



**HAL**  
open science

# A Scalable Data Dissemination Protocol Based on Vehicles Trajectories Analysis

Sabri Allani, Taoufik Yeferny, Richard Chbeir

► **To cite this version:**

Sabri Allani, Taoufik Yeferny, Richard Chbeir. A Scalable Data Dissemination Protocol Based on Vehicles Trajectories Analysis. *Ad Hoc Networks*, 2018, 71, pp.31-44. 10.1016/J.ADHOC.2017.12.003 . hal-01906763

**HAL Id: hal-01906763**

**<https://univ-pau.hal.science/hal-01906763>**

Submitted on 15 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A scalable data dissemination protocol based on vehicles trajectories analysis

Sabri Allani<sup>a,b</sup>, Taoufik Yeferny<sup>a,c</sup>, Richard Chbeir<sup>b</sup>, Sadok Ben Yahia<sup>a</sup>

<sup>a</sup>*LIPAH-LR 11ES14, University of Tunis El Manar, 2092 Tunis*

<sup>b</sup>*LIUPPA LAB, University of Pau and Adour Countrie ,64600 Anglet, France*

<sup>c</sup>*College of Science, Northern Border University, Arar, Saudi Arabia*

---

## Abstract

The emergence of affordable wireless devices in vehicle ad hoc networks was a key step towards improving road safety as well as transport efficiency. Through informing vehicles about interesting safety and non-safety events, the challenge to take up would be to avoid, as far as possible, the broadcast storm problem. Thus, the design of an efficient dissemination protocol has been of paramount importance. The thriving challenge would be to maximize the reachability ratio. At the crossroads of the wealthy number of literature approaches stands the definition of the zone of relevance for event dissemination. Unfortunately, the definition of this zone of relevance is carried out in an ad hoc manner. In this paper, and to palliate this shortage, we introduce a new infrastructure-less geocast approach, called data *Dissemination Protocol based on Map Splitting* (DPMS). The main thrust of the DPMS stands on the formalization of the zone of relevance through the mining of correlations between kept track connections of vehicle trajectories versus regions. These correlations are faithfully described through a set of formal concepts. This latter notion stands at the corner stone of the *Formal Concept Analysis* (FCA) field. Therefore, we provide an efficient determination of the zone of relevance through the instantiation of the well known cover-set problem. The performed experiments show that DPMS outperforms its competitors in terms of effectiveness and efficiency.

*Keywords:* VANET, Data dissemination, Geocast, FCA

---

## 1. Introduction

Owing to the embedded devices into modern cars, the 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 information, as an accident on the road, an empty place in a parking, bottling, obstacles, weather, road cut, to cite but a few.

---

\*Corresponding author. Tel.: +216 53 652 717; fax: +216 71 885 190

*Email addresses:* sabri.allani@gmail.com (Sabri Allani), yeferny.taoufik@gmail.com (Taoufik Yeferny), richard.chbeir@univ-pau.fr (Richard Chbeir), sadok.benyahia@fst.rnu.tn (Sadok Ben Yahia)

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. In this respect, a core and challenging issue in vehicular networks is the design of an efficient data dissemination protocol able to inform vehicles about interesting events. The thriving challenge would be to maximize the reachability ratio, i.e. by only informing the interested vehicles, and avoiding as far as possible the broadcast storm problem <sup>1</sup>. A careful scrutiny of the pioneering vehicle-to-vehicle data dissemination approaches highlights that the latter could be done through broadcast or geocast techniques. Worth of mention, the main moan that can be addressed to the broadcast technique stands on the costly dissemination of messages to all the vehicles in the networks without exception (i.e., vehicles interested in the event or not). This drawback leads us to naturally opt for the geocasting technique, which requires delivering information to vehicles inside a specific region unlike what broadcasting does. 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 approach based on the geocast technique has to fulfill with the following requirements:

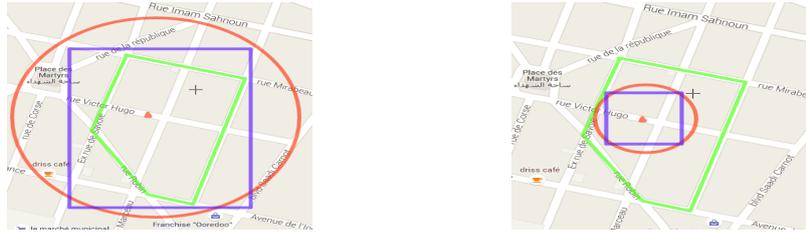
1. determining the geocast area, also called *Zone Of Relevance (ZOR)* of an event;
2. delivering the message to all vehicles within the ZOR; and
3. keeping the geocasted message alive in the network for a desired delay, such that the disseminated information could reach all the arriving vehicles.

Even though the literature witnesses a wealthy number of geocast based techniques for data dissemination, only few of them consider all of the three above mentioned requirements. Indeed, the most challenging issue for geocast protocols is the definition of the ZOR for event dissemination. Unfortunately, the existing geocast protocols define the ZOR in an ad hoc manner as a rectangle [1] [2] [3], a circle or a polygon [4]. However, none of the used shapes has matched the ZOR as closely as possible. For example, in Figure 1, 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 (a)) several non-interested vehicles receive the message and a lot of unnecessary messages are exchanged. Whereas, in the second case (Figure 1 (b)) many interested in vehicles do not receive the message.

In order to cope with the above-mentioned requirements, we propose here a new infrastructure-less geocast approach for urban area, called data Dissemination Protocol based on Map Splitting (*DPMS*). The latter aims to reach a high reachability ratio as well as a high geocast precision by sending messages only to vehicles in the ZOR with a minimum overhead cost. 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 comes back to an instantiation of the cover set problem [5], 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.

---

<sup>1</sup>The latter is known to lead to network saturation as well as conflict and collision issues



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

Figure 1: Zone of relevance vs target region specifications

The remainder of this paper is organized as follows. In section 2, we describe the pioneering approaches of the literature in order to show their limitations/drawbacks. In section 3, we thoroughly describe the main idea of our dissemination protocol. In section 4, the simulation settings and the evaluation of the proposed DPMS protocol are then presented. Besides, a comparison between DPMS versus the other surveyed geocast protocols is also presented. The last section concludes this paper and pins down several issues of future work.

