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

José-Luis Zechinelli-Martini
CENTIA, University of las Américas
LAFMIA UMI CNRS 3175
Exhacienda Sta. Catarina Mártir s/n
72820 Cholula - 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) [1] is a system for enabling vehicles to share information in vehicular ad-hoc networks (VANETs). The originality of VESPA is to process and disseminate 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 events to potentially interested vehicles and evaluating their relevance once received in order to determine, for instance, whether the driver should be warned or not.

This paper concentrates on knowledge extraction in VESPA. In particular, it focusses on how to exploit data exchanged among vehicles to produce knowledge to be used later on by drivers. Existing systems only use exchanged data to produce warnings for drivers when needed. Then, data is considered obsolete and is deleted. In contrast, we propose to aggregate 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 neighboring vehicles. For example, by aggregating events it is possible to dynamically detect potentially dangerous road segments or to determine the areas where the probability to find an available parking space is high.

[1]See: `http://www.univ-valenciennes.fr/ROI/SID/tdelot/vespa/` for more information.

General Terms

VANETs, data storage, data aggregation

1. INTRODUCTION

Today, the car is indisputably the most popular mode of transportation. Unfortunately, its popularity leads to numerous issues concerning, for example, safety and the environment. In spite of significant efforts to reduce the number of people dying on the road, this number remains quite high, mainly due to human factors (e.g., accident-prone behavior 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 work concentrates 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 [12] and provide bandwidth in the range of Mbps. Using such communication networks, a car driver can receive information from his/her neighbors. 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 for finding available parking spaces, avoiding traffic congestions, being aware of real-time traffic conditions, etc.

In the last years, research works have investigated on data dissemination and data relevance evaluation in inter-vehicle networks. This paper presents our approach that considers both fresh data for warning drivers but that also maintains and aggregates data histories for disseminating knowledge among vehicles. Indeed, once data is considered no relevant or obsolete, for example because it has been already used to warn the driver, we propose to aggregate it to produce additional knowledge to be used later on. This approach enables for instance the real-time detection of potentially dangerous areas on the roads due to bad weather conditions; or the identification of those places where there is a high probability to find parkings with available places. Such knowledge can be determined using previously received warnings even if no new event has arrived during a period of time.

Two approaches can be used to aggregate events: centralized and decentralized. 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 contrast, in a decentralized approach, vehicles construct their own aggregates. T Even if these aggregates are less precise than in centralized approaches they are exactly those needed by the vehicle. Vehicles do not require to have good connectivity for constructing or querying summaries. This paper presents a fully decentralized approach improved by aggregate exchanges between vehicles. Accordingly, the rest of this paper is organized as follows. Section 2 presents some related works. Section 3 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) [9], followed later by the NoW project[2] (finishing at the end of 2007), and CarTalk [13, 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 [10], CarTalk used no existing infrastructure. Those are very interesting projects which mainly focus on the network level of inter-vehicle communication.

Regarding data dissemination protocols, there are several projects that are worth mentioning. The *TrafficView* system [15], proposes dissemination protocols for broadcasting data periodically. In [1, 2], study the way to determine the *dissemination area*, which is the area where the data should be broadcast. [26] proposes a *carry and forward* strategy for keeping information stored in the car until it can be transmitted to another one. Finally, the *Mobi-Dik* project [24], considers the spatio-temporal relevance of data. Thus, a vehicle with a certain piece of information acts as a disease carrier, and "contaminates" the nearby vehicles along its route; a similar approach is proposed in [16].

[2]http://www.network-on-wheels.de/

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 to disseminate information relative to events (accident, obstacle, etc.) or resources (available parking slots, etc.) to other vehicles. Existing solutions are specific to a particular type of event (accidents, available parking slots, etc.), adapted to an specific application type, 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 must be modified if they should 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 in cars. Therefore, VESPA uses an encounter probability to determine the relevance of an event [5, 7].

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, 19]), 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:

- [20], proposes the *Region-based Location Service Management Protocol* (*RLSMP*), 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).
- [18] focusses 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 issues 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. Two types of aggregation are proposed. *Syntactic aggregation* which mainly reduces the overhead of message headers. *Semantic aggregation* which is information-specific and saves more bandwidth at the cost of loss of information. It The solution presented in the paper, however, cannot be used 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.

- [14] distinguishes 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.
- [8] considers vehicles which aggregate *warning data* if they have several messages describing the same event. It additionally proposes 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.
- [11] studies hierarchical data aggregation. 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 [21] as a means to avoid the *distinct counting problem* for count queries and sum queries.
- [6] proposes a *Location Based Aggregation* (*LBAG*) protocol. In this protocol, 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).

[23, 4] are works that consider aggregation in the context of vehicular networks exist. Also, [25] proposes approaches to manage temporal aggregates with different time granularities. [17] proposes an approach to count the (approximate) number of distinct trajectories within a specific spatio-temporal range, based on the computation of an aggregate function, and compared with the sketch-based approach in [21]. Finally, summaries have also been considered in the context of databases, to handle growth and keep high-level knowledge (e.g., see [22]).

3. EVENTS AGGREGATION

The main existing solutions related to inter-vehicle communications focus on information exchanges among 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 neighboring 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 describing an event should not only be used to produce a warning for the drivers. Once stored in 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 neighboring 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 <*timestamp, type of event, location*>. 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, neither in terms of computational cost nor of memory usage.

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 V. Each *spatio-temporal cell* of that matrix contains a number of events and a confidence value ranging between 0 and 1 (see Figure 1).

