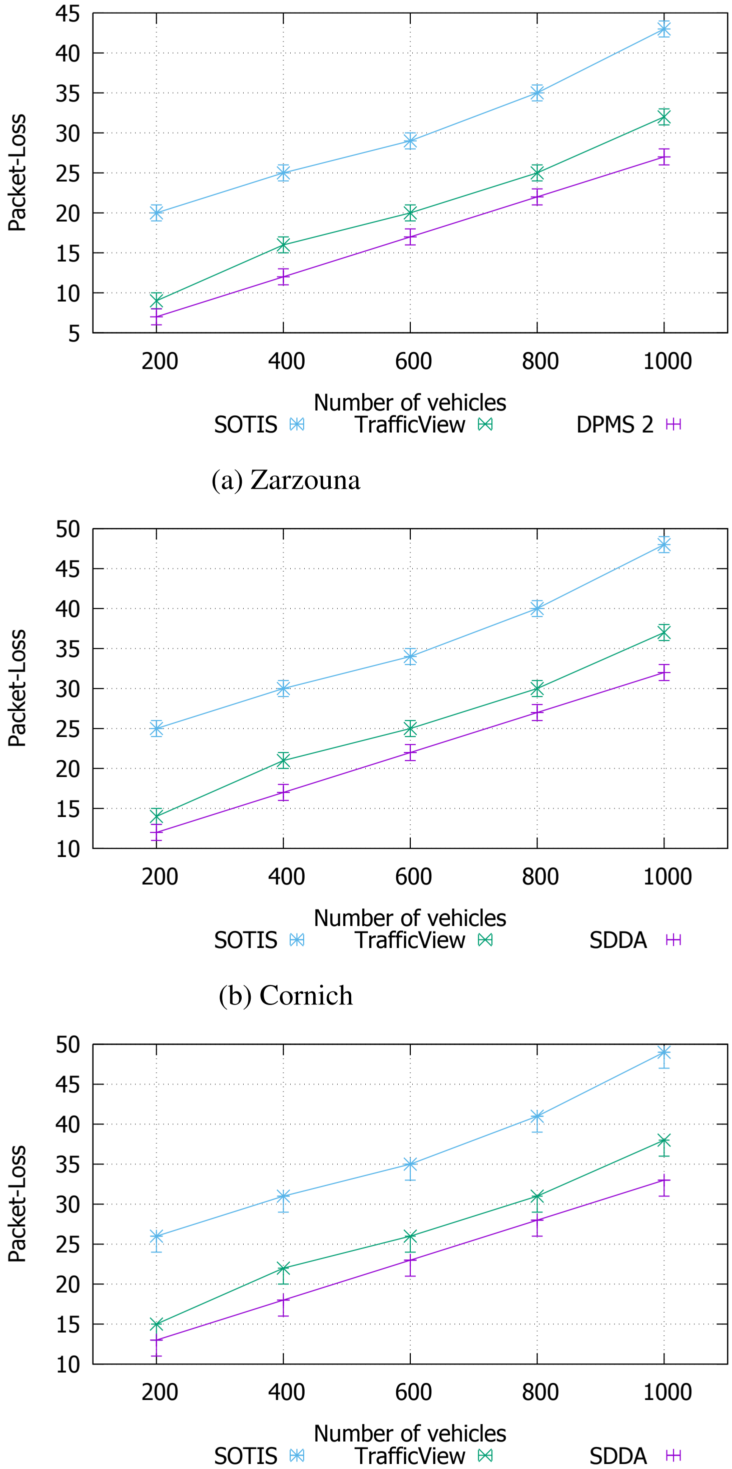


(b) Cornich



(c) Menzel Bourguiba

Fig. 4.10 Variation of average latency values w.r.t. number of vehicles



(a) Zarzouna

(b) Cornich

(c) Menzel Bourguiba

Fig. 4.11 Variation of the packet loss values w.r.t. number of vehicles

4.5 Conclusion

In this chapter, we have introduced SDDA as a generic smart directional data aggregation protocol to aggregate data about traffic in a VANET. The main thrust of our protocol stands on the use of direction and road speed limitation in data aggregation process. Doing so has allowed meeting several goals, namely reaching a low overload ratio as well as a high aggregation precision. Extensive experimental work has shown that SDDA obtains very interesting results in comparison with those provided by pioneering approaches of the literature.