## 2. Related work

VANET security applications aim to improve 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 [4]. The Infrastructure-based protocols [6, 7, 8], use RoadSide Units (RSU) in junctions and along the roads to store and disseminate VANET messages. These protocols achieve generally good results. However, they rely on costly infrastructures. In this respect, infrastructure-less protocols have been recently introduced to disseminate information without relying on a costly infrastructure. They are known as broadcast and geocast data dissemination protocols [9]. In the following, we discuss most recent ideas including intelligent broadcasting and geocasting techniques. Applications widely range from emergency messaging to exchanging traffic information.

### 2.1. Broadcast-based data dissemination

Broadcasting techniques are frequently used in VANETs for data sharing, traffic information, weather, entertainment and commercial announcements applications, which aim to disseminate information to all vehicles, without exception, using a blind or a moderated flooding mechanism. However, they are unsuitable for safety events dissemination, which are of interest to some vehicles near the event location. 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 known as the broadcast storm problem [10]. In the literature, several research studies have attempted to improve this naive method by using suppression techniques, which are probabilistic (e.g., weighted p-persistence) [10], timer-based (i.e., Slotted 1-persistence) [10], or hybrid (i.e., Slotted p-persistence) [10]. These techniques aim 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  $p$ -weighted persistence [10], a forwarding probability is assigned to each neighboring vehicle according to its distance from 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. In Slotted 1-persistence [10], 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 probability 1 at the assigned time slot if it receives the packet for the first time and has not received any duplicates before its assigned time slot; otherwise, it discards the packet. In Slotted  $p$ -persistence [10], 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. In [10] the authors concluded, after performing several simulations, that the Slotted-1 persistence is the most efficient technique. In [11], the authors addressed the multi-directional dissemination issue in an urban scenario within a dense network. Therefore, they introduced an optimized version of the Slotted 1-persistence where the main optimization was to give a high priority of dissemination to the farther vehicles having the same direction as the event message.

As already pointed out in early work on vehicular networks, geocasting techniques can help overcoming the broadcast storm problem more effectively [9]. Furthermore, they have practical applications more especially for disseminating safety event. In the following, we provide an overview of some recent works that have tackled the problem of data geocasting in vehicular networks.

## 2.2. Geocast-based data dissemination

Geocast protocols aim to disseminate data only to vehicles inside a specific geographical area, called the ZOR [12]. It is worth mentioning that, in geocast protocols – and after determining the ZOR – probabilistic, timer-based, priority-based or hybrid suppression broadcasting techniques could be used to disseminate the message inside the ZOR. The vehicles receiving the message outside the specified area simply ignore the message [13].

In this respect, Ibrahim et al. [1] introduced a probabilistic and beacon-based approach named  $p$ -IVG (*probabilistic Inter-Vehicular Geocast*). The  $p$ -IVG approach would tackle the drawbacks of timer based broadcasting approach, *e.g.* the Slotted 1-persistence [10] 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 will re-broadcast the packet at the same time. This will 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, which 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. [2] and Kheawchaoom et al. [3] introduced, respectively, two versions of the Dynamic Time-Stable Geocast protocol (DTSG) for the highways scenario. The main focus was to ensure the message delivery with a low cost 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 comes back to geocasting the message to the specific region. Whenever a vehicle detects such a critical event, then 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. After this period, the protocol then 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. [4] split the dissemination area, which is equivalent to a ZOR, into a set of sub-ZORs to deal with the temporal network fragmentation issue. Moreover, they used simple geometrical forms (*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 provided a technique allowing to determine whether some sub-ZORs were in the same direction to address them the same message. It is worth mentioning that the authors delegated the affectation of the ZOR to a competent 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.* [1] [2] [3] to cite but a few. Whereas, in some other approaches, it has been defined as a circle or a polygon [4]. Although Jochle et al. [14] concluded that a circular geocast area performed better in most of the considered scenarios they used, the ZOR specification still a thriving challenge in a geocast protocol.

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. [15] introduced a dissemination protocol that followed 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, then the message will be considered as relevant enough to be rediffused by the receiving vehicle. Otherwise, the message will be 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 moan that can be adressed to this approach stands in the fact that it does not keep the event alive within the DDA and the suppression technique of use, 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 undoubtedly would lead to a local broadcast storm. Alsubaihi et al. [16] introduced the SAS-GP protocol, which initially executes an algorithm for locally determining 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 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 that lead up to the event location. Consequently, several non-interested vehicles will receive the event message. In fact, the vehicles that drive in a road leading to the event location are not necessarily interested in the event.

Scrutiny of the above mentioned work highlights that all the above mentioned approaches cannot effectively fulfill all the three predefined requirements of geocast protocols. For the sake of ensuring a high reachability ratio as well as a high geocast precision by only sending messages 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:

- Efficiently determine 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, a smart algorithm determines the set of regions composing its ZOR. Hence, a centralized database is dedicated to store the result of map decomposition and ZOR association.
- Keep the event alive in the ZOR during the event lifetime. Indeed, the geocasted message should remain in the ZOR for a desired delay, such that arriving vehicles can be informed. Only few approaches, [2] and [3], addressed this issue by rebroadcasting the event during a certain delay, which significantly increases the network overload. To tackle this issue, in our protocol, whenever a vehicle enters in a new region, it can retrieve events received by other vehicles in the current region. By doing so, pertinent events are kept alive inside the region, without the need for rebroadcasting event messages as the existing geocast approaches do.
- Avoid the broadcast storm during the dissemination of events to vehicles within the ZOR by using a moderated Slotted-1 persistence technique. In fact, this suppression technique avoids the broadcast storm problem and guarantees a high *reachability*.

### 3. Protocol description

In the following, we start by defining the main concepts that are of use for the remainder.

**Definition 1. Road( $rd$ ):** is a path that connects two junctions. We denote by  $rd^*$  the set of the two junctions connected by 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 a common junction, i.e.,  $rd_i^* \cap rd_j^* \neq \emptyset$

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

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

Where:

- $rd_i$  denotes a road in the region  $r$ .
- $\forall rd_i \in r \exists rd_j \in r$  such that  $rd_i^* \cap rd_j^* \neq \emptyset$

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

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

Where:

- $r_i$  denotes a region in the city  $c$ .
- $\forall r_i \in c \exists r_j \in c$  such that  $\exists rd_p \in r_i \exists rd_k \in r_j$  and  $rd_p^* \cap rd_k^* \neq \emptyset$

**Definition 4. (Zone Of Relevance):**

The the zone of relevance of a region  $r$ , denoted  $ZOR(r)$ , is the set of regions where events that arise in the region  $r$  are also pertinent for vehicles in those regions. Formally, let  $R$  be a set of regions and  $\mathcal{P}(R)$  the powerset of  $R$ , we define  $ZOR(r)$  as follows:

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

Where events that arise in the region  $r$  are also pertinent for the for vehicles in the regions  $r_i$  ( $1 \leq i \leq n$ ).