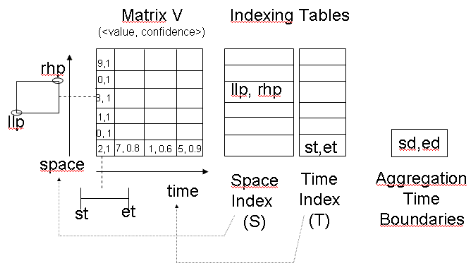

Figure 1: Storage structure of the events

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 S 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.

Regarding 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 T associated with the type of event defines the chosen temporal representation. It contains the starting time (st) and the ending time (et). For each matrix, the starting date (sd) and the ending date (ed) of the aggregation is also kept.

For each cell of V, a confidence value indicates whether it is guaranteed that the events have really been observed in that cell or, on the contrary, the information is 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 neighboring 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 matrix V 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 center, 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. Some of the confidence values in that row may 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 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, which corresponds to the spatial dimension of c_1, one rectangle at the right, one at the left, one below and one underneath) and up to 3 for the temporal one (the common interval, the preceding interval and the following interval). The aggregated value of c_2 is then distributed between all the created cells, since the aggregation process leads to the loss of precise information. Therefore, the value and confidence of the shared rectangle, which we call sub_{c_2}, is given by:

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

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 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)
It 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 Exchanging summaries among 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 there is a match, then the exchange takes place. Thus, vehicle A sends vehicle B a matrix m_A. The receiving vehicle B will need to fusion its previous matrix m_B with the receiving matrix m_A. For this, the following methods are used:

procedure fusionmatrix(m_1, m_2)
m1 is the target matrix in the fusion process. All the cells of matrix m1 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(c_1, c_2) returns c_3
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[3]:

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

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

In order to avoid exchanging summaries which have not been updated since the last exchange, timestamps are used. 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 $LT(V2, e)$ with $ET(V_1, V_2, e)$.

The amount of data exchanged for the summaries is signifi-

[3]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)]*.

cant compared with the exchange of 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.

4. CONCLUSION & PERSPECTIVES

We presented the main principle of our approach based on data aggregation in VANETs to generate additional knowledge for drivers. Such an approach allows to exchange summaries among vehicles in addition to information about traditional events (emergency braking, accidents, etc.). Our current 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*, 2004.

[4] M. Caliskan, D. Graupner, and M. Mauve. Decentralized discovery of free parking places. In *3rd Int. Workshop on Vehicular ad hoc networks*, pages 30–39, 2006.

[5] N. Cenerario, T. Delot, and S. Ilarri. Dissemination of information in inter-vehicle ad hoc networks. In *IEEE Intelligent Vehicles Symposium*, 2008.

[6] C. Chen. Location-based data aggregation in mobile ad hoc networks. Master's thesis, Institut für Parallele und Verteilte Systeme, 2003.

[7] T. Delot, N. Cenerario, and S. Ilarri. Estimating the relevance of information in inter-vehicle ad hoc networks. In *IEEE Int. Conf. on Mobile Data Management – Workshops*, 2008.

[8] S. Eichler, C. Merkle, and M. Strassberger. Data aggregation system for distributing inter-vehicle warning messages. In *31st IEEE Conf. on Local Computer Networks)*, pages 543–544, November 2006.

[9] A. Festag, H. FüSSler, H. Hartenstein, A. Sarma, and R. Schmitz. Fleetnet: Bringing car-to-car communication into the realworld. In *9th ITS World Congress on Intelligent Transport Systems*, 2002.

[10] 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.

[11] C. Lochert, B. Scheuermann, and M. Mauve. Probabilistic aggregation for data dissemination in VANETs. In *4th ACM Int. Workshop on Vehicular Ad Hoc Networks*, pages 1–8, 2007.

[12] 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*, 2005.

[13] 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.

[14] 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.

[15] 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.

[16] S. Nittel, M. Duckham, and L. Kulik. Information dissemination in mobile ad-hoc geosensor networks. In *3rd Int. Conf. on Geographic Information Science*, 2004.

[17] S. Orlando, R. Orsini, A. Raffaetà, A. Roncato, and C. Silvestri. Spatio-temporal aggregations in trajectory data warehouses. In *9th Int. Conf. on Data Warehousing and Knowledge Discovery*, pages 66–77, 2007.

[18] F. Picconi, N. Ravi, M. Gruteser, and L. Iftode. Probabilistic validation of aggregated data in vehicular ad-hoc networks. In *3rd Int. Workshop on Vehicular Ad Hoc Networks*, pages 76–85, 2006.

[19] R. Rajagopalan and P. Varshney. Data-aggregation techniques in sensor networks: a survey. *IEEE Communications Surveys & Tutorials*, 8(4):48–63, 2006.

[20] H. Saleet and O. Basir. Location-based message aggregation in vehicular ad hoc networks. In *2007 IEEE Global Communications Conference Workshops*, pages 1–7, November 2007.

[21] Y. Tao, G. Kollios, J. Considine, F. Li, and D. Papadias. Spatio-temporal aggregation using sketches. In *20th Int. Conf. on Data Engineering*, page 214, 2004.

[22] 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.

[23] 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.

[24] 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.

[25] D. Zhang, D. Gunopulos, V. J. Tsotras, and B. Seeger. Temporal aggregation over data streams using multiple granularities. In *8th Int. Conf. on Extending Database Technology*, pages 646–663, 2002.

[26] J. Zhao and G. Cao. VADD: Vehicle-assisted data delivery in vehicular ad hoc networks. *IEEE Transactions on Vehicular Networks*, To appear, 2008.