

# Data aggregation in VANETs: the VESPA approach

Bruno Defude  
TELECOM & Management  
SudParis  
9 rue Charles Fourier  
91011 EVRY Cedex - France  
Bruno.Defude@int-  
edu.eu

Thierry Delot  
LAMIH - UMR CNRS 8530  
University of Valenciennes  
Le Mont Houy  
59313 Valenciennes - France  
tdelot@univ-  
valenciennes.fr

Sergio Ilarri  
IIS Dept  
University of Zaragoza  
Maria de Luna 1  
Zaragoza, 50018 - Spain  
silarri@unizar.es

Jose-Luis Zechinelli  
Martini  
CENTIA  
University of las Américas  
Puebla - Mexico  
joseluis.zechinelli@udlap.mx

Nicolas Cenerario  
LAMIH - UMR CNRS 8530  
University of Valenciennes  
Le Mont Houy  
59313 Valenciennes - France  
ncenerar@univ-  
valenciennes.fr

## ABSTRACT

VESPA (Vehicular Event Sharing with a mobile P2P Architecture)<sup>1</sup> is a system designed for vehicles to share information in vehicular ad-hoc networks (VANETs). The originality of VESPA is to support any type of event (e.g., available parking spaces, accidents, emergency braking, information relative to the coordination of vehicles in emergency situations, etc.). The basic functions of VESPA are both disseminating the events to the potentially interested vehicles and evaluating their relevance once received (for instance to determine whether the driver should be warned or not).

In this paper, we concentrate on knowledge extraction in VESPA. In particular, we focus on how to exploit data exchanged between vehicles to produce knowledge to be used later on by drivers. Existing systems only use the data exchanged to produce a warning for the drivers when it is needed. Then, the data is considered obsolete and is deleted. On the contrary, we propose here to aggregate the data once it becomes obsolete. Our objective is to produce additional knowledge to be used by drivers when no relevant data has been communicated by neighbouring vehicles. For example, it becomes so possible to dynamically detect potentially dangerous road segments or to determine the areas where the probability to find an available parking space is high.

## General Terms

VANETs, Data Storage, Data Aggregation

<sup>1</sup>See: <http://www.univ-valenciennes.fr/ROI/SID/tdelot/vespa/> for more information.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. MobiQuitous 2008, July 21-25, 2008, Dublin, Ireland.

Copyright ©2008 ICST ISBN 978-963-9799-27-1

## 1. INTRODUCTION

Today, the car is indisputably the most heavily used mode of transportation. Unfortunately, its popularity has been accompanied by numerous problems, for example, in the areas of safety and the environment. In spite of significant efforts to reduce the number of persons dying on the road, this number remains quite high, mainly due to the human factor (e.g., accident-prone behaviour or low response time). To reduce the number of accidents, a variety of programs, generally involving “Intelligent Transport Systems”, have been initiated in Japan, Europe and the United States, attracting the interest of researchers both in academia and in industry.

Our works concentrate on data management in inter-vehicle ad hoc networks. These wireless networks rely on the use of short-range networks (about a hundred meters), like IEEE 802.11 or Ultra Wide Band (UWB) standards for vehicles to communicate [13] and provide bandwidth in the range of Mbps. Using such communication networks, the driver of a car can receive information from its neighbours. Many pieces of information may be exchanged in the context of inter-vehicle communications, for instance to warn drivers when a potentially dangerous event arises (accident, emergency braking, obstacle on the road, etc.) or to try to assist them (available parking spaces, traffic congestions, real-time traffic conditions, etc.).

These last years, numerous research works have investigated how to disseminate data in inter-vehicle networks and how to evaluate their relevance. In this paper, our approach is quite different. Contrary to existing solutions, we do not aim at considering data exchanged between vehicles only relevant to produce a warning to the drivers. Once considered as no relevant any more, for example because it has been already used to warn the driver or evaluated as obsolete, we aim at aggregating the data in order to produce additional knowledge to be used later on. This enables, among other things, the real-time detection of the potentially dangerous areas on the roads (for example becoming dangerous because of bad weather conditions) or to determine places with a high probability to find an available parking space when none has

been communicated by other vehicles.