**Definition 5. (Split Map)**

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

$$\begin{aligned} SP: C &\rightarrow R \times \mathcal{P}(R) \\ c &\mapsto \{(r_1, ZOR(r_1)), \dots, (r_n, ZOR(r_n))\}. \end{aligned}$$

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

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

### 3.1. Architecture

As depicted in Figure 2, the architecture of our proposal consists of two tiers of nodes:

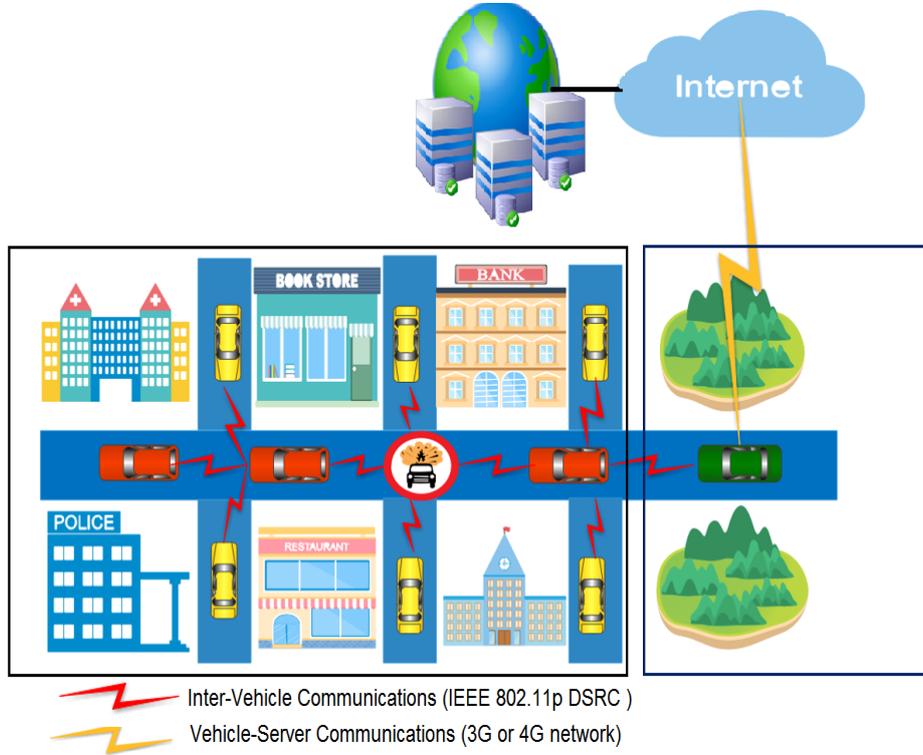


Figure 2: General architecture

- **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 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  $c$ , noted  $SP(c)$ ) is then stored in a central knowledge base. To do this, the server firstly sends a query to vehicles driving in a city  $c$  to gather their traces files (i.e., each file contains a set of vehicle trajectories, where each vehicle trajectory is a set of roads crossed by the vehicle). Thereafter, it runs a smart algorithm that computes the ZOR of each region in the city based on the gathered vehicles trajectories. It is possible to update the knowledge base in order to improve the quality of the ZORs computation. This task requires gathering new vehicles trajectories and then running the ZORs computation algorithm, which needs a sophisticated scheduling strategy. It is worth of note that the update issue is beyond the scope of this paper.
- **Vehicles:** they disseminate events using our protocol DPMS. Indeed, in DPMS the data dissemination process is done only through V2V communications by using IEEE 802.11p DSRC technology [17]. However, DPMS requires the split map  $SP(c)$ , available in the server, of the city  $c$  in which the vehicle is moving. Hence, to avoid vehicle/server communications during the event dissemination time, vehicles download the split map  $SP(c)$ ,

via cellular network technology such as 3G or 4G, whenever they enter in an unvisited city (e.g., the green car in Figure 2). Finally, we also require that vehicles are able to determine their respective positions on the road using, e.g., the Global Positioning System (GPS).

In the following subsections 3.2, 3.3 we respectively detail our ZOR computation method and our DPMS protocol.

### 3.2. 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 determination is a thriving challenge leading to increase the reachability ratio and to overcome the broadcast storm problem. In this respect, we introduce 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, noted  $ZOR(r_i)$ . For example, in Figure 3 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 matches the green region that is considered as the ZOR in Figure 1.



Figure 3: Map Splitting and ZOR determination

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) [18], 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 that of 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. 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 from the FCA.

#### 3.2.1. Key notions

##### **Definition 6. (FORMAL CONTEXT)**

A formal context is a triplet  $\mathcal{K} = (\mathcal{V}, \mathcal{R}, \mathcal{I})$ , where  $\mathcal{V}$  represents a finite set of vehicles trajectories,

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

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

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

**Definition 7. (GALOIS CONNECTION)**

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

$$\begin{aligned} \psi : \mathcal{P}(\mathcal{V}) &\rightarrow \mathcal{P}(\mathcal{R}) \\ V &\mapsto \psi(V) = \{r \in \mathcal{R} \mid \forall v \in V, (v, r) \in \mathcal{I}\} \end{aligned}$$

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

$$\begin{aligned} \phi : \mathcal{P}(\mathcal{R}) &\rightarrow \mathcal{P}(\mathcal{V}) \\ R &\mapsto \phi(R) = \{v \in \mathcal{V} \mid \forall r \in R, (v, r) \in \mathcal{I}\} \end{aligned}$$

The coupled applications  $(\psi, \phi)$  form a Galois connection between the power-set of  $\mathcal{V}$  and of  $\mathcal{R}$  [19].

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

**Definition 8. (FORMAL CONCEPT)**

A pair  $\langle V, R \rangle \in \mathcal{V} \times \mathcal{R}$  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,

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

**Example 3.** With respect to the formal context shown by Table 1, Figure 4 (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 4 (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  $\langle \{v_1v_2\}, \{r_3r_5r_6\} \rangle$ .