Chapter 5

Conclusion and future directions

5.1 Conclusion

The emergence of VANET networks has allowed a great progress of advanced Intelligent Transportation System (ITS) services including various safety and non-safety applications. Due to the VANET environment, information has to be exchanged between mobile vehicles in an efficient way by avoiding as far as possible the broadcast storm problem, which needs efficient data dissemination and aggregation protocols.

In this respect, we have introduced in the first part of this thesis a map splitting-based dissemination protocol, called DPMS. The main thrust of our protocol stands for an adequate targeting of the zone of relevance of the disseminated messages. Doing so has allowed us to meet our goals, namely reaching a high delivery ratio as well as a high geocast precision. Extensive experimental work has shown that DPMS has obtained very encouraging results versus those got by pioneering approaches of the literature.

In the second part of this thesis, we have introduced SDDA as a generic smart directional data aggregation model protocol to exchange information about traffic in a VANET to overcome several limitations related to existing approaches. The main thrust of our protocol stands for an adequate targeting of using direction and road speed limitation. Doing so has allowed meeting several goals, namely reaching a low overload ratio as well as a high aggregation precision. Extensive experimental work has shown that SDDA obtains very interesting results in comparison with those provided by pioneering approaches of the literature.

5.2 Future directions

This thesis can constitute a basis for a lot of future work, which can be articulated on several axes. The details of this future work are as follows:

Managing the scalability issue for the extraction of the ZORs. Even though the size of the considered datasets is by far larger than those considered by the literature approaches, they are far from being firmly grounded within the era of big data. Thus, the scalability issue of the ZOR computation is still a thriving issue to carry extensive experiments by considering a higher number of vehicles. In addition, we will also consider the incremental maintenance of building ZORs.

Introducing a data aggregation mechanism to be used before broadcasting traffic information events. Indeed, without aggregation, a vehicle can use DPMS to send a warning message reporting the condition to vehicles inside ZOR. Other vehicles in the traffic jam can also start generating such warning messages, which will undoubtedly lead to a local broadcast storm.

Integrating the DPMS protocol within a signal phase and time information supporting a "green driving" for all vehicles and a safe and comfortable crossing of intersections even by blind and visually impaired pedestrians.

The scenarios that we consider in our simulations and our experiments in real environment are still limited, for example in the number of vehicles or in the number and size of objects. To ensure the effectiveness of our proposals in all possible scenarios, it may be particularly interesting to carry out large-scale tests. In the same way, we have considered in the various scenarios that the created objects are of the same size and that the preferences of the users are often uniformly distributed. Considering objects with different sizes will be interesting since it is the case in a real environment. The enthusiasm of users for certain types of content would also be fairly representative and quite favorable to our proposals. Likewise, studying and analyzing the load induced by the exchange of the beacon messages we use is important in order to show that it does not have a real impact and does not cause a too big overhead in the network.

In addition, dissemination depends essentially on the cooperation between the nodes to communicate and share the contents. However, nodes usually have a selfish behavior which can influence the performance of dissemination protocols. Therefore, taking into account this factor in dissemination is important and proposing solutions encouraging nodes to cooperate is a very interesting extension of our work.

In addition, validating the experiment carried out using mobile nodes (cars) equipped with the standard 802.11p will certainly be very informative. This will make it possible to highlight important parameters such as the distance between the nodes, the transmission range, the variability in the channel and the flow rate, thus validating the estimation of contact times in a real environment. Finally, estimating future contacts can improve dissemination in vehicular networks. An estimate based on a centralized method can be studied. Indeed, the periodic exchanges between the nodes and the central equipment, performed through the cellular network, allow the latter to predict future contacts. Nevertheless, in these approaches it will be important to study and analyze the overload induced on the cellular network because of these exchanges, the freshness of the information, and the load in terms of traffic and calculations that the central element will undergo.

We can first note that although vehicles are equipped with short-range communications, this is still insufficient because of the variability in the environment (urban, rural, highway) and road traffic (rush hour, day, late evening). This has an impact on the communication frequencies between the nodes and on the density of the network. To do this, proposing a communication paradigm between vehicular networks and telecommunications networks such as LTE will be useful for these two types of networks. Indeed, the vehicular networks can be a means to offload the large volume of traffic that undergoes LTE. The latter offers long-range communications to reach isolated nodes and instant communications when needed. Another aspect that could prolong the work of this thesis will be the integration of other types of networks with vehicular networks. The communication of a vehicle should not be limited to other vehicles that move. Thus, it will be interesting to allow even a pedestrian equipped with a smartphone in the street to be able to communicate and share contents with road users. Similarly, sensors installed at the edge of the road can communicate and inform vehicles automatically with the information collected. This makes it possible to extend the vehicular networks and minimize the intermittencies that may occur.