Two approaches can be used to aggregate events, a centralized one or a decentralized one. In centralized approaches, vehicles send their events to a central server when they have access to a network. The server has powerful resources and can construct large and precise summaries. This approach does not consume local resources but it supposes that vehicles have good connectivity (to send their events and to query the server when needed). It is also difficult at the server side to construct all aggregates needed by vehicles. In a decentralized approach, vehicles construct their own aggregates. Of course these aggregates are less precise than in centralized approaches but are exactly those needed by the vehicle. Vehicles do not require to have good connectivity for constructing or querying summaries. In this paper we present a fully decentralized approach improved by aggregate exchanges between vehicles.

The rest of this paper is organized as follows. Section 2 presents some related works. Section 3 is dedicated to data aggregation. More precisely, it presents our approach as well as the structures and methods used in order to aggregate information. Finally, section 4 offers our conclusions and gives some ideas for prospective research.

## 2. RELATED WORK

In this section, we review some related works. First, we consider some interesting projects in the field of vehicular networks in general and then we focus on data aggregation.

### 2.1 Inter-vehicle Communications

Inter-vehicle communication is a recent field of research. Nevertheless, some studies have already made significant contributions. In this section, we review some of these works.

The FleetNet project (2000-2003) [10], followed later by the NoW project<sup>2</sup> (finishing at the end of 2007), and CarTalk [15, 3] (2001-2004) worked to exploit inter-vehicle communications to make driving safer. Both FleetNet and CarTalk used multi-hop communication techniques, but while FleetNet was supported by a partial fixed infrastructure [11], CarTalk used no existing infrastructure. Those are very interesting projects which focus mainly on the network level of inter-vehicle communication. They use Geocast communication protocols [14, 7], which allow to determine the geographical area where messages have to be conveyed, either through direct inter-vehicle communication or through multi-hop communication.

Regarding data dissemination protocols, there are several projects that are worth mentioning. In the context of the *TrafficView* system [16, 17], the authors present different dissemination protocols for periodic broadcasting of data. In [1, 2], the authors study the way to determine the *dissemination area*, which is the area where the data should be broadcast. In [27], a *carry and forward* strategy is proposed, which aims at keeping the information stored on the car until it can be transmitted to another one. Finally, in the *Mobi-Dik* project [25], which considers the spatio-temporal relevance of data, a vehicle with a certain piece of information

<sup>2</sup><http://www.network-on-wheels.de/>

acts as a disease carrier, and “contaminates” the nearby vehicles along its route; a similar approach is proposed in [18].

Summing up, the main goal of existing V2V communication solutions is to limit the number of messages exchanged to avoid overloading the network, which is indeed crucial if the correct functioning of the applications is to be guaranteed. The existing protocols and dissemination techniques mentioned above are generally interesting. Nevertheless, they only focus on how and when the information relative to events (accident, obstacle, etc.) or resources (available parking slots, etc.) should be disseminated to other vehicles. Existing solutions are all specific to a particular type of event (accidents, available parking slots, etc.), presented as an application example, and thus cannot support other event types. For instance, *Mobi-Dik* provides a very interesting solution for the problem of information-sharing inside a restricted spatio-temporal area. Although these techniques are very well adapted for sharing information between cars about available parking spaces, they cannot be exploited to relay information about an accident or an emergency braking situation. Thus, it is really difficult to compare existing solutions since they are well adapted to a particular type of event but cannot support other types of events. Since it is really difficult to imagine embedded systems in cars dealing with only one type of event and not with the others, VESPA aims at proposing a single data sharing approach valid for all types of events, in order to deploy a generic system into cars. Therefore, VESPA uses an encounter probability to determine the relevance of an event [5, 8].

### 2.2 Data Aggregation in Vehicular Networks

Data aggregation has been proposed as a way of optimizing bandwidth/storage usage or as a way of managing high-level information when fine-granularity information is not required. Many data aggregation strategies have been developed in the context of sensor networks (e.g., see [6, 20]), where the main purpose is to reduce energy consumption. However, the high mobility of vehicles and the large number of vehicles that may be present in a geographical area render these strategies very difficult to apply in the context of vehicular networks. Moreover, energy consumption is not an issue in vehicular networks. In the following, we indicate some works related to data aggregation in the specific context of vehicular networks:

- In [21], the *Region-based Location Service Management Protocol (RLSMP)* is proposed, which considers the use of message aggregation (i.e., information is aggregated into larger packets) and geographical clustering to minimize the number of location updates and querying messages for location management of vehicles. The authors indicate that, although message aggregation will improve scalability, it can also lead to: 1) more packet collisions and retransmissions (because larger packets are transmitted), and 2) longer delays (as the information must be buffered for aggregation before being sent).
- In [19], the focus is on security aspects, particularly on how to detect attackers that purposely disseminate erroneous aggregated records. Aggregation can alleviate the bandwidth problem but also makes security

more challenging. The proposed solution is based on the use of a tamper-proof service on vehicles and on asking an aggregator for a randomly chosen original aggregated record. The authors distinguish between *syntactic aggregation* (which mainly reduces the overhead of message headers) and *semantic aggregation* (which is information-specific and saves more bandwidth at the cost of loss of information), focusing on the first type of aggregation. The solution presented in the paper, however, cannot be used (according to the authors) for events such as accidents (only for information about cars, such as speeds and locations). Similarly, re-aggregation is not considered and the omission of records by malicious cars (to compute an aggregate) cannot be detected.

- In [16], the authors distinguish between *data compression* and *data aggregation*: Only data aggregation considers the semantics of the data. Several algorithms for data aggregation are proposed. The *ratio-based aggregation algorithm* considers a division of the road in front of the vehicle in segments, each associated to a certain *aggregation ratio*. On the other hand, the *cost-based aggregation algorithm* considers the cost (based on the error introduced during the merging, the number of vehicles affected by the aggregation, etc.) of aggregating records to merge pairs of records with minimum cost.
- In [9], they consider vehicles which aggregate *warning data* if they have several messages describing the same event. They additionally propose the use of *revocation messages* when a vehicle does not detect a hazard when entering an area which is dangerous according to a stored aggregate.
- Hierarchical data aggregation is studied in [12]. The motivation is that a vehicle needs detailed data about its surroundings but only coarse aggregates about farther areas. An algorithm is proposed, based on the use of a modified *Flajolet-Martin* sketch to store approximate information. This approach provides interesting benefits. For example, it is possible to merge two aggregates (even if they have some overlapping) while avoiding the appearance of duplicates. The aggregation hierarchy is pre-defined in the data map, grouping areas according to their natural relation (e.g., by districts or roads). In the context of spatio-temporal applications, sketches are also proposed in [22] as a means to avoid the *distinct counting problem* for count queries and sum queries.
- In [6], a *Location Based Aggregation (LBAG)* protocol is proposed. In this protocol, the data aggregation relies on a hierarchy of static locations instead of considering a tree-structure of nodes (which would be very challenging to maintain due to the high mobility of vehicles). A geocast protocol (based on position-based routing and optimized flooding) is used to deliver a message to a target area. *Suppression messages* are used to minimize the probability of redundant messages (several vehicles could in principle generate aggregates for the same area).

Other works that consider aggregation in the context of vehicular networks exist, such as [24, 4]. We would also like to mention the work presented in [26], where they propose approaches to manage temporal aggregates with different time granularities. Finally, summaries have also been considered in the context of databases, to handle growth and keep high-level knowledge (e.g., see [23]).

### 3. AGGREGATION OF EVENTS

The main existing solutions related to inter-vehicle communications focus on information exchanges between vehicles. All those solutions only consider a piece of data as an element to diffuse to a set of potentially interested vehicles. When the data is received by a vehicle, its relevance is evaluated, generally using both spatial and temporal criteria. Then, the data is diffused again to neighbouring vehicles and the driver is warned when necessary. Once used, the data produced to describe an event is considered obsolete and deleted. As discussed in Section 2, other works have addressed data aggregation in VANETs. The goal is then to minimize the amount of data to be exchanged between vehicles.

In this paper, our approach is quite different. We consider that the data received by a vehicle to describe an event should not only be used to produce a warning for the drivers. Once stored on a vehicle, it is indeed possible to use the data collected to produce, at the vehicle level, additional knowledge (relative to the environment) that can be exploited by the driver. To illustrate our purpose, let us mention the example of available parking spaces. When no available parking space has been received from neighbouring vehicles, it may be interesting for the driver to know the places where the probability to find one is the highest (according to the day and the hour). In another context, it is possible to use the different messages received about accidents, emergency braking, obstacles in the roads, etc. to dynamically determine the dangerous road segments and indicate them to the driver. Other types of data exchanged between vehicles may be aggregated to generate additional knowledge like for example traffic information.