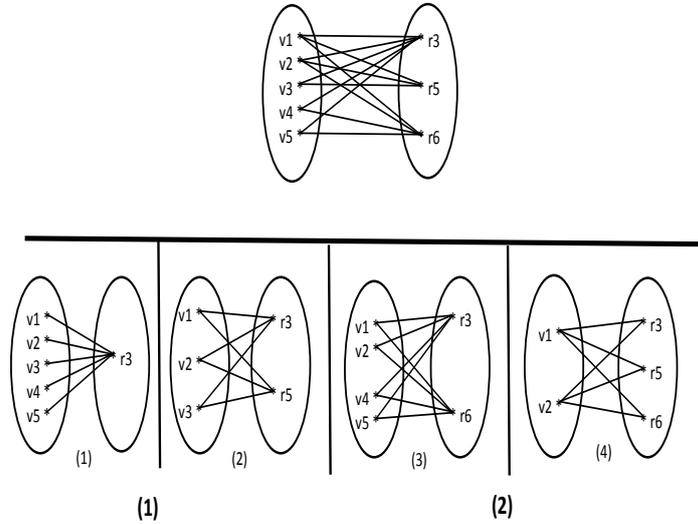


Figure 4: **(Up)**  $PC_{v_1r_3}$ : Pseudo-concept associated to couple  $(v_1, r_3)$ , **(Bottom)** Set of formal concepts extracted from  $PC_{v_1r_3}$ : **(1)**  $\langle \{v_1v_2v_3v_4v_5\}, \{r_3\} \rangle$ ; **(2)**  $\langle \{v_1v_2v_3\}, \{r_3r_5\} \rangle$ ; **(3)**  $\langle \{v_1v_2v_4v_5\}, \{r_3r_6\} \rangle$ ; **(4)**  $\langle \{v_1v_2\}, \{r_3r_5r_6\} \rangle$

The main thrust of the notion of formal concepts stands for the fact that they vehicles 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 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 interested vehicles.

Interestingly enough, the computation of ZORs comes back to the extraction of 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 [5]. In the following, we present our approach for the determination of ZORs, which is an instantiation of that was presented in [5].

### 3.2.2. 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, consequently having a reduced intent part. In the ZORFINDER algorithm, we propose to rely on the following selection metric  $\mathcal{M}$ , where  $|\cdot|$  is the cardinality operator :

$$\mathcal{M}(\langle V, R \rangle) = (|V| \times |R|) - (|V| + |R|). \quad (1)$$

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 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 :*

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

*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 of formal concepts including the couple  $(v_1, r_3)$ . With respect to the gain function 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 pertinent formal concepts. As a greedy approach to set 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 pertinent 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  $\mathcal{M}$  is considered as pertinent and is added to the coverage.

**Example 5.** *If we consider the formal context presented in Table 1, the ZOR set denoted  $ZOR_{\mathcal{M}}$  associated to the selection metric  $\mathcal{M}$  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  $ZOR_{\mathcal{M}}$ , including  $r_i$  in their intent respective parts. Formally,  $ZOR(r_i) = \{\cup r_j \mid \langle V, R \rangle \in ZOR_{\mathcal{M}} \wedge r_j, r_i \in R\}$ . For example, we have  $ZOR(r_2) = \{r_1r_2r_6r_7\}$ . In the following, we present a thorough description of of the DPMS protocol.

### 3.3. The protocol

The flowchart diagram Figure 5 depicts the different operations of the DPMS protocol. Indeed, DPMS is permanently listening to three different events then it executes a set of operations when 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  $SP(c)$ . In case where  $v$  does not have  $SP(c)$ , it downloads it from the server then it stores it locally. Indeed, vehicles only need to download the split maps of not previously visited cities. After that,  $v$  must cheek if it enters in a new region  $r$  of the city  $c$ . In this case, it runs the *SearchForPertinentEvents* operation in order to retrieve pertinent 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 interested vehicles. Furthermore, the pertinent events are kept alive inside the region, without the need for rebroadcasting event messages as the existing geocast approaches do.

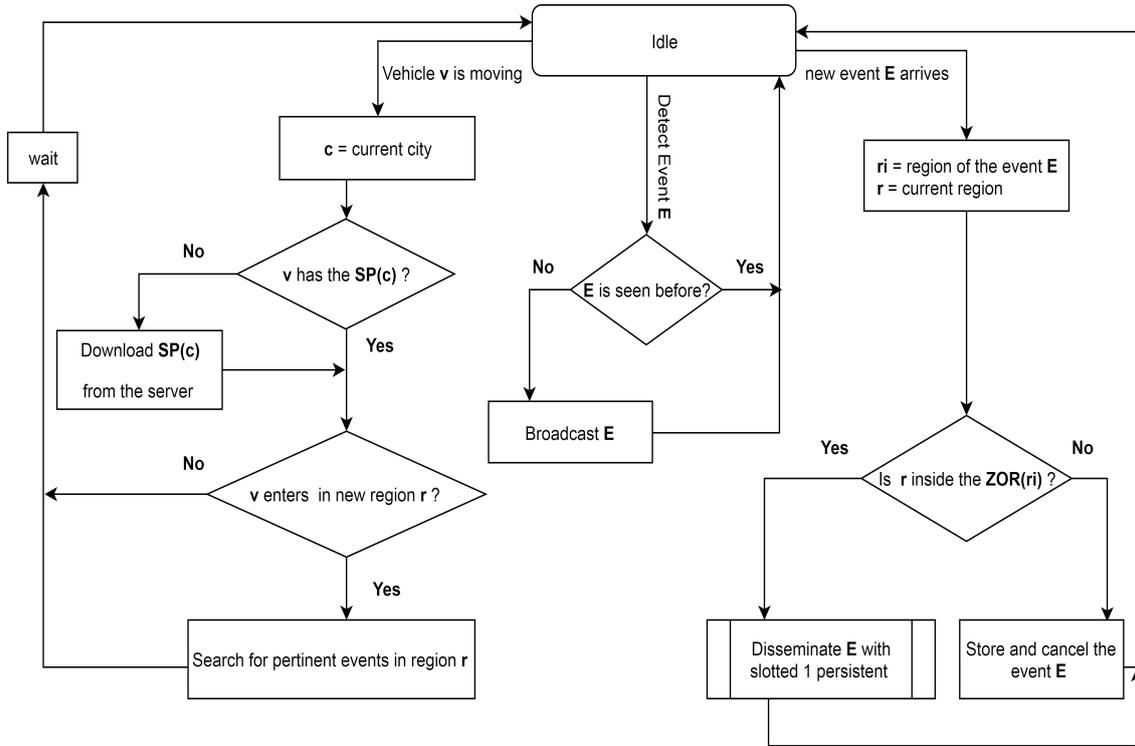


Figure 5: Protocol Flowchart Diagram

2. *Event detected*: Whenever a vehicle detects a new event, it generates an event message according to following structure [sender ID, message ID, event region ID, event type, event time]<sup>2</sup>. 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 was happened, from the event message. Thereafter, it checks if the current region  $r$  of the vehicle is belong to the zone of relevance  $ZOR(r_i)$  of the event region  $r_i$ . In this case, the Slotted 1-Persistence suppression technique will be 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, within 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-broadcasts the event if it has not received any duplicate packet during the waiting time slot; otherwise, it will discard it. Given all of the relative distance between the vehicles  $i$  and  $j$  ( $D_{ij}$ ), the average transmission range  $R$  and the predetermined number of slots  $N_s$ ,  $T_{S_{ij}}$  is computed as follows:

<sup>2</sup>The size of the message is less than 2321 Bytes, and the maximum allowed size of the message is 802.11p standard [20]

$$T_{S_{ij}} = N_s(1 - [\frac{\min(D_{ij}, R)}{R}]) \times \tau \quad (2)$$

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

#### 4. Performance evaluation

In this section, we present the performance evaluation of our proposal versus the Slotted 1-Persistence [2]. We choose the Slotted 1-Persistence as a baseline approach thanks to its high reachability ratio [10] [9]. Worth of mention that in a sake of leading a fair comparison, we limit 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. [14]). 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. Thus, we assume that a region is defined as a set of connected roads. 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 in consideration the granularity of the concept or 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 [17]) 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.

##### 4.1. Simulation settings

The network simulation is performed by OMNeT++ [21] along with the physical layer modeling toolkit MiXiM<sup>3</sup>, 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 [22], 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<sup>4</sup>. 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 2. In the MAC layer, we set the transmission power to  $40mW$  to achieve approximately  $300m$  of interference range. In 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.

<sup>3</sup><http://mixim.sourceforge.net/>

<sup>4</sup>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/>

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 <sup>2</sup>
Data message size	2313 bytes
Data message frequency	0.5 Hz

Table 2: Simulation settings

#### 4.2. 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 6, representing portions of the urban areas of Corniche, Zarzouna and Menzel Bourguiba cities. Table 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: Road topologies characteristics

#### 4.3. 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.

##### 4.3.1. 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[14]). Formally, the reachability metric is defined as follows:

$$Reachability(e) = \frac{|IIV|}{|IV|} \quad (3)$$

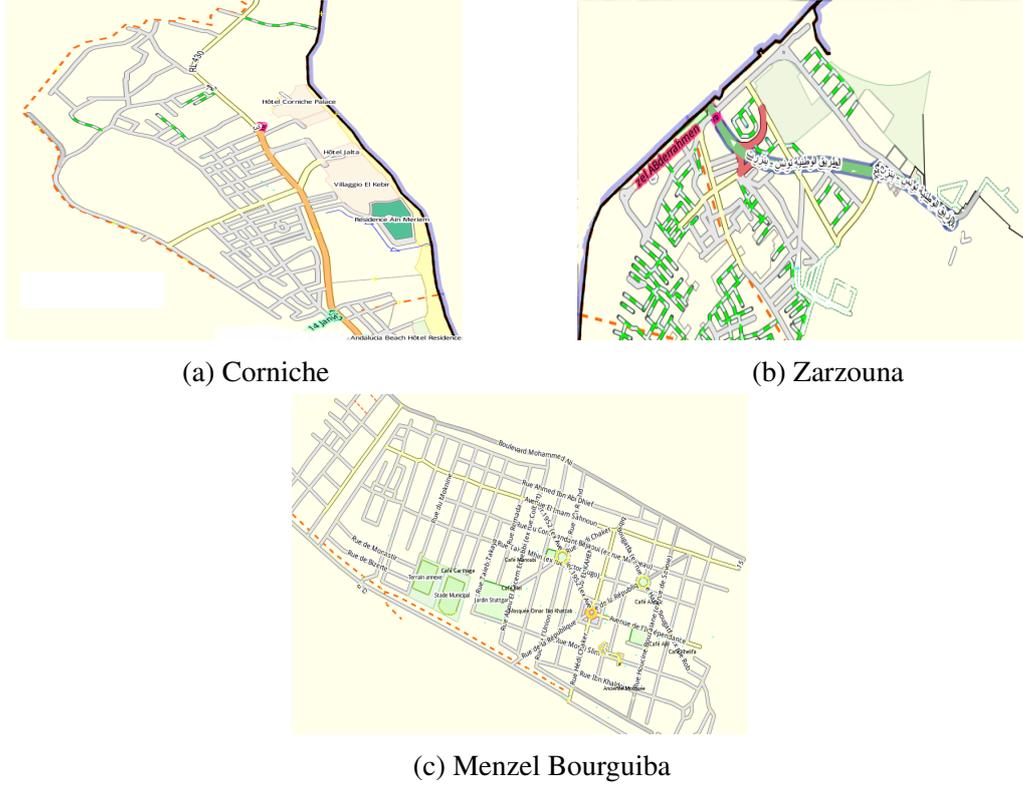


Figure 6: Example of different road topologies

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

$$AverageReachability = \frac{\sum Reachability(e)}{NumberOfEvents} \quad (4)$$

**Precision:** This metric assesses to what extent the protocol is able to only inform pertinent vehicles that are actually interested in 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 the determination of the geocasting area <sup>5</sup>. Formally, the precision metric is defined as follows:

$$Precision = \frac{|IIV|}{|AIV|} \quad (5)$$

where  $IIV$  stands for the set of interested informed vehicles, *i.e.* only pertinent vehicles for an event  $e$ , and  $AIV$  stands for the set of all informed vehicles, *i.e.* pertinent as well as not pertinent vehicles for an event  $e$ . The average precision is defined as follows:

$$AveragePrecision = \frac{\sum Precision(e)}{NumberOfEvents} \quad (6)$$

<sup>5</sup>Geocast is a special case of multicast where data should be only disseminated to a special geographic area.

**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.*,

$$F - score = \frac{2 \times (Precision \times Reachability)}{Precision + Reachability} \quad (7)$$

#### 4.3.2. 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 in the remainder.

**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 [10] 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:

$$AL = \frac{\sum(t_i - T)}{NumberOfInterestedVehicles} \quad (8)$$

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.