It cannot be denied that VANETs have been designed to bring a number of benefits, such as: reducing road traffic accidents, driving comfort and travel for drivers and passengers, easy payment for certain services such as parking, gas, etc. These networks also implement safety, maintenance and comfort applications such as Internet access, online and network games, and audio and video downloads. All these applications involve the exchange of messages such as emergency messages, warnings about incidents, road conditions at specific times, and driving assistance information. All these exchanges involve computer data and the content of the messages can influence the behavior of drivers, hence changing the topology of the network. This implies a risk of attack by malicious users who can interceptes the traffic messages exchanged on the network. Some attacks that can be observed on VANETs are: traffic blocking attacks, replay attacks, lie attacks on transmitted information, denial of service attacks, masquerade attacks, identity robbery attacks, illusion attacks, communication equipment attacks, etc.

For these reasons, security in these networks presents a very important challenge for researchers. There is a great deal of research on the safety of wireless vehicular networks to combat and protect these networks from threats and attacks. For the same purpose, in the near future we are interested in the security of the VANET network, the different types of attacks that this network may undergo, and the means of preventing these attacks.

Last but not least, our avenues of future work for data aggregation are as follows:

- We are actually working on real-word validation scenarios and we plan to extend the model to deal with other aggregation issues in VANETs like parking spaces, road conditions, and trip travel time.
- We provide a generic compression algorithm for FCD messages in order to reduce the bandwidth usage
- We integrate other dissemination protocols in our aggregation model in order to provide a more generic aggregation and dissemination protocol.

References

[1] Allal, S. and Boudjit, S. (2012). Geocast routing protocols for vanets: Survey and guidelines. In *Proc. of the Inl. Conf. on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, pages 323–328.

[2] Allani, S., Chbeir, R., Yeferny, T., and Yahia, S. B. (2018). Smart directional data aggregation in vanet. In *Proceedings - International Conference on Advanced Information Networking and Applications*.

[3] Allani, S., Yeferny, T., and Chbeir, R. (2017). A scalable data dissemination protocol based on vehicles trajectories analysis. *Ad Hoc Networks*, 71:31–44.

[4] Allani, S., Yeferny, T., Chbeir, R., and Ben Yahia, S. (2016a). A novel VANET data dissemination approach based on geospatial data. In *The 7th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2016)/The 6th International Conference on Current and Future Trends of Information and Communication Technologies*, pages 572–577.

[5] Allani, S., Yeferny, T., Chbeir, R., and Yahia, S. B. (2016b). Dpms: A swift data dissemination protocol based on map splitting. In *Proceedings - International Computer Software and Applications Conference*, volume 1, pages 817–822.

[6] Alsubaihi, B. and Boukerche, A. (2015). Semantic and self-decision geocast protocol for data dissemination over vanet (sas-gp). In *Wireless Communications and Networking Conference (WCNC)*, pages 1948–1953.

[7] Andras, V. (2010). Omnet++. In Wehrle, Klaus, Günes, Mesut, Gross, and James, editors, *Modeling and Tools for Network Simulation*, pages 35–59.

[8] Bako, B. and Weber, M. (2011). Efficient Information Dissemination in VANETs. In *Adv. Veh. Netw. Technol.*, page 20.

[9] Barbut, M. and Monjardet, B. (1970). *Ordre et classification. Algèbre et Combinatoire*.

[10] Bernhard, G. and Rudolfe, W. (1999). *Formal Concept Analysis.*

[11] Cappiello, A., Chabini, I., Nam, E. K., Lue, A., and Abou Zeid, M. (2002). A statistical model of vehicle emissions and fuel consumption. *IEEE 5th Conference on Intelligent Transportation Systems*, pages 801–809.

[12] Commission, F. C. (2004). Amendments of the commission's rules regarding dedicated short-range communication services in the 5.9 ghz band. *Federal Communications Commission Proceedings*, pages 10–809.

[13] Corson, S. and Macker, J. (1999). Mobile ad hoc networking (manet): Routing protocol performance issues and evaluation considerations", rfc 2501.

[14] Delot, T., Ilarri, S., Cenerario, N., and Hien, T. (2011). Event sharing in vehicular networks using geographic vectors and maps. *Mob. Inf. Syst.*, 7:21–44.

[15] Duan, X., Liu, Y., and Wang, X. (2017). Sdn enabled 5g-vanet: Adaptive vehicle clustering and beamformed transmission for aggregated traffic. *IEEE Communications Magazine*, 55(7):120–127.