Thus, a vehicle may produce and receive messages describing road events. In the following, we will describe the structures we use to aggregate the data. We consider an event described as a tuple  $\langle \text{timestamp}, \text{type of event}, \text{location} \rangle$ . When disseminated in the network, the events generally have a more complex representation (values associated with the types of events, direction indication, etc.) that will not be considered here. The aggregation process should verify the following properties:

- It should summarize the events according to the fundamental dimensions that are location and time;
- It should be incremental and the volume of data stored should be limited;
- Each driver should be able to choose the types of events s/he is interested in, as well as the spatial and temporal scales to be used to aggregate the data;
- It should be possible to exchange the aggregated data between vehicles in order to improve their respective

knowledge (even if the aggregated data exchanged have different spatial and/or temporal extents);

- The aggregation process should not be costly.

In the following, we present our approach relying on an aggregation of the number of events relative to the spatial and temporal dimensions as well as the type of events.

### 3.1 Aggregation structure used for the events

In addition to the “active” events received, each vehicle now stores a summary of the previously received events. This summary is structured as follows: each type of event is associated with a two-dimensional (spatial and temporal) matrix. Each *spatio-temporal cell* of that matrix contains a number of events and a confidence value ranging between 0 and 1 (see Figure 1).

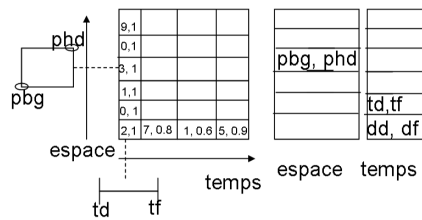


Figure 1: Storage structure of the events

To manage the spatial dimension, we consider the space as split in a set of rectangles. Such a representation does not necessarily cover the whole physical space (a driver is not always interested in the whole space) and the size of the rectangles is not necessarily homogeneous (some areas may require a more detailed information than others). Therefore, each matrix used to aggregate one particular type of event is associated to a table representing the set of defined rectangles. Each cell of the table contains the left low point (llp) and the right high point (rhp) of the rectangle.

As concerns the temporal dimension, two items have to be specified. The first one is the observation window determining the events to aggregate. The time representation may also be non-homogeneous and so has to be specified. A table associated with the type of event defines the chosen temporal representation. It contains the starting time (st) and the ending time (et) as well as the starting date (sd) and the ending date (ed) of the aggregation.

A confidence value indicates whether the information is guaranteed to be certain (the events have really been observed in that cell) or approximate (the events have been observed in a larger spatial area or temporal dimension).

To illustrate how the summaries are created, let us consider the itinerary between a driver’s house and the university where s/he works. Along its displacement, her/his vehicle may receive each day a set of messages transmitted by

neighbouring vehicles and indicating close available parking spaces. An example of summary generated thanks to those messages is described in Figure 1. The first column of the matrix contains 6 aggregations distributed in the space at the first defined temporal interval, representing the aggregation of available parking spaces detected in 6 distinct places (parking at the university, city centre, etc.). All these observations are sure since the confidence value is equal to 1. Each row corresponds to the situation of a cell in different time intervals. For example, the first row presents the aggregation of the events observed on the first spatial rectangle for all the time intervals. The confidence values may then be inferior to 1. Potentially, each matrix may contain an important number of 0 because no occurrence of the event has been observed in that cell. Storage techniques dedicated to sparse matrices may then be used.

The scales (i.e. the lower and upper bounds) used for both the temporal and spatial dimensions may be determined either statically, using for example fixed (spatial and temporal) intervals or dynamically, for example by determining clusters of events according to a (spatial and/or temporal) proximity relationship.