### 4.4. Results

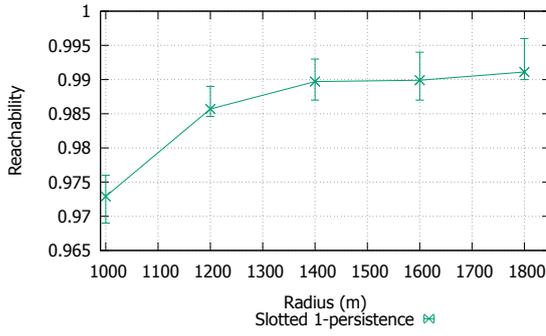
#### 4.4.1. Effectiveness and Efficiency of Slotted 1-persistence under different values of ZOR radius

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.

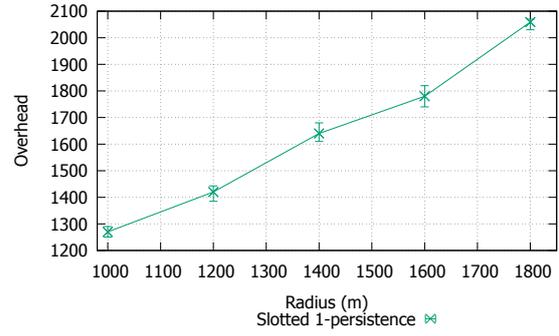
Figure 7, using the map of Menzel Burguiba city, depicts that the reachability and the overload of Slotted 1-persistence grow as far as the radius of the ZOR increases. Indeed, we observe that within a radius of 1400m the reachability becomes stable (around 0.99). However, the overload sharply rises 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 tradeoff between effectiveness and efficiency. Hence, in the remainder, we set the ZOR radius of Slotted 1-persistence to 1400m and we compare its effectiveness and efficiency against our protocol.

#### 4.4.2. Effectiveness of dissemination protocol

Figures 8, 9 and 10 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, 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

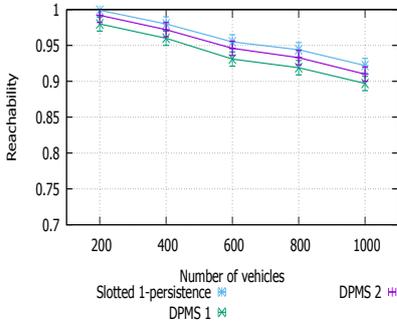


(a) Reachability

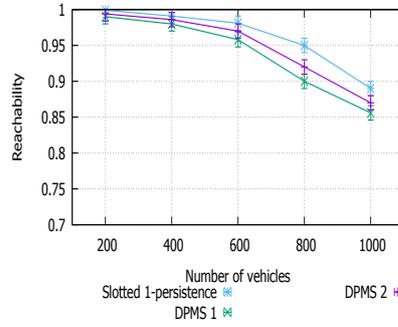


(b) Overload

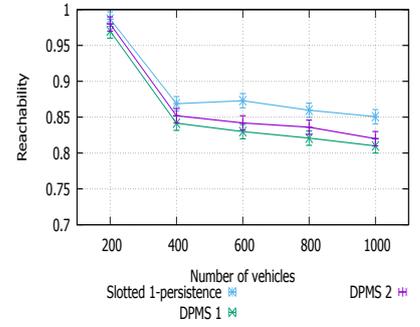
Figure 7: Variation of the average Reachability and Overload of Slotted 1-persistence w.r.t the variation of ZOR's radius



(a) Corniche

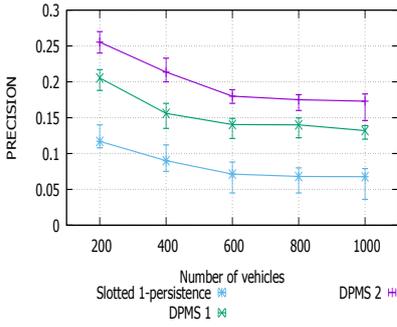


(b) Zarzouna

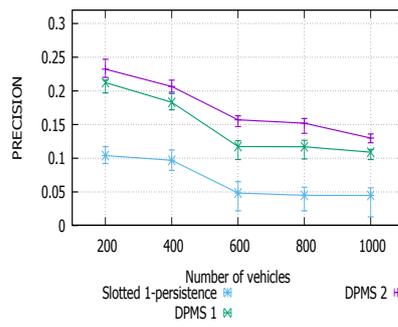


(c) Menzel Bourguiba

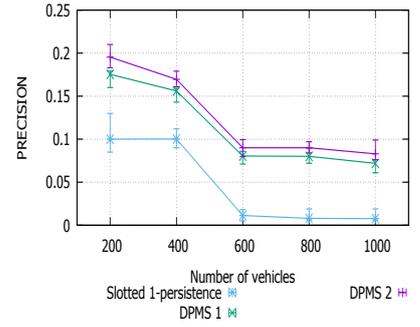
Figure 8: Variation of the average Reachability w.r.t the variation of the number of vehicles under different maps



(a) Corniche



(b) Zarzouna



(c) Menzel Bourguiba

Figure 9: Variation of the average Precision w.r.t the variation of the number of vehicles under different maps

is slightly sharper than our protocol (i.e., Figure 8). 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 1400m will include the whole

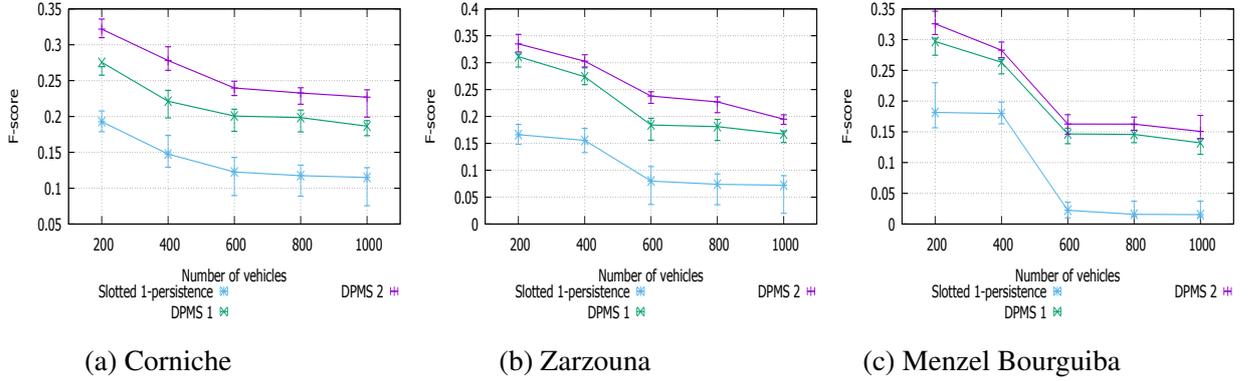


Figure 10: Variation of the F-score w.r.t the variation of the number of vehicles