[16] George, B. and Kim, S. (2013). Spatio-temporal networks. *Database Management and Information Retrieval*.

[17] George, B. and Shekhar, S. (2008). *Time-Aggregated Graphs for Modeling Spatio-temporal Networks*. Springer Berlin Heidelberg.

[18] Ibrahim, K. (2011). Data aggregation and dissemination in vehicular ad-hoc networks by data aggregation and dissemination in vehicular ad-hoc networks. (4):1–120.

[19] Ibrahim, K., Weigle, M., and Abuelela, M. (2009). p-ivg: Probabilistic inter-vehicle geocast for dense vehicular networks. In *Vehicular Technology Conference (VTC)*, pages 1–5.

[20] Ibrahim, K. and Weigle, M. C. (2008). Cascade: Cluster-based accurate syntactic compression of aggregated data in vanets. *2008 IEEE Globecom Workshops, GLOBECOM 2008*, pages 1–10.

[21] Jayapal, C. and Roy, S. (2016). Road traffic congestion management using vanet. In *Proc. of the International Conference on Advances in Human Machine Interaction (HMI)*, pages 1–7.

[22] Jochle, T., Wiedersheim, B., Schaub, F., and Weber, M. (2012). Efficiency analysis of geocast target region specifications for vanet applications. In *Proc. of the Intl. Vehicular Networking Conference (VNC)*, pages 250–257.

[23] Kaisser, F., Gransart, C., and Berbineau, M. (2012). Simulations of vanet scenarios with opnet and sumo. In *Communication Technologies for Vehicles*, volume 7266 of *Lecture Notes in Computer Science*, pages 103–112. Springer Berlin Heidelberg.

[24] Kenney, J. (2011). Dedicated short-range communications (dsrc) standards in the united states. 99(7):1162–1182.

[25] Kenney, J. B. (2010). *Standards and Regulations, in VANET: Vehicular Applications and Inter-Networking Technologies*, pages 365–429.

[26] Kheawchaoom, P., Kittipiyakul, S., and Ayutayay, K. S. N. (2012). idtsg time-stable geocast for post crash notification in vehicular highway networks. In *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 1–4.

[27] Kumar, R. and Dave, M. (2013). A framework for handling local broadcast storm using probabilistic data aggregation in vanet. *Wireless Personal Communications*, page 315–341.

[28] Li, X. and Li, H. (2014). A survey on data dissemination in vanets. *Chinese Science Bulletin*, 59(32):4190–4200.

[29] Lim, Y., Ahn, S., and Cho, K.-H. (2011). Abiding geocast for commercial ad dissemination in the vehicular ad hoc network. In *Proc. of the Intl. conference on Consumer Electronics (ICCE)*, pages 115–116.

[30] Luo, J., Gu, X., Zhao, T., and Yan, W. (2010). A mobile infrastructure based vanet routing protocol in the urban environment. In *Proc. of the Intl. conference on Communications and Mobile Computing (CMC)*, volume 3, pages 432–437.

[31] Maurer, J., Fugen, T., Schafer, T., and Wiesbeck, W. (2004). A new inter-vehicle communications (ivc) channel model. In *Proc. of the Intl. conference on Vehicular Technology Conference (VTC)*, pages 9–13.

[32] Mohanty, S. and Jena, D. (2012). Secure data aggregation in vehicular-adhoc networks: A survey. *2nd International Conference on Communication, Computing and Security ICCCS-2012*, 6:922 – 929.

[33] Mouakher, A. and Yahia, S. B. (2016). Qualitycover: Efficient binary relation coverage guided by induced knowledge quality. *Information Sciences*, 355-356:58–73.

[34] Nadeem, T., Dashtinezhad, S., Liao, C., and Iftode, L. (2004). Trafficview : Traffic data dissemination using car-to-car communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, (3):1–8.

[35] Rahbar, H., Naik, K., and Nayak, A. (2010). Dtsg: Dynamic time-stable geocast routing in vehicular ad hoc networks. In *Proc. of the Inl. 9th IFIP Annual Mediterranean Ad Hoc Networking (Med-Hoc-Net)*, pages 1–7.

[36] Saleet, H., Basir, O., Langar, R., and Boutaba, R. (2011a). Region-based location-service-management protocol for vanets. *IEEE Transactions on Vehicular Technology*, pages 917–931.

[37] Saleet, H., Langar, R., Naik, K., Boutaba, R., Nayak, A., and Goel, N. (2011b). Intersection-based geographical routing protocol for vanets: A proposal and analysis. *IEEE Transactions on Vehicular Technology*, pages 4560–4574.