### 3.2 Mapping between cells

The aggregation matrices are used to store and manipulate summaries of known events. The basic operation on these matrices is the reduction of a cell according to another one. The *reduction* operation may be used to facilitate the exchange of summaries between vehicles, to restructure a summary or to evaluate a query on a summary. It consists of the comparison of two cells having potentially different spatio-temporal frames of reference. The main method is:

*function reduction* ( $c_1, c_2$ ) returns *sub-cell*

where  $c_1$  (target cell) and  $c_2$  (cell to be mapped) are two cells of a matrix containing the same type of event. Cell  $c_1$  is the target cell and so  $c_2$  may have to be split in several sub-cells. Among those generated sub-cells, one will have the same spatio-temporal frame of reference as  $c_1$ . More precisely, the following cases have to be considered:

- if  $c_2$  is “included” in  $c_1$  then  $c_2$  is returned;
- if  $c_2$  and  $c_1$  have no intersection then the empty cell is returned;
- if  $c_1$  is “included” in  $c_2$  then  $c_2$  has to be split to be adapted to  $c_1$ . It is thus necessary to create up to 5 cells for the spatial dimension (one shared rectangle, one rectangle at the right, one at the left, one below and one underneath) and 3 for the temporal one (the common interval, the preceding interval and the following interval). The aggregated value of  $c_2$  is then uniformly distributed between all the created cells, since the aggregation process leads to the loss of precise information:

$$value(sub_{c_2}) = \frac{value(c_2)}{nb}, conf(sub_{c_2}) = \frac{conf(c_2)}{nb}$$

where  $nb$  corresponds to the number of cells created. The sub-cell  $sub_{c_2}$  is then returned;

- if  $c_1$  and  $c_2$  have a non-empty intersection then only the intersection area between  $c_1$  and  $c_2$  is considered. The procedure described at the preceding item can then be applied on that area.

### 3.3 Exploiting the aggregations

Each cell of a matrix may be seen as an abstract event comparable to one directly generated by a vehicle, except that it is uncertain. Each observed event has a probability of 1, whereas a cell with an aggregated value of 0 is associated an event probability of 0. For non-empty cells, the probability (i.e. confidence) is estimated using both the number of observed events and the confidence value.

Concerning the needs regarding the search of interesting events, they may be relatively complex, ranging from simple filters on the type of event to complex correlations on several types of events. At this stage, we focus on simple filters. More precisely, we consider that the filter may be defined by a simple query retrieving the number of events observed in a given spatio-temporal area. We therefore use the following functions:

*function query(area1, matrix1) returns (value, conf)*  
which applies the *reduction* function on all the cells of matrix1 to obtain all the sub-cells of matrix1 having a non-empty intersection with area1. The *unioncells* function (defined below) is then used to compute the union of those sub-cells.

*function unioncells(set of cells) returns cell*  
The area defining the result cell is a subset of the area defining the query. The aggregated value is obtained by computing the sum of the values of the different cells. The confidence factor is the weighted average (obtained considering the temporal and spatial extents of the cells) of the confidence factors.

For example, if a driver is searching for an available parking space while none has been communicated by the other vehicles, these functions can be used to determine the interest of the area where the vehicle is (i.e., is there a high probability of finding an available parking space here?). Then, if the area is interesting, a shorter area has to be determined, where the vehicle should go because it is the most interesting one. Otherwise, another big area (farther from the current position of the vehicle) should be identified and the previous functions should be applied again.

### 3.4 Exchange of summaries between vehicles

To increase the knowledge exploitable by a vehicle, one solution consists in exchanging summaries between vehicles. These exchanges may be performed either by using relays (for example, data servers located close to roads with an important traffic) or directly (when the vehicles are close enough during a sufficiently large period of time). Thus, each vehicle may decide to publish the summaries it manages to make them available to other vehicles. Reciprocally, a vehicle may be interested in subscribing to some of the summaries published by the others. To simplify, we only consider here public subscriptions and publications, that is the vehicles publish/subscribe towards all the other vehicles.

The publication implies defining which summaries are published (possibly by aggregating them). The goal of the aggregation is here to regroup cells. The subscription requires defining filters to select only the interesting types of events, and sometimes also to restrict that set of interesting events to specific spatio-temporal windows. For example, if a Parisian driver wants to be warned when s/he encounters a dangerous area, it may be relevant for the vehicle to collect the summaries with type “Accident” relative to Paris published by the other vehicles.

The first step of the exchange phase consists in a comparison between the publications of a vehicle A and the subscriptions of a vehicle B. If at least one of the matrices matches, the fusion of matrices A and B is calculated. The main methods are:

*procedure fusionmatrix(m1, m2)*

$m_1$  is the target matrix in the fusion process. All the cells of matrix  $m_1$  are successively compared with those of  $m_2$  using the *reduction* function. If the value returned is different from the empty cell, the *cellfusion* function is applied on the corresponding cell of  $m_2$  and the sub-cell returned by reduction.