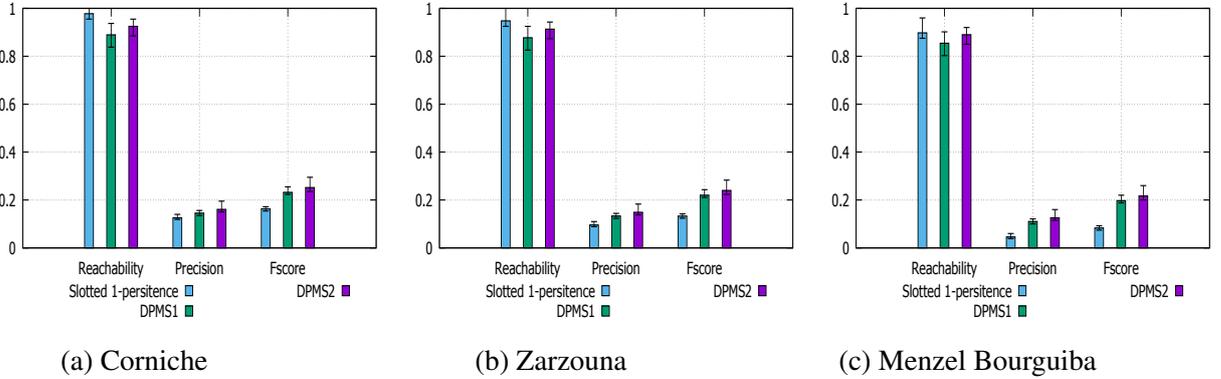


Figure 11: Average Reachability, Precision and F-score

map in most cases. Hence, a higher number of vehicles are got in touch, 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 9 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% (c.f., Figure 11). This encouraging performances are owed to the fact that our protocol closely matches the ZOR. Therefore, within our smart ZOR determination method, a less number of non interested vehicles receive the events leading to a higher precision than the circular ZORs do. In addition, Figure 10 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% (c.f., Figure 11).

Studying the impact of the road topology on the effectiveness of the dissemination protocols is of paramount importance. Indeed, Figures 8, 9 and 10 show that the effectiveness of the different protocols varies slightly according to the road topology. We observe that the reachability, the precision and the F-score are more better within the map of the city that has a less number of roads and junctions. This is owing to the fact that in a such 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,  $DPMS_2$  increases respectively the average reachability, precision and F-score of  $DPMS_1$  by around 1.3%, 11% and 10% (c.f. Figure 11). This can simply be explained by the fact that in  $DPMS_2$ , we associate for each road its associated ZOR (i.e., since each region is considered as a road), which 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 determination process giving birth to challenging scalability issues.

#### 4.4.3. Efficiency of dissemination protocol

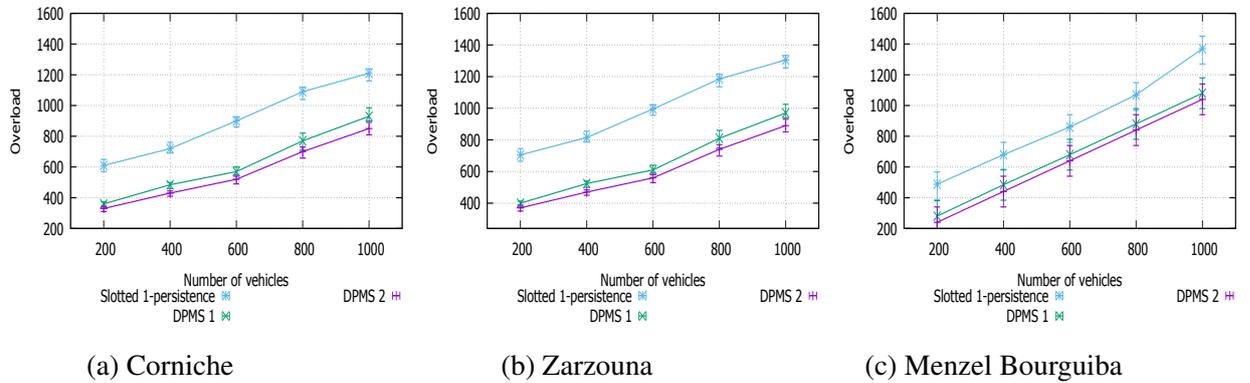


Figure 12: Variation of the Overload w.r.t the variation of the number of vehicles

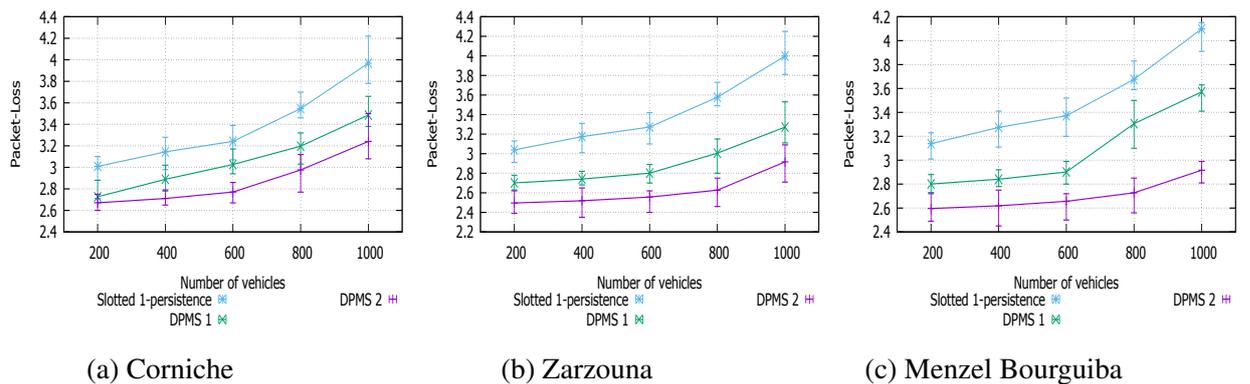


Figure 13: Variation of the Packet loss w.r.t the variation of the number of vehicles