[38] Schwartz, R. S., Scholten, H., and Havinga, P. (2013). A scalable data dissemination protocol for both highway and urban vehicular environments. *EURASIP Journal on Wireless Communications and Networking*, 2013:1–19.

[39] Sommer, C. and Dressler, F. (2015a). Information Dissemination in Vehicular Networks. In *Vehicular Communications and Networks: Architectures, Protocols, Operation and Deployment*, pages 75–93.

[40] Sommer, C. and Dressler, F. (2015b). Information Dissemination in Vehicular Networks. In *Vehicular Communications and Networks: Architectures, Protocols, Operation and Deployment*, pages 75–93. Woodhead.

[41] Sommer, C., German, R., and Dressler, F. (2011a). Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis. *IEEE Transactions on Mobile Computing*, 10(1):3–15.

[42] Sommer, C., German, R., and Dressler, F. (2011b). Bidirectionally coupled network and road traffic simulation for improved ivc analysis. *IEEE Transactions on Mobile Computing*, 10:3–15.

[43] Tian, B., Hou, K. M., and Li, J. (2016). Trad: Traffic adaptive data dissemination protocol for both urban and highway vanets. In *Proc. of the Intl. conference on Advanced Information Networking and Applications (AINA)*, pages 724–731.

[44] Tsai, H. W. (2013). Aggregating data dissemination and discovery in vehicular ad hoc networks. *Telecommunication Systems*, page 285–295.

[45] Tyagi, S. S. (2013). International Journal of Advanced Research in Computer Science and Software Engineering. 3(5).

[46] Wischoff, L., Ebner, A., Rohling, H., Lott, M., and Halfmann, R. (2003). Sotis - a self-organizing traffic information system. In *The 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring.*, pages 1–6.

[47] Wisitpongphan, N., Tonguz, O., Parikh, J., Mudalige, P., Bai, F., and Sadekar, V. (2007a). Broadcast storm mitigation techniques in vehicular ad hoc networks. *Wireless Communications, IEEE*, 14(6):84–94.

[48] Wisitpongphan, N., Tonguz, O., Parikh, J., Mudalige, P., Bai, F., and Sadekar, V. (2007b). Broadcast storm mitigation techniques in vehicular ad hoc networks. *Wireless Communications, IEEE*, 14(6):84–94.

[49] Yamada, A. and Ishihara, S. (2017). Data exchange strategies for aggregating geographical distribution of demands for location-dependent information using soft-state sketches in VANETs. *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, pages 728–736.

[50] Yeferny, T. and Allani, S. (2018). Mpc: A rsus deployment strategy for vanet. *International Journal of Communication Systems*, 31(12).

[51] Yeferny, T., Arour, K., and Bouzeghoub, A. (2013a). An efficient peer-to-peer semantic overlay network for learning query routing. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, pages 1025–1032. IEEE.

[52] Yeferny, T., Arour, K., and Bouzeghoub, A. (2013b). An efficient peer-to-peer semantic overlay network for learning query routing. In *27th IEEE International Conference on Advanced Information Networking and Applications, AINA 2013, Barcelona, Spain, March 25-28, 2013*, pages 1025–1032.

[53] Yu, Q. and Heijenk, G. (2008). Abiding geocast for warning message dissemination in vehicular ad hoc networks. In *Proc. of the Intl. Communications Workshops (ICC Workshops)*, pages 400–404.

[54] Zekri, D., Defude, B., and Delot, T. (2014). Building , sharing and exploiting spatio-temporal aggregates in vehicular networks. *Hindawi Mobile Information Systems*, 14:259–285.

[55] Zhang, L., Gao, D., Gao, S., and Leung, V. (2013). Smartgeocast: Dynamic abnormal traffic information dissemination to multiple regions in vanet. In *Proc. of the Inl. on Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1750–1755.

[56] Zhao, J. and Cao, G. (2006). Vadd : Vehicle-assisted data delivery in vehicular ad hoc networks. *Proc. Intl. Conference on Computer Communications (INFOCOM)*, pages 1–12.

[57] Zhu, Y., Zhao, Q., and Zhang, Q. (2015). Delay-constrained data aggregation in vanets. *IEEE Transactions on Vehicular Technology*, pages 2097–2107.

[58] Zongjian, H. and Daqiang, Z. (2017). Cost-efficient traffic-aware data collection protocol in vanet. *Ad Hoc Networks*, 55:28–39.