*function cellfusion(c1, c2) returns c3*

where  $c_1$  and  $c_2$  have the same spatio-temporal frame of reference. The aggregated values of both cells have to be merged and the new confidence factor has to be determined. The events related to  $c_2$  could be integrated into  $c_1$  if they do not already appear in  $c_1$ . Since no more precise information is available, we approximate the aggregated value with a reduction by using the max of both values<sup>3</sup>:

$$value(c_3) = \max(value(c_1), value(c_2));$$

$$conf(c_3) = \frac{value(c_1) * conf(c_1) + value(c_2) * conf(c_2)}{value(c_1) + value(c_2)}$$

In order to avoid exchanging again summaries when they have not been updated since the last exchange, we maintain timestamps on each vehicle. So, each vehicle  $v$  maintains locally and for each summary matrix (i.e. each type of event  $e$ ), the timestamp of the last update  $LT(v, e)$ , together with its type (local or by exchange). Furthermore it maintains the timestamp  $ET(v, v_i, e)$  of the last exchange for all known vehicles  $v_i$  (or at least the  $k$  more recent ones). Thus, a vehicle  $V_1$  can decide whether it is interesting or not to consider the summary of a vehicle  $V_2$  for an event  $e$ , comparing  $LT(V_2, e)$  with  $ET(V_1, V_2, e)$ .

Obviously, the amount of data exchanged for the summaries is much more important than for classical events. It is so really important not to saturate the bandwidth, which explains the publication/subscription process used. Besides, it is generally more interesting to exchange summaries rather than query results obtained by a particular vehicle using its aggregated data because each vehicle can query the summary according to the drivers' needs.

<sup>3</sup>It would be possible to obtain a more precise approximation by defining an interval  $[\max(value(c_1), value(c_2)), value(c_1) + value(c_2)]$ .

## 4. CONCLUSION & PERSPECTIVES

In this paper, we have presented the foundations of our solution based on data aggregation in VANETs to generate additional knowledge for drivers. Our approach allows the exchange of summaries between vehicles in addition to the one of information about traditional events (emergency braking, accidents, etc.). Our actual work consists in defining scenarios in order to experimentally evaluate our solution and refine the choice of scales used for the generation of summaries.