Figures 12, 13 and 14 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

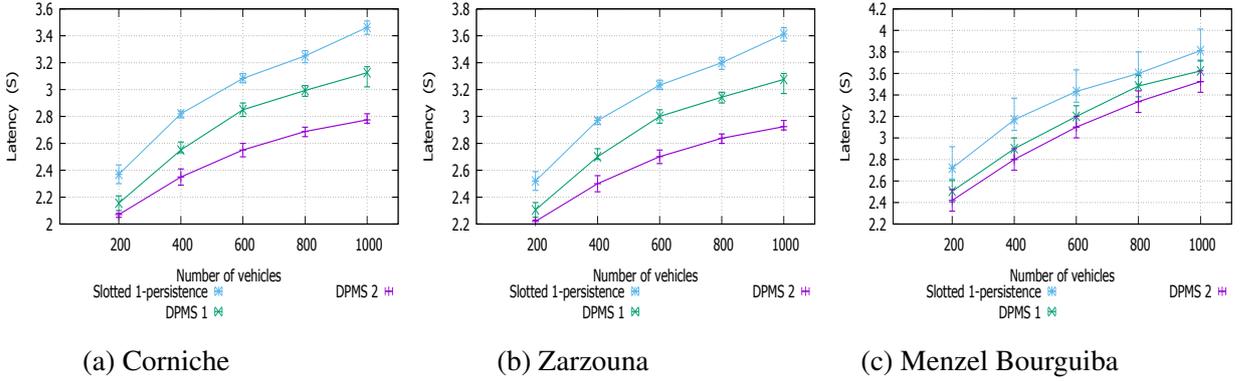


Figure 14: Variation of the average Latency w.r.t the variation of the number of vehicles

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 12 and Figure 13 demonstrate that our protocol flags out a minimum network overload as well as a packet loss than the Slotted-1 persistence protocol do 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 algorithm of the ZOR determination, the dissemination area is too narrow than 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 pertinent 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.

Finally, Figures 12, 13 and 14 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.

## 5. Conclusions and future work

In this paper, we have introduced DPMS as a MAP-splitting-based dissemination protocol to exchange information about events in a VANET. 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. Avenues of future work are as follows:

1. The scalability issue is of paramount importance 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 in the aim of carrying extensive experiments by considering a higher number of vehicles. In addition, we will also consider the incremental maintenance of the ZORs building.

2. Introduce a data aggregation mechanism that will 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 the ZOR. Other vehicles in the traffic jam also start generating such warning messages, which undoubtedly leads to a local broadcast storm.
3. Integrate 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.

## References

- [1] K. Ibrahim, M. Weigle, M. Abuelela, p-ivg: Probabilistic inter-vehicle geocast for dense vehicular networks, in: Vehicular Technology Conference (VTC), 2009, pp. 1–5.
- [2] H. Rahbar, K. Naik, A. Nayak, DtsG: Dynamic time-stable geocast routing in vehicular ad hoc networks, in: Proc. of the Intl. 9th IFIP Annual Mediterranean Ad Hoc Networking (Med-Hoc-Net), 2010, pp. 1–7.
- [3] P. Kheawchaom, S. Kittipiyakul, K. S. N. Ayutayay, idtsG time-stable geocast for post crash notification in vehicular highway networks, in: Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2012, pp. 1–4.
- [4] S. Allal, S. Boudjit, Geocast routing protocols for vanets: Survey and guidelines, in: Proc. of the Intl. Conf. on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012, pp. 323–328.
- [5] A. Mouakher, S. B. Yahia, Qualitycover: Efficient binary relation coverage guided by induced knowledge quality, *Information Sciences* 355-356 (2016) 58–73.
- [6] J. Luo, X. Gu, T. Zhao, W. Yan, A mobile infrastructure based vanet routing protocol in the urban environment, in: Proc. of the Intl. conference on Communications and Mobile Computing (CMC), Vol. 3, 2010, pp. 432–437.
- [7] C. Jayapal, S. Roy, Road traffic congestion management using vanet, in: Proc. of the International Conference on Advances in Human Machine Interaction (HMI), 2016, pp. 1–7.
- [8] B. Tian, K. M. Hou, J. Li, 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), 2016, pp. 724–731.
- [9] C. Sommer, F. Dressler, Information Dissemination in Vehicular Networks, in: Vehicular Communications and Networks: Architectures, Protocols, Operation and Deployment, Woodhead, 2015, pp. 75–93.
- [10] N. Wisitpongphan, O. Tonguz, J. Parikh, P. Mudalige, F. Bai, V. Sadekar, Broadcast storm mitigation techniques in vehicular ad hoc networks, *Wireless Communications, IEEE* 14 (6) (2007) 84–94.
- [11] R. S. Schwartz, H. Scholten, P. Havinga, A scalable data dissemination protocol for both highway and urban vehicular environments, *EURASIP Journal on Wireless Communications and Networking* 2013 (2013) 1–19.
- [12] B. Bako, M. Weber, Efficient Information Dissemination in VANETs, in: *Adv. Veh. Netw. Technol.*, 2011, p. 20.
- [13] X. Li, H. Li, A survey on data dissemination in vanets, *Chinese Science Bulletin* 59 (32) (2014) 4190–4200.
- [14] T. Jochle, B. Wiedersheim, F. Schaub, M. Weber, Efficiency analysis of geocast target region specifications for vanet applications, in: Proc. of the Intl. Vehicular Networking Conference (VNC), 2012, pp. 250–257.
- [15] T. Delot, S. Ilarri, N. Cenerario, T. Hien, Event sharing in vehicular networks using geographic vectors and maps, *Mob. Inf. Syst.* 7 (2011) 21–44.
- [16] B. Alsubaihi, A. Boukerche, Semantic and self-decision geocast protocol for data dissemination over vanet (sas-gp), in: *Wireless Communications and Networking Conference (WCNC)*, 2015, pp. 1948–1953.
- [17] J. Kenney, Dedicated short-range communications (dsrc) standards in the united states 99 (7) (2011) 1162–1182.
- [18] G. Bernhard, W. Rudolfe, *Formal Concept Analysis.*, 1999.

- [19] M. Barbut, B. Monjardet, *Ordre et classification. Algèbre et Combinatoire*, 1970.
- [20] C. Sommer, R. German, F. Dressler, Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis, *IEEE Transactions on Mobile Computing* 10 (1) (2011) 3–15.
- [21] V. Andras, Omnet++, in: Wehrle, Klaus, Gijnes, Mesut, Gross, James (Eds.), *Modeling and Tools for Network Simulation*, 2010, pp. 35–59.
- [22] F. Kaiser, C. Gransart, M. Berbineau, Simulations of vanet scenarios with opnet and sumo, in: *Communication Technologies for Vehicles*, Vol. 7266 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2012, pp. 103–112.