## 5. REFERENCES

- [1] C. Adler. *Information Dissemination in Vehicular Ad Hoc Networks*. PhD thesis, University of Munchen, 2006.
- [2] C. Adler and M. Strassberger. Putting together the pieces - a comprehensive view on cooperative local danger warning. In *13th IST World Congress on Intelligent Transport Systems and Services*, 2006.
- [3] D. D. Bruin, J. Kroon, R. Klaveren, and M. Nelisse. Design and test of a cooperative adaptive cruise control system. In *Intelligent Vehicles Symposium (IV'04)*, 2004.
- [4] M. Caliskan, D. Graupner, and M. Mauve. Decentralized discovery of free parking places. In *3rd international workshop on Vehicular ad hoc networks (VANET'06)*, pages 30–39, New York, NY, USA, 2006. ACM.
- [5] N. Cenerario, T. Delot, and S. Ilarri. Dissemination of information in inter-vehicle ad hoc networks. In *Intelligent Vehicles Symposium (IV'08)*, 2008.
- [6] C. Chen. Location-based data aggregation in mobile ad hoc networks. Master's thesis, Institut für Parallele und Verteilte Systeme, Abteilung Verteilte Systeme, 2003.
- [7] A. Cheng and K. Rajan. A digital map/GPS based routing and addressing scheme for wireless ad hoc networks. In *Intelligent Vehicle Symposium (IV'03)*, 2003.
- [8] T. Delot, N. Cenerario, and S. Ilarri. Estimating the relevance of information in inter-vehicle ad hoc networks. In *IEEE International Conference on Mobile Data Management (MDM'08) - Workshops*, 2008.
- [9] S. Eichler, C. Merkle, and M. Strassberger. Data aggregation system for distributing inter-vehicle warning messages. In *31st IEEE Conference on Local Computer Networks*, pages 543–544. IEEE Computer Society, November 2006.
- [10] A. Festag, H. Füßler, H. Hartenstein, A. Sarma, and R. Schmitz. Fleetnet: Bringing car-to-car communication into the realworld. In *World Congress on Intelligent Transport Systems (ITS)*, 2002.
- [11] C. Lochert, H. Hartenstein, J. Tian, H. Füßler, D. Hermann, and M. Mauve. A routing strategy for vehicular ad hoc networks in city environments. In *IEEE Intelligent Vehicles Symposium*, 2003.
- [12] C. Lochert, B. Scheuermann, and M. Mauve. Probabilistic aggregation for data dissemination in VANETs. In *4th ACM International Workshop on Vehicular Ad Hoc Networks (VANET'07)*, pages 1–8, New York, NY, USA, 2007. ACM.
- [13] J. Luo and J.-P. Hubaux. A survey of research in inter-vehicle communications. In *Embedded Security in Cars - Securing Current and Future Automotive IT Applications*. Springer-Verlag, 2005.
- [14] C. Maihofer. A survey of geocast routing protocols. *IEEE Communications Surveys & Tutorials*, 6(2):32–42, 2004.
- [15] P. Morsink, R. Hallouzi, I. Dagli, C. Cseh, L. Schafers, M. Nelisse, and D. D. Bruin. Cartalk 2000: Development of a cooperative adas based on vehicle-to-vehicle communication. In *Intelligent Transport Systems and Services*, 2003.
- [16] T. Nadeem, S. Dashtinezhad, C. Liao, and L. Iftode. TrafficView: Traffic data dissemination using car-to-car communication. *ACM Sigmobile Mobile Computing and Communications Review, Special Issue on Mobile Data Management*, 8(3):6–19, 2004.
- [17] T. Nadeem, P. Schankar, and L. Iftode. A comparative study of data dissemination models for VANETs. In *3rd Annual Int. Conf. on Mobile and Ubiquitous Systems*, pages 1–10, 2006.
- [18] S. Nittel, M. Duckham, and L. Kulik. Information dissemination in mobile ad-hoc geosensor networks. In *3rd Int. Conf. on Geographic Information Science*, 2004.
- [19] F. Picconi, N. Ravi, M. Gruteser, and L. Iftode. Probabilistic validation of aggregated data in vehicular ad-hoc networks. In *3rd International Workshop on Vehicular Ad Hoc Networks (VANET'06)*, pages 76–85, New York, NY, USA, 2006. ACM.
- [20] R. Rajagopalan and P. Varshney. Data-aggregation techniques in sensor networks: a survey. *IEEE Communications Surveys & Tutorials*, 8(4):48–63, 2006.
- [21] H. Saleet and O. Basir. Location-based message aggregation in vehicular ad hoc networks. In *2007 IEEE Global Communications Conference (Globecom) Workshops*, pages 1–7. IEEE Computer Society, November 2007.
- [22] Y. Tao, G. Kollios, J. Considine, F. Li, and D. Papadias. Spatio-temporal aggregation using sketches. In *20th International Conference on Data Engineering (ICDE'04)*, page 214, Washington, DC, USA, 2004. IEEE Computer Society.
- [23] W. A. Voglozin, G. Raschia, L. Ughetto, and N. Mouaddib. Querying a summary of database. *Journal of Intelligent Information Systems*, 26(1):59–73, 2006.
- [24] L. Wischhof, A. Ebner, and H. Rohling. Information dissemination in self-organizing intervehicle networks. *IEEE Transactions on Intelligent Transportation Systems*, 6(1):90–101, March 2005.
- [25] B. Xu, A. M. Ouksel, and O. Wolfson. Opportunistic resource exchange in inter-vehicle ad-hoc networks. In *5th Int. Conf. on Mobile Data Management*, 2004.
- [26] D. Zhang, D. Gunopulos, V. J. Tsotras, and B. Seeger. Temporal aggregation over data streams using multiple granularities. In *8th International Conference on Extending Database Technology (EDBT'02)*, pages 646–663, London, UK, 2002. Springer-Verlag.
- [27] J. Zhao and G. Cao. VADD: Vehicle-assisted data delivery in vehicular ad hoc networks. *IEEE Transactions on Vehicular Networks*, To appear, 2